Commercial DNA tests and police investigations: a broad bioethical perspective

Nina F de Groot,[1] Britta C van Beers,[2] Gerben Meynen[1,3]

[1]Philosophy, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
[2]Legal Theory and Legal History, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
[3]Willem Pompe Institute for Criminal Law and Criminology, Utrecht University, Utrecht, The Netherlands

Correspondence to
Nina F de Groot, Philosophy, Vrije Universiteit Amsterdam, Amsterdam 1081 HV, The Netherlands; n.f.de.groot@vu.nl

Check for updates

ABSTRACT

Over 30 million people worldwide have taken a commercial at-home DNA test, because they were interested in their genetic ancestry, disease predisposition or inherited traits. Yet, these consumer DNA data are also increasingly used for a very different purpose: to identify suspects in criminal investigations. By matching a suspect's DNA with DNA from a suspect's distant relatives who have taken a commercial at-home DNA test, law enforcement can zero in on a perpetrator. Such forensic use of consumer DNA data has been performed in over 200 criminal investigations. However, this practice of so-called investigative genetic genealogy (IGG) raises ethical concerns. In this paper, we aim to broaden the bioethical analysis on IGG by showing the limitations of an individual-based model. We discuss two concerns central in the debate: privacy and informed consent. However, we argue that IGG raises pressing ethical concerns that extend beyond these individual-focused issues. The very nature of the genetic information entails that relatives may also be affected by the individual customer's choices. In this respect, we explore to what extent the ethical approach in the biomedical genetic context on consent and consequences for relatives can be helpful for the debate on IGG. We argue that an individual-based model has significant limitations in an IGG context. The ethical debate is further complicated by the international, transgenerational and commercial nature of IGG. We conclude that IGG should not only be approached as an individual but also—and perhaps primarily—as a collective issue.

INTRODUCTION

In 1988, a newborn boy, wrapped in blankets, was left under a tree in a Connecticut parking lot. He later died of the frigid temperatures outside. Police did not succeed in finding who had left the infant there. More than three decades later, in 2020, police knocked on the door of a woman who they had identified as the mother of the child. The woman admitted that she had left the boy there after hiding her pregnancy and delivery from friends and family.

The unusual breakthrough in this very old case came as a result of comparing the DNA profile of the infant with DNA data of people who had taken a genetic test with a commercial company and uploaded it into the genealogical website GEDmatch.[1] Through corporations such as Ancestry.com, 23andMe and MyHeritage, customers hope to find information about their ancestry, distant relatives or risks for certain diseases. The genetic testing kits are even popular Christmas gifts. Already around 30 million people have taken a direct-to-consumer (DTC) genetic ancestry test.[2]

These DTC genetic databases are, thus, not only interesting for people wanting to find out about their distant relatives or their genetic risk for 'coriander taste aversion' but could also be helpful to law enforcement to identify a suspect. Such a search takes place by uploading a genetic profile from a suspect, found at the crime scene, in either the genealogical website GEDmatch or the DTC company FamilyTreeDNA to search for distant relatives.[i] Through this distant relative, an expert genealogist examines the intersections between the family trees of the relative and the crime scene DNA. The group of individuals identified can then be narrowed down by using information such as the age and gender of the suspect, or where this person was living at the time of the crime until one can eventually zero in on a single individual or a number of siblings.[3] Depending on how complex and distant found relations are, finding a suspect can take less than 2 hours[4] but may also take up to 6 months (as was the case in a Canadian search for a murder suspect).[5]

In the past few years, this investigative genetic genealogy (IGG) has been used in over 200 law enforcement investigations in the USA.[6] It can be used not only to identify perpetrators but also to identify crime victims or unidentified human remains. One of the most high-profile cases is the conviction of the Golden State Killer, a serial rapist and killer who committed his crimes in the 1970s and 1980s. There is an increasing interest in IGG also outside the USA. In 2020, IGG was performed in a Swedish crime investigation, which led to the arrest and confession of a double-murder suspect.[7] This was probably the first time in Europe that IGG led to an arrest. Earlier, also in Sweden, investigators had found investigative leads to help identify human remains of an unknown individual by using GEDmatch.[8] In the Netherlands, a pilot is currently being set up with this method to identify unidentified human remains of people whose death is not caused by a crime. In July 2020, it was revealed that in the Philippines, IGG had been used to successfully track down several men, most likely sex tourists, who had fathered children with local women during their visit to the region, and that IGG will possibly also be used to identify aid workers who sexually abuse minors.[9] In October 2020, police

---

[i]FamilyTreeDNA and GEDmatch are, thus, two different types, but in our ethical analysis, we discuss them together as the concerns apply to both types. The difference is mostly relevant from a legal perspective, which lies outside the scope of our paper.

forces in Canada revealed for the first time that IGG had helped them to solve a 1984 murder cold case.[10]

The potential impact of such forensic use of DTC databases for crime investigations is enormous. For example, it has been estimated that a genetic database consisting of only 2% of all US citizens of European descent will be able to find a third-degree cousin in more than 99% of all US persons of that ethnicity.[11] As the DNA of millions of people is in these databases, the chance of finding a distant relative will possibly become almost certain. If an increasing amount of these DNA data would be accessible for law enforcement, it can have major implications for criminal investigations. Although currently, most consumers of DTC companies are US citizens, the kits are also increasingly popular in other countries around the world. Therefore, it is likely that IGG will become more and more effective in identifying suspects outside of the USA. For example, a UK study was able to identify 4 out of 10 anonymous volunteers by uploading their genealogical DNA profile to GEDmatch.[12]

Since 2018, in the aftermath of the Golden State Killer arrest, IGG has received attention in the academic debate as well, especially from genetic genealogy and/or legal perspectives, that often also cover ethical concerns.[4 13–21] Regulation of IGG, at the moment, is largely absent and several authors have called or made suggestions for regulatory oversight for IGG.[17 18 20 22–26] Although there have been interesting discussions on ethical aspects of IGG,[6 13 18 23 27] including justice, public safety, risk for false positives, privacy and consent, a more systematic analysis from a bioethical perspective has not been performed. A recent extensive review study on IGG recommends that for proper regulation, much more research by bioethicists, among others, on the ethical issues is needed.[28]

In this paper, we aim to contribute to the bioethical analysis on IGG by analysing the limitations of an individual customer-based approach to IGG. The paper runs as follows. First, we discuss two central ethical concerns regarding IGG that may be considered autonomy related: privacy and informed consent. Second, we argue that IGG raises pressing ethical concerns that extend beyond these concepts. In that process, we explore to what extent the ethical approach in the medical clinical genetics context can be helpful for the debate on IGG. We argue that an individual-based model has significant limitations in an IGG context. The debate is further complicated by the transgenerational, transnational, and commercial nature of IGG. We conclude that IGG should be approached not only as an individual but also—and perhaps even more so—as a collective issue.

PRIVACY

A first ethical concern regarding IGG is privacy. We will discuss three privacy-related issues of IGG. The first concerns the *sensitivity of information* that may come to light. Notably, there is a fundamental difference between genetic profiles used in the standard forensic setting, and genetic profiles, used in commercial genealogy. Identification of a relative within an IGG context is fundamentally different from relative identification within the standard forensic setting. The vast majority of the human genome is identical for all humans, but there are also places in the DNA that differ from person to person. This includes so-called 'non-coding' regions in the DNA that consists of repeats of DNA units, varying in length across individuals. One type of repeat is called short tandem repeat (STR). A standard forensic DNA profile is established out of the length of these STRs on several places of the DNA, usually around 20 different places.[29] By contrast, a genetic genealogical profile established

by DTC genetic companies encompasses far more different places in the DNA. It specifically looks at one variation of a single nucleotide—the basic building block of DNA—in many places across the entire genome. These variations are called single-nucleotide polymorphisms (SNPs). A commercial genealogical SNP profile encompasses around 600 000 to 700 000 different SNPs. Therefore, these profiles contain much more information compared with forensic STR profiles. STR profiles are established from non-coding regions of the DNA and, therefore, do not entail detailed information about physical appearance, risks for diseases, etc. Indeed, the different places in the DNA that forensic STR profiles represent have been precisely *chosen* because they do not contain such information. Forensic STR profiles, thus, contain no information relevant for medical interests and this has been of significance in securing public support for forensic DNA profiling.[30]

Matching forensic STR profiles may be compared with checking whether barcodes on two cans of chili beans match (do they have the same numbers, in the same order?). Apart from knowing whether two chili bean cans are identical to one another, there is not much other information to be retrieved from just looking at the chili bean can. By contrast, a commercial SNP profile contains a wide variety of personal information such as risk for diseases, physical appearances or geographical origin. Access to this information depends, however, on *who* sees the information in the DTC databases and whether and how it can be analysed. In that sense, a SNP profile might be compared with a QR code on the chili bean can: the dots and patterns of the QR code could be decoded into valuable extensive information, but this information cannot be retrieved without having the knowledge or techniques to decipher the code. Therefore, some authors have argued that one of the factors that mitigate privacy concerns regarding IGG is that 'raw genetic data are not disclosed to law enforcement'. and that '(...) only the possible genetic kinship among individuals is shown'.[31] However, it should be noted that it is not clear whether every individual and company that is active in commercial genealogy does comply with this.[ii] In other words, *if* people with expert knowledge could obtain the commercial SNP data, they may be able to draw conclusions far beyond mere kinship. In short, the risk that *sensitive information* comes to light is greater in the IGG context, compared with traditional forensic databases. Furthermore, given the pace of developments in genetic technologies, it might well be that in 5 or 10 years from now, much more sensitive information can be extracted from the commercial SNP profiles.

Related to this is the problem of sensitivity of familial relations that may be exposed. In various countries, familial searching in forensic databases is increasingly conducted, but it is often strictly regulated. With a forensic STR profile, one can search the forensic database to potentially—in an ideal case, because this is often quite hard—identify a parent, child or sibling of the uploaded STR profile.[32] By contrast, SNP profiles used by DTC companies are generally capable of identifying third cousins (having great-great-grandparents in common) or even more distant relatives. Of note, one individual has on average almost 5000 fifth cousins.[13] The chance of finding at least one third cousin in a database of one million people is reportedly already

---

[ii]As single-nucleotide polymorphism profiles are being held by commercial companies, their employees could potentially have access to the raw data.

as high as 90%.[33] The possibility of using commercial SNP profiles to identify both close and distant relatives may bring unwelcome (or welcome) information to light for a family: there could be revelations of misattributed parentage or siblings who turn out to not be siblings at all. An in-depth forensic search in a commercial genealogical database may expose this information. Obviously, exposing this information is also a risk when taking a commercial DNA test. However, the extensive family tree search that is performed by a genetic genealogical expert in IGG is *not* standardly conducted for regular customers. Notably, this family search is usually conducted unbeknownst to the innocent family members; therefore, this information may come to light to the police, but presumably not to the family members themselves. Still, it *could* be disclosed; and this risk is greater in comparison to traditional forensic databases because in the latter the amount of sensitive information that can be retrieved is very limited.

The second privacy-related concern of IGG is technological *security*. Forensic databases often have strict privacy and security regulations. DTC services, however, often do not have that level of security. This entails a greater risk for technological privacy breaches. In fact, two recent studies show that there are multiple security threats to commercial genetic genealogy companies. Ney *et al*[34] point out that GEDmatch has a vast amount of features to get an in-depth view of one's genetic profile and that an adversary can extract a large number of genetic markers of other users, including hundreds of markers that encompass sensitive medical information. Moreover, Edge and Coop[35] show that there are multiple ways 'to reveal users' raw genetic data on the basis of the locations of shared DNA blocks'. This might disclose sensitive information about an uploaded DNA profile of a suspect. Both publications also suggest possible interventions to reduce the security risks. In July 2020, GEDmatch was hacked with the effect that *all* of the approximately one million users, who had opted out of law enforcement access, were now labelled as opted in.[36] Moreover, in January 2021, a data breach caused a restoration of earlier deleted user data for a period of 2 days.[37]

Third, and partially related to the previous point, the *regulatory landscape* is completely different for forensic genetic databases, compared with IGG. For forensic genetic databases, extensive regulations and legal frameworks exist, including guidelines for familial searching within these databases.[17] However, the regulatory framework for DTC companies, including privacy protection, is often opaque. In 2018, a best practice guideline was established by the Future of Privacy Forum, supported by several major DTC companies, including 23andMe and Ancestry, which covers topics like integrity, accountability, transparency, law enforcement access, consent and privacy.[38] However, GEDmatch and FamilyTreeDNA are not involved in the forum.[38] It has been noted that DTC companies' terms of services/privacy policies and law enforcement's self-restricting regulations—as far as they are in place—are currently the only obstacle to forensic access in the USA.[39] Specific legislation on IGG is often absent in Europe, probably because it is a novelty in these countries (although Sweden has allowed it in a pilot setting).[7] Furthermore, a review of 22 DTC companies' policies concluded that all companies communicated to consumers the possibility of law enforcement access or disclosure if required by law, but that only 4 out of 22 companies provided information about *how* law enforcement should request access to their databases.[40] It has been argued that familial searching in DTC databases provides a possibility to *circumvent* the existing longstanding regulations for forensic DNA databases.[32] According to a recent report on IGG by a UK advisory committee, '(...) the

whole process is unregulated'.[22] Of note, it has been pointed out that no official accreditation for genetic genealogists exists and that 'this lack of professionalization and accountability increases the risk of ethical and privacy breaches'.[20] Moreover, third-party genetic interpretation services (companies who help customers to interpret their raw genetic data obtained from DTC companies) have recently emerged.[41] These novel services could further complicate the regulation of IGG as law enforcement might in the future also access raw genetic data through these companies.

In sum, IGG creates relevant concerns regarding privacy that differ in important ways from the privacy protection in forensic genetic databases. Still, one could argue that if strict privacy protections and regulations were set in place, many of the privacy concerns could be partially mitigated. However, even if this were—or already largely is, as some have argued[31]—realised, a perhaps more pressing issue of IGG remains: consent.

INFORMED CONSENT

Concerns about consent and IGG are frequently raised.[4 6 13 20 23 27 42 43] The discussion often revolves around the question as to whether or not consumers have validly consented to the use of their genetic data for forensic purposes. We will discuss three concerns about informed consent in an IGG context. Subsequently, we will explore the limitations of consent in this context.

The first concern regarding informed consent and IGG is that there is reason to doubt whether customers of commercial genetic databases are *well informed* about law enforcement access. According to Murphy,[17] the question is whether the seeming public acceptance of IGG is the result of having insufficient information. If we do not know how often the police accesses these databases, under which conditions and regulations they are allowed to do so, and how many innocent people thereby become subjects of extensive police investigations, can we really make an informed decision about the issue? Indeed, there is reason to believe that the information provided to customers about the possibility of IGG often falls short.[4] One study shows that 38 out of 55 studied DTC genetic companies provided information about governmental or law enforcement access, but that the 'majority of companies simply stated that personal information may be disclosed "as required by law" without further explanation'.[44] The commercial genealogy company FamilyTreeDNA, which stores an estimated two million profiles, states the following on its website's homepage, next to an image of a vault: 'We won't share your DNA. We believe your DNA belongs to YOU and only you... period'.[43] However, FamilyTreeDNA has adopted an opt-out policy for law enforcement access.[43] This means that the 'default' option for consumers is that their sample can be used by law enforcement—unless one opts out (EU users prior to March 2019 were automatically opted out).[43] All users have to actively go to their privacy settings and unclick the box to participate in law enforcement access.[43] Over 96% of US-based FamilyTreeDNA users are opted in to law enforcement access.[21] In contrast to FamilyTreeDNA, GEDmatch had abandoned its opt-out policy for law enforcement access and required customers to opt in. This was the result of a controversy following the access of the database in a case of severe assault, which lead to the arrest of a 17-year-old.[27] Only around 280 000 of the 1.45 million GEDmatch users chose to opt in for these law enforcement searches.[45] However, in January 2021, GEDmatch changed its terms of service, opting in the *entire* database for law enforcement matching for unidentified human remains (users still have to opt in for seeking suspects of violent crimes).[37]

Furthermore, it can be questioned whether consumers actually *read* the conditions of DTC services at all.[23 43] DTC companies often make use of so-called 'wrap contracts', where consumers use the service without seeing or reading the (often very elaborate) terms and privacy statements.[46] One survey study of DTC consumers found that only under half of the participants reported they had read the privacy policies and that customer's expectations were often not consistent with the companies' policies.[47] These wrap contracts are far removed from the standard of informed consent that is employed when genetic data collection and analysis take place in a traditional *medical* setting. Within a clinical medical context, informed consent needs to be obtained from an individual before a genetic test is performed. Obtaining informed consent in accordance with high professional standards is widely regarded as an essential requirement that follows from the bioethical principle of respect for *autonomy*. Also, in DTC genetic testing, informed consent is crucial as it protects consumers against harm and enables autonomous decision-making.[48] However, at present, the regulations and legal requirements that those in the medical setting have to comply with, do not apply to DTC companies.

A second issue regarding IGG and informed consent concerns the problem of adherence. Even when a customer *has* made an informed decision about whether or not to consent to law enforcement access, it can be questioned whether this decision will also be adhered to. Laestadius *et al*[49] reported that 25 of the 30 companies (83.3%) that they examined specifically stated in their terms that they may be required to disclose data to legal authorities *without* the consent of the customer. Indeed, in the case of FamilyTreeDNA, even when a customer has opted out, this is not a guarantee that their profile will not be used by law enforcement. The company's privacy statement reads: 'Please note, if you decline to participate in law enforcement matching, we may still be required to share your Personal Information to comply with a valid legal process (…)'.[50] Thus, while it has been argued concerning IGG that 'only data voluntarily uploaded and explicitly made public are searched',[31] this argument does not hold in all cases; also without one's permission, one's data could, theoretically, be accessed. Meanwhile, two major companies, Ancestry and 23andMe, have stated that they resist law enforcement access, although they acknowledge that they may be required by law to comply with a legal order.[43] Notably, in 2019, Florida police obtained a warrant for a particular case to search the entire GEDmatch database, thus obtaining access to the data of *all* the database's users.[51] By contrast, the problem of adherence is less problematic in the *clinical* setting; here, personal health information falls under extensive legal protection. In the rare cases that law enforcement attempted to obtain DNA from biological materials kept within a medical setting, this has been declined by court rulings multiple times.[52]

A third concern regards the possibility of change: even if DTC companies have accurate and well-informed opt-in policies *now*, it could well be that they will change these in the future. There is often a possibility of *unilateral change* of terms by DTC companies without any notification.[46] An example is the previously discussed GEDmatch change in terms, without informing customers of the implications. In other words, customers might have given consent at a certain point in time based on the terms that were applicable at that moment, but this may be changed a month or a year later, without them ever knowing this. Such a change is virtually impossible in a medical setting, where strict regulations are in place. In sum, there are currently significant concerns with respect to valid (informed) consent for law enforcement access.

It has been argued that consent procedures for DTC services could be significantly improved and that some concerns may, thus, be partially overcome,[4] even though it is not yet clear how exactly this can be achieved or enforced. Yet, *even if* companies offered clear, sustainable and transparent information, and even if consumers expressly gave valid consent for police usage, an important issue remains. When an individual gives informed consent to the use of her profile for law enforcement access, this individual decision might not only entail consequences for herself but *also for her close and distant relatives*.[18] Thus, someone who has never taken a DNA test could become part of a criminal investigation, just because a distant relative happened to have her DNA tested and consented to police access.[20] Even if the police does not have any genetic information about the people that have been identified as part of the suspect's family tree (like their risks for diseases), an individual might still be part of the chain that leads to the suspect. It could be argued that an individual is, thus, consenting also 'on behalf of' her distant family members.[42] This is a vital point in the debate on IGG; individual consent has implications for relatives. In the next section, we will analyse this issue further and argue for a more collective ethical approach to IGG.

CONSENT AND RELATEDNESS IN THE BIOMEDICAL CONTEXT

The genetic relatedness between individuals complicates the issue of individual consent in the IGG context. In the ethical debate about genetic testing in a *healthcare* setting, consequences for relatives of the consenting person are a central issue. Therefore, for our debate on IGG, it may be helpful to briefly look at genetic testing in a healthcare context to examine how individual informed consent and its consequences for relatives are approached in this context.

In a clinical setting, a patient might receive information about having a hereditary type of breast cancer, which can also have consequences for her relatives (eg, that relatives also have an increased genetic risk to develop breast cancer). An ethical question that arises is what should be done when a patient refuses to share this information with her relatives. Should we approach the issue from the viewpoint that it entails 'personal information' with the responsibility lying with the patient to disclose relevant information to relatives? Parker and Lucassen[53] have suggested a 'joint account model', in which the physician shares genetic information with other family members, similar to how a bank manager shares information about a bank account with the other account holders. The justification of this joint account model lies primarily in the bioethical principle of justice: as genetic information is *inevitably familial,* it seems unjust to let one family member benefit from genetic information, while this is being denied for other family members.[53] Furthermore, overall, the regular sharing of genetic information might result in more beneficial and fewer harmful outcomes. A sister of the patient might, for example, receive timely breast cancer screening.

Could this joint account model also be relevant for IGG? That does not seem to be the case, at least not straightforwardly: investigating the family tree of a woman X for a criminal investigation will not directly result in potential beneficial outcomes for her relatives, but rather in *harmful* consequences, they may become subject to privacy infringements (discussed in the previous section), surveillance, a criminal investigation or prosecution—and it may turn out that they are completely innocent. In healthcare, no patient is prosecuted or punished. Moreover, the choice of woman X to know her risk for a hereditary type of cancer may have a beneficial result for this individual (eg, her life

is saved because she can act on this information), but the choice of woman X to participate in law enforcement access does in general not result in a beneficial result for herself *personally*.[iii] Of course, there could be outcomes that can be regarded as beneficial: successful crime solving, improved safety of society, punishment of a perpetrator. However, these will be primarily societal benefits on the *collective* level, and *not*—as is the case with clinical genetics—direct benefits for the individual person and her family members. In short, IGG differs from the clinical genetic testing context in terms of both the actors (physicians vs law enforcement) as well as the goals (individual vs collective) and, therefore, the existing accounts in the clinical context are difficult to transpose to the IGG setting.

Multiple authors have written about informed consent and/ or the obligation to inform relatives specifically in a DTC genetic context. Parsons and Baker[54] argue that when a particular disease risk to a relative of a DTC consumer is actionable, it should be disclosed to this relative by the DTC genetic test provider even if the consumer objects to this. Loi[55] has provided a libertarian argument for the right to take a DTC test, based on the moral right to self-ownership, which argues that individuals are 'sovereign over all uses of their own genome that do not harm others'. Bunnik et al[56] argue that for DTC genetic testing, informed consent is *not* required for non-medical entertainment testing, but it is required for genetic tests that might cause harm, such as knowing about an untreatable disease. Although all these accounts vary from one another, they have one element in common: the accounts depart from possible harms and benefits for the individual or others (such as an actionable, treatable disease risk to a relative). However, these accounts primarily look at the medical and/or individual beneficences of DTC testing, *not* possible forensic use.

Thus, the accounts on informed consent and the consequences for relatives in a healthcare or DTC setting differ significantly from the IGG context, because IGG is not about potential beneficence to the DTC consumer and her relatives, but about criminal investigation and possibly prosecution, which may have severe consequences for these family members. Moreover, they are based on the consequences of more close relatives and *not* on very distant relatives as in the case of IGG. Therefore, these accounts are not directly helpful to resolve the concerns surrounding informed consent and IGG. This is not to argue that the accounts cannot be helpful at all in the context of IGG, but that further arguments are required to show their usefulness.

A COLLECTIVE PERSPECTIVE
So how should we approach IGG? A recent article on the misconceptions of IGG states that 'genetic genealogy database participants are themselves the ultimate check on the indiscriminate conduct of IGG'.[21] In the following, we will argue that this check is insufficient; ethical decision-making on IGG should *not* primarily be left to individual consumers.

It can be argued that to use the term 'informed consent' when people do *not* have the possibility to opt out is inappropriate.[57] As discussed, even if consumers in the case of IGG have the right to opt out, their *relatives* do not. If we suppose a (perhaps utopian) situation in which consent for police access is given by consumers, in a way that miraculously meets all

the criteria for informed consent, the problem still exists that large groups of people who have *not consented* can be affected by it (most likely without their knowledge; a relative could be a 'key' that is needed to build the rest of the family of the suspect).

As already mentioned, it has been estimated that a genetic genealogy database has to include only 2% of the US population of European descent to find a third-degree cousin of almost anyone in that population.[11] Thus, when only a fraction of a population would give their voluntary and well-informed consent to police accessing their DTC genetic data, the investigation would also encompass the family trees of all the non-consenting people in that population. With a genealogical search in a DTC database, there are often hundreds of relatives who need to be examined.[12] A study by Erlich et al[11] found that a genealogical search would have initially 850 hits, which can eventually be narrowed down— on the basis of geographical location, age and sex—to a list of 16 or 17 possible suspects. This group of possible suspects can then be further limited by traditional detective work. Another study found an initial list of 3000 individuals.[12] In one murder investigation, police used a false story about the finding of an unidentified body in order to retrieve DNA samples from relatives for IGG—the son of one of those 'participating' relatives was subsequently arrested.[58] The genealogical search will, thus, not only involve the suspect but also many innocent people (most likely unbeknownst to them). By contrast, a forensic genetic database will solely affect the suspect or his direct family members (if a familial search is performed).

Still, it has been argued that in a regular criminal investigation, also many innocent people are involved and perhaps even more than in an investigation that uses IGG.[21] That may well be true, as IGG has a high potential to single-out a suspect, especially in cold cases that would otherwise require a large criminal investigation covering many innocent people. However, our point does not involve the number of people targeted in a *single* criminal investigation. Rather, we look at IGG as a whole: if only a small percentage of the population makes their genetic profile accessible for law enforcement access, this could potentially cover the *entire* population. In this respect, the fears expressed of a *de facto* population-wide genetic database might not be far-fetched. In arguing *against* the claim that IGG could result in a *de facto* universal database, Brown[4] emphasises that the databases are not set up by the government, and that 'individuals are freely, voluntarily and enthusiastically populating these non-governmental databases'. However, as discussed, it is questionable that all people are freely and voluntarily permitting law enforcement access, and, second, that *even if they do,* many more—who might be affected by the decision—*have not*.

Some authors have suggested to implement a universal population-wide DNA database (which includes every citizen), because this will result in fewer privacy concerns, compared with both forensic familial testing[59] *and* compared with IGG.[60] Indeed, a survey study of almost 1.600 respondents showed that law enforcement access to a universal database was regarded as being less intrusive compared with accessing a database of a DTC company.[39] Arguments in favour of a universal database are that it is not racially biased, are more effective, have fewer privacy concerns, that potential family ties are not disclosed (and are not needed to chase, because everyone is in the database), and that it does not include other highly sensitive information (because it uses forensic STR instead of genealogical SNP profiles).[59 60] In view of these considerations, it could be argued that even a universal database, already the subject of much debate and generally not regarded as an acceptable option, may be more

---
[iii] Or it could be that a person feels satisfied from knowing they have contributed to the solving of a crime.

appealing than IGG. Or, differently put: IGG may be *even more problematic* than a universal database. Although further arguments are needed to support this claim, the suggestion underlines the extent of ethical problems of IGG.

It has been noted that the concept of informed consent is not adequate in the case of collective groups of people.[57] As Widdows[61] writes, when it comes to the governance of genetic issues, individuals do not only function as independent subject, but also as 'connected selves' because of their genetic relations. Consequently, in such genetic contexts, the focus on the individual, which underlies notions such as informed consent, fails to offer adequate ethical guidance. This critique of the individual choice model seems also quite pertinent in the context of IGG. Thus, an ethical focus on autonomous choice might have significant limitations. Notably, improved informed consent has been presented in the literature as a way of alleviating some ethical concerns of IGG.[4 23] In a recent interview study by Samuel and Kennett,[6] some participants seemed to regard informed consent as an 'ethical panacea'; they were of the opinion that most ethical problems could be mitigated if consent was implemented properly. In reality, it may well be that too much weight is given in the ethical debate on IGG to the concept of informed consent. Notably, Samuel and Kennett state that 'focusing on individual-based consent narrows ethical discussion by shielding other substantive political and societal issues from critical scrutiny'.[6] There is, thus, reason to believe that the value of informed consent as the solution to pressing IGG problems is questionable. Moreover, this may not only apply to the ethical debate but also to the legal one; the UK Biometrics and Forensics Ethics Group recently stated in a report about IGG that the legality of obtaining this kind of sensitive information on the legal basis of informed consent alone, is questionable, considering the EU Law Enforcement Directive and Article 8, protecting the right to private life, of the European Convention on Human Rights.[22]

In sum, we have argued that an individual consent model has significant limitations when it comes to IGG and needs to be complemented with a more collective approach.

BROADENING THE SCOPE
We have discussed two individual-based ethical concerns regarding IGG: privacy and informed consent. However, even if privacy and valid individual decision-making on IGG were adequately protected, not all pressing ethical issues would be resolved. Therefore, we have argued that the ethical debate urgently needs to be extended *beyond* the interests and decisions of the individual customer. This is needed to prevent the debate from primarily revolving around issues such as *who* can see *which* sensitive information or whether participants really did consent. This is not to say that these are not very relevant topics to address, but that we must be careful that they will not turn out to be serving as a smokescreen, hiding the issue that is fundamentally at stake: that IGG is inherently also a collective issue and not only an individual one.

There are other factors on the collective level that further complicate the ethical debate about IGG. These factors cannot be sufficiently addressed when only an individual-based model is used in the ethical analysis of IGG. First, there are consequences on a group level: the effectiveness of IGG in solving a case depends on the number of people who have made their data available for law enforcement access. As most DTC customers are US citizens of European ancestry, IGG is far less likely to identify a suspect that is not part of that ethnic group. Therefore, it has been argued that IGG could potentially reduce the ethnic

imbalance currently seen in governmental forensic databases (which skew minority groups) because the overrepresentation of white people in DTC databases could complement these.[17 25] However, it might be too early to immediately assume that IGG reduces the currently existing ethnic imbalance because IGG is also used to identify unidentified crime *victims*. Therefore, with IGG, an unidentified African American victim will be less likely to be identified than a victim of European descent, all the more so because this minority group is often absent from genealogy records.[62] Of note, the vast majority of solved IGG cases concerns white victims.[63]

Second, there is a *transgenerational* issue that needs to be addressed as well. For instance, the individual decision of a person to participate in law enforcement access may equally have consequences for her children and their children and their children too. Furthermore, future generations might also be affected by today's individual decisions to make one's genetic data accessible to law enforcement.

In addition, there is a clear *transnational* aspect to IGG.[18 20] Family trees with thousands of people often include people from different countries and, thus, different legislations. As such, even if one country may decide, for example, to completely ban IGG, it is questionable that this will be effective as distant family relations do not stop at a country's border. When a US citizen, for example, participates in GEDmatch law enforcement access, her distant relative in Europe might, perhaps unknowingly, be the key that eventually leads to the perpetrator (or is even the perpetrator himself). That this perpetrator may subsequently get the death penalty in the USA,[28] is only one of many ethical issues involved. The transnational nature of IGG poses significant challenges for both ethical and legal governance.

Furthermore, the *commercial* nature of genealogy companies is relevant here. This may entail not only more challenges for regulations but also the risk that commercial interests of companies overshadow ethical concerns. For example, GEDmatch was originally set up as a hobby by two genealogists but was sold in December 2019 to the US forensic sequencing company Verogen.[12] The engagement of 23andMe in a partnership with a pharmaceutical company resulted in a public debate over the use of data from people 'who technically consented but were not truly informed'.[64] Moreover, the genealogical search, including the establishment of family trees, is usually conducted by private firms, such as Parabon and Othram, which offer their services to law enforcement on a commercial basis. By contrast, forensic genetic databases are not set up to produce financial profits but have the security of society as their aim. In sum, the transgenerational, transnational and commercial aspects of the practice might be additional complicating factors to the ethical debate on IGG. They further underline the importance of a broad ethical analysis on the topic.

These are all additional ethical concerns that move beyond the dichotomy of individual privacy and decision-making versus the safety of society. Eventually, it might be the case that public discussion will focus on how to weigh the societal interests of crime solving against the ethical concerns mentioned. Of note, professionals in the criminal justice system and those working in forensic genetics tend to see forensic DNA testing as highly beneficial for crime fighting, but at the same time, they relatively devalue the ethical risks.[65] Several suggestions have been made about when IGG may be allowed, such as in cold cases where other investigative methods have failed,[23] major crimes against the person[24] or in the case of identification of unidentified human remains.[22] Although the vast number of IGG cases currently concern murder or sexual assault and the identification

of unidentified remains of a deceased, there is also an example in which a DTC database was accessed for another type of crime, namely, an assault.[27] It is often emphasised that the aim of genetic genealogy is to produce investigative leads and *not* to convict.[31] Still, in many cases, IGG has led to an essential investigative lead *and* a likely perpetrator. The fact that in the end, a *traditional* forensic STR profile is established and used in court does not diminish IGG's significance.

Clearly, IGG could potentially be very beneficial to solve crimes—which is of great importance to a society. Meanwhile, it has been noted that in the USA (which is the location of the vast majority of IGG cases), a significant proportion of cases solved by genealogical searches could have been solved much earlier if all convicted criminals would have been in the forensic genetic database.[22] In addition to this backlog, familial searching in forensic databases is only very limitedly used in the USA.[28] Therefore, the question arises as to what the remaining additional benefits of IGG would be, if already existing forensic practices were fully used.

Balancing benefits and harms of IGG is highly relevant. Discussing these considerations merely in terms of individual consent and privacy has significant limitations. Therefore, it can be argued that balancing harms and benefits may better be left to a democratic decision-making process than to the individual consent of only a fraction of the population—for example, the 2% mentioned above—deciding to participate in genetic databases accessible by law enforcement.

CONCLUSION

In this paper, we aimed to contribute to the bioethical analysis of IGG by broadening the ethical perspective of IGG beyond the dichotomy of safeguarding privacy and consent versus the safety of society. We discussed two issues related to the individual: privacy and informed consent. Justified concerns about privacy exist. Yet, arguably the most interesting and pressing issue is the fact that individual consent, although surely relevant, may not be sufficient to deal with the concerns about IGG. Existing ethical frameworks for the clinical biomedical context cannot sufficiently address these concerns. The choice of an individual consumer could have important consequences for many close and distant relatives. Moreover, the decision of only a fraction of the population could have far-reaching consequences for the population as a whole. This has to do with the reach of this technology where DNA from the few may have consequences for the many, making it a collective issue. In other words, discussing IGG in terms of consent has significant limitations. The debate is further complicated by other factors extending beyond the individual, including the transgenerational, transnational and commercial nature of IGG. We conclude that there is an urgent need to further develop a broad bioethical perspective for IGG that extends beyond individual customer-based approaches: IGG is also—and perhaps primarily—a collective issue.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement There are no data in this work.

REFERENCES

1 Iati M. Three decades after she left her baby for dead, police knocked on her door. She had been waiting. *The Washington Post* 17 Jan 2020
2 Greely HT. The future of DTC genomics and the law. *J Law Med Ethics* 2020;48(1):151–60.
3 Greytak EM, Moore C, Armentrout SL. Genetic genealogy for cold case and active investigations. *Forensic Sci Int* 2019;299:103–13.
4 Brown TR. Why we fear genetic informants: using genetic genealogy to catch serial killers. *Columbia Sci Technol Law Rev* 2019;21(1):1.
5 White P. Genetic genealogy helps Toronto police crack landmark 1984 Christine Jessop cold case in Ontario. *The Globe and Mail*, 2020. Available: https://www.theglobeandmail.com/canada/article-genetic-genealogy-helps-toronto-police-crack-landmark-1984-christine/
6 Samuel G, Kennett D. *Problematizing consent: searching genetic genealogy databases for law enforcement purposes*. New Genetics and Society, 2020: 1–21.
7 Tillmar A, Fagerholm SA, Staaf J, *et al*. Getting the conclusive lead with investigative genetic genealogy - A successful case study of a 16 year old double murder in Sweden. *Forensic Sci Int Genet* 2021;53:102525.
8 Tillmar A, Sjölund P, Lundqvist B, *et al*. Whole-genome sequencing of human remains to enable genealogy DNA database searches - A case report. *Forensic Sci Int Genet* 2020;46:102233.
9 Batha E. Could genealogy websites help catch aid worker sex abusers? : Reuters, 2020. Available: https://news.trust.org/item/20200716031354-blleb/ [Accessed 1 Oct 2020].
10 Creet J. Christine Jessop's killer identified: Solved cold case raises questions about genetic privacy. *The Conversation*, 2020. Available: https://theconversation.com/christine-jessops-killer-identified-solved-cold-case-raises-questions-about-genetic-privacy-148266
11 Erlich Y, Shor T, Pe'er I, *et al*. Identity inference of genomic data using long-range familial searches. *Science* 2018;362(6415):690–4.
12 Thomson J, Clayton T, Cleary J, *et al*. An empirical investigation into the effectiveness of genetic genealogy to identify individuals in the UK. *Forensic Sci Int Genet* 2020;46:102263.
13 Syndercombe Court D. Forensic genealogy: some serious concerns. *Forensic Sci Int Genet* 2018;36:203–4.
14 Phillips C. The golden state killer investigation and the nascent field of forensic genealogy. *Forensic Sci Int Genet* 2018;36:186–8.
15 Ram N, Guerrini CJ, McGuire AL. Genealogy databases and the future of criminal investigation. *Science* 2018;360(6393):1078–9.
16 Ram N, Roberts JL. Forensic genealogy and the power of defaults. *Nat Biotechnol* 2019;37(7):707–8.
17 Murphy E. Law and policy oversight of familial searches in recreational genealogy databases. *Forensic Sci Int* 2018;292:e5–9.
18 Scudder N, McNevin D, Kelty SF, *et al*. Policy and regulatory implications of the new frontier of forensic genomics: direct-to-consumer genetic data and genealogy records. *Current Issues in Criminal Justice* 2019;31(2):194–216.
19 Meulenbroek L, Aben D. *Een hooiberg vol spelden - het gebruik van genealogische DNA-databanken bij opsporing en identificatie*. Uitgeverij Paris, 2019.
20 Kennett D. Using genetic genealogy databases in missing persons cases and to develop suspect leads in violent crimes. *Forensic Sci Int* 2019;301:107–17.
21 Guerrini CJ, Wickenheiser RA, Bettinger B, *et al*. Four misconceptions about investigative genetic genealogy. *J Law Biosci* 2021;8(1).
22 UK Biometrics and Forensics Ethics Group. *Should we be making use of genetic genealogy to assist in solving crime? A report on the feasibility of such methods in the UK*. Biometrics and Forensics Ethics Group, 2020.
23 Berkman BE, Miller WK, Grady C. Is it ethical to use genealogy data to solve crimes? *Ann Intern Med* 2018;169(5):333–4.
24 Wickenheiser RA. Forensic genealogy, bioethics and the golden state killer case. *Forensic Sci Int Synergy* 2019;1:114–25.
25 Hendricks-Sturrup RM, Prince AER, Lu CY. Direct-To-Consumer genetic testing and potential loopholes in protecting consumer privacy and Nondiscrimination. *JAMA* 2019;321(19):1869–70.
26 Scudder N, Daniel R, Raymond J, *et al*. Operationalising forensic genetic genealogy in an Australian context. *Forensic Sci Int* 2020;316:110543.
27 Katsanis SH. Pedigrees and perpetrators: uses of DNA and genealogy in forensic investigations. *Annu Rev Genomics Hum Genet* 2020;21:535–64.
28 Kling D, Phillips C, Kennett D, *et al*. *Investigative genetic genealogy: current methods, knowledge and practice*. 102474. Forensic Science International: Genetics, 2021.
29 Butler JM, Willis S. Interpol review of forensic biology and forensic DNA typing 2016-2019. *Forensic Sci Int Synergy* 2020;2.

30 Williams R, Wienroth M. Social and ethical aspects of forensic genetics: a critical review. *Forensic Sci Rev* 2017;29(2).
31 Greytak EM, Kaye DH, Budowle B, *et al*. Privacy and genetic genealogy data. *Science* 2018;361(6405):857.
32 Machado H, Granja R. *Emerging DNA technologies and stigmatization. forensic genetics in the governance of crime*. Singapore: Springer Singapore, 2020: 85–104.
33 Khan R, Mittelman D. Consumer genomics will change your life, whether you get tested or not. *Genome Biol* 2018;19(1):120.
34 Ney P, Ceze L, Kohno T. *Genotype extraction and false relative attacks: security risks to third-party genetic genealogy services beyond identity inference*. San Diego, USA: Network and Distributed System Security Symposium (NDSS), 2020.
35 Edge MD, Coop G. Attacks on genetic privacy via uploads to genealogical databases. *Elife* 2020;9:e51810.
36 Murphy H. Why a Data Breach at a Genealogy Site Has Privacy Experts Worried. *New York Times* 1 Aug 2020, 2020. Available: https://www.nytimes.com/2020/08/01/technology/gedmatch-breach-privacy.html [Accessed 8 Sep 2020].
37 Mullin E. 'Deleted' DNA Data Just Reappeared on a Popular Database: Future Human, 2021. Available: https://futurehuman.medium.com/deleted-dna-data-just-reappeared-on-a-popular-database-e1f43587f7ec
38 Majumder MA, Guerrini CJ, McGuire AL. Direct-To-Consumer genetic testing: value and risk. *Annu Rev Med* 2021;72(1):151–66.
39 Slobogin C, Hazel J. 'A World of Difference?': Law Enforcement, Genetic Data and the Fourth Amendment. *Duke Law Journal* 2020;70.
40 Skeva S, Larmuseau MH, Shabani M. Review of policies of companies and databases regarding access to customers' genealogy data for law enforcement purposes. *Per Med* 2020;17(2):141–53.
41 Guerrini CJ, Wagner JK, Nelson SC, *et al*. Who's on third? regulation of third-party genetic interpretation services. *Genet Med* 2020;22(1):4–11.
42 Moran KS. Damned by DNA - Balancing personal privacy with public safety. *Forensic Sci Int* 2018;292:e3–4.
43 Ram N. Genetic privacy after carpenter. *Va L Rev* 2019;105.
44 Hazel JW, Slobogin C, What WK. And when: a survey of the privacy policies Proffered by US direct-to-consumer genetic testing companies. *Cornell JL & Pub Pol'y* 2018;28.
45 Arnold C. Crimefighting with family trees. *Nature* 2020;585:178–81.
46 Phillips AM. 'Only a click away - DTC genetics for ancestry, health, love…and more: A view of the business and regulatory landscape'. *Appl Transl Genom* 2016;8:16–22.
47 Christofides E, O'Doherty K. Company disclosure and consumer perceptions of the privacy implications of direct-to-consumer genetic testing. *New Genet Soc* 2016;35(2):101–23.
48 Bunnik EM, Janssens ACJW, Schermer MHN. A tiered-layered-staged model for informed consent in personal genome testing. *Eur J Hum Genet* 2013;21(6):596–601.
49 Laestadius LI, Rich JR, Auer PL. All your data (effectively) belong to us: data practices among direct-to-consumer genetic testing firms. *Genet Med* 2017;19(5):513–20.
50 FamilyTreeDNA Privacy Statement,. Available: https://www.familytreedna.com/legal/privacy-statement
51 Moreau Y. Crack down on genomic surveillance. *Nature* 2019;576(7785):36–8.
52 de Groot NF, van Beers BC, Decock L, *et al*. Accessing medical biobanks to solve crimes: ethical considerations. *J Med Ethics* 2021;47:502-509.. doi:10.1136/medethics-2020-106133
53 Parker M, Lucassen AM. Genetic information: a joint account? *BMJ* 2004;329(7458):165–7.
54 Parsons JA, Baker PE. From proband to provider: is there an obligation to inform genetic relatives of actionable risks discovered through direct-to-consumer genetic testing? *J Med Ethics* 2020:medethics-2020-106966.
55 Loi M. Direct to consumer genetic testing and the libertarian right to test. *J Med Ethics* 2016;42(9):574–7.
56 Bunnik EM, Janssens ACJW, Schermer MHN. Informed consent in direct-to-consumer personal genome testing: the outline of a model between specific and generic consent. *Bioethics* 2014;28(7):343–51.
57 Hansson SO. Informed consent out of context. *Journal of Business Ethics* 2006;63(2):149–54.
58 Schuppe J. 'They lied to us': Mom says police deceived her to get her DNA and charge her son with murder. *NBC News*, 2020. Available: https://www.nbcnews.com/news/us-news/they-lied-us-mom-says-police-deceived-her-get-her-n1140696
59 Dedrickson K. Universal DNA databases: a way to *improve* privacy? *J Law Biosci* 2017;4(3):637–47.
60 Hazel JW, Clayton EW, Malin BA, *et al*. Is it time for a universal genetic forensic database? *Science* 2018;362(6417):898–900.
61 Widdows H. *The connected self: the ethics and governance of the genetic individual*. Cambridge: Cambridge University Press, 2013.
62 Phillips D. Modern Crime-Solving Methods vs. the Mystery of World War II Deaths. *The New York Times*, 2021. Available: https://www.nytimes.com/2021/04/06/us/soldiers-death-dna-identification.html2021
63 Stern J, Zhang S. The Victims Left Behind by Genetic Genealogy. *The Atlantic*, 2021. Available: https://www.theatlantic.com/science/archive/2021/01/genetic-genealogy-race/616171/
64 Wiggins A, Wilbanks J. The rise of citizen science in health and biomedical research. *Am J Bioeth* 2019;19(8):3–14.
65 Machado H, Silva S. What influences public views on forensic DNA testing in the criminal field? A scoping review of quantitative evidence. *Hum Genomics* 2019;13(1):23.