



Data Sovereignty & Forensic Investigative Genetic Genealogy (FIGG): A Path Forward For Humanitarian & Mass Graves Investigations

Dowdeswell TL*

Professor, Department of Criminology & Legal Studies, Douglas College, Canada

*Corresponding author: Tracey Leigh Dowdeswell, Professor, Department of Criminology & Legal Studies, Douglas College, Canada, Email: dowdeswellt@douglascollege.ca

Investigation Paper

Volume 8 Issue 2

Received Date: April 11, 2023

Published Date: May 10, 2023

DOI: 10.23880/ijfsc-16000300

Abstract

This paper addresses several issues concerning the ethical governance of forensic investigative genetic genealogy (FIGG) in humanitarian investigations that seek to identify decedents in mass graves, disaster victims, and to reconstruct past atrocities. FIGG is better suited to human remains investigations than existing forensic DNA methods, such as partial matching in CODIS databases, and for this reason its use has increased. However, survivor communities may not benefit from the use of FIGG to reconstruct past events and promote the goals of healing and reconciliation unless we first address several pressing issues with the ethical governance of FIGG in humanitarian investigations. These include a lack of trust on the part of survivor communities, concerns over privacy, autonomy, informed consent, and the future uses of genetic data that generate an unwillingness to provide DNA for forensic investigations. This paper looks at the movement of Indigenous data sovereignty, which posits that control over data should be put in the hands of those who are most affected by its use, and its potential to be used as a blueprint for the ethical governance of FIGG in all humanitarian investigations. This is illustrated through recent examples of data sovereignty being applied by FIGG investigators: the private, non-profit DNA Justice database, and the mass graves investigations at the Mother and Baby Home in Tuam, Ireland.

Keywords: Mass Graves; Sovereignty; Tuam

Abbreviations: FIGG: Forensic Investigative Genetic Genealogy; MRCA: Most Recent Common Ancestors.

Introduction: Truth & Reconciliation

The Tulsa Race Massacre of 1921 devastated the African American community in an era of widespread racial violence. An estimated 300 people were killed, and more than 1200 homes were looted and burned.

The district of Greenwood – one of the most prosperous and culturally-significant Black neighbourhoods in America

– was destroyed, and its loss has set the community back for generations [1]. In an effort to uncover the truth of the massacre and bring some measure of justice for the survivors, the City of Tulsa embarked on an effort to exhume several of the mass graves and identify the victims through forensic investigative genetic genealogy (FIGG) [2]. They chose to partner with the non-profit laboratory, Intermountain Forensics. City and lab officials called upon local residents and survivors of the massacre to provide their DNA to assist in the forensic identifications of victims being unearthed from mass graves at the Oaklawn Cemetery. Justice for Greenwood, a non-profit organization that advocates for

victims of the Tulsa Race Massacre asked survivors not to provide their DNA. Their main concerns were a lack of privacy protections, and a mistrust of city officials who might misuse a DNA database consisting solely of Black residents of Tulsa [3].

To date, FIGG has been used mainly to identify unknown decedents and suspects in criminal cases. FIGG uses SNP-sequencing to identify distant relatives of the person of interest in GEDmatch and Family Tree DNA – the two direct to consumer genealogical databases that permit users to consent to the forensic use of their data (D2C databases). FIGG allows forensic genetic genealogists to identify a person of interest by finding relative matches in a D2C database, identifying their most recent common ancestors (MRCA), and then using traditional genealogical records to draw their family trees down to the present-day [4]. FIGG has proven to be a very powerful technique for identifying unknown persons – finding the proverbial needle in the haystack – but it has been used infrequently to identify victims buried in mass graves, disaster victims, or investigations into past atrocities (which I refer to in this paper as ‘humanitarian investigations’). Only two disaster victims have been identified using FIGG, both victims of the 2018 Paradise Camp Fire in California [5].

This is beginning to change. At present, two humanitarian investigations are underway that are using FIGG: the investigation into the Tulsa Race Massacre [3], and the investigation into the deaths and clandestine burials at the Bon Secours Mother and Baby Home in Tuam, Ireland [6]. FIGG is particularly well-suited to humanitarian identifications of this kind, and its use is therefore expected to increase.

However, the use of FIGG in humanitarian investigations raises a special set of ethical and privacy concerns. These incidents leave a legacy of pain and inter-generational trauma that can render survivors highly vulnerable, and survivor communities have experiences, worldviews, and cultural practices that need to be respected if healing is to take place [7].

Statement of the Problem

FIGG is better suited to human remains investigations than existing forensic DNA methods, and so it has the potential to make a positive contribution to humanitarian investigations. However, survivor communities will not benefit from the use of FIGG to reconstruct the truth of past events and promote the goals of healing and reconciliation unless we first address several pressing issues with the ethical governance of FIGG in humanitarian investigations.

Summary of Paper

This paper draws upon the movement of Indigenous data sovereignty to resolve several issues with the ethical governance of FIGG in humanitarian investigations, including privacy, autonomy, access to and control over data, and future unintended uses of data. Indigenous data sovereignty is a movement that seeks to situate control of data in the hands of those who are most affected by its use [8-11]. Indigenous data sovereignty principles should govern humanitarian investigations into Indigenous decedents, such as the mass graves at residential school sites in Canada [8]. More than this, the lessons learned from Aboriginal peoples’ struggle to gain control over their data is instructive for all investigations.

This paper will first describe the principles of Indigenous sovereignty. Next, I will describe the current limitations of existing forensic DNA methods for the identification of human remains, such as the degradation of DNA, and the necessity to first obtain DNA samples from first-degree relatives. FIGG can overcome many of these limitations. However, there are other limitations in humanitarian investigations that cannot be resolved through scientific or technical means – they are concerns over trust, privacy, control, and the autonomy of survivors and survivor communities. This has hampered the use of forensic DNA in humanitarian investigations, as well as our ability to deliver the kind of truth and reconciliation that is most needed by survivors. These issues are best resolved through the ethical governance of FIGG.

The paper will then explore two recent examples of data sovereignty being used to better govern FIGG investigations.

Indigenous Data Sovereignty

Most – but by no means all – of the decedents so far identified by FIGG are of European ancestry; the D2C databases reflect the demographics of their users, including those who have chosen to opt-in for law enforcement [4]. As of the end of 2022, 266 decedents are recorded as being identified using FIGG: 24 were of African American ancestry, 10 of Central/South American (Hispanic) ancestry, 9 were Indigenous, and 6 were of other, non-European ancestries (Middle Eastern, East Asian, Indian Subcontinental, and Sub-Saharan African) [5]. FIGG is therefore capable of identifying decedents of non-European ancestries, although to date the number of successful case resolutions does not reflect the numbers of decedents who need to be identified [4].

FIGG has begun to make some progress in the identification of Indigenous decedents. One of these is Shirley Ann Soosay, a woman who was found murdered in

Kern County, California. In May of 2022, Soosay was returned to her family on the Samson Cree Nation in Alberta, making her the first missing and murdered Indigenous woman to be returned home with the assistance of FIGG [9]. For 2023, it also includes the significant case of Daisy Mae Tallman – a young woman who went missing in 1987 in a remote area of the Yakama Nation – and who was identified by the Yakama Nation Tribal Police in January of 2023 [10]. This makes her the first Indigenous person to be identified by FIGG in an investigation led by her own tribal police force. Tallman's case demonstrates how useful FIGG can be when local communities have control over the data and its use during the investigation.

The identification of Indigenous decedents is most likely to be effective when performed in line with the principles of Indigenous data sovereignty. Indigenous data sovereignty is a movement that seeks to situate control of data in the hands of those who are most affected by its use [11]. One popular formulation of Indigenous data sovereignty is the CARE principles put forward by Carroll et al. They state that Indigenous data sovereignty is the “articulation of Indigenous Peoples’ rights and interests in data about their peoples, communities, cultures, and territories” and that it is “part of reclaiming control of data, data ecosystems, data science, and data narratives in the context of open data and open science” [12].

The principles articulated by Carroll et al. are: collective benefit – the data should be used for the benefit of the community and improved governance and citizen engagement; authority over data governance is situated in the community itself, including rights to control, use and remove the data, and to restrict its future use; responsibility – those governing the data are responsible for using it in ways that improve relationships and that are consistent with Indigenous customs, beliefs, and worldviews; and ethics – the data is used to promote justice, harms to Indigenous communities are minimized and benefits, including self-determination and equity, are maximized [12].

There are numerous formulations of Indigenous data sovereignty. For example, the First Nations Information Governance Centre in Akwesasne, Ontario has developed the First Nations Principles of OCAP, describing the principles of ownership, control, access, and possession of data that should be adhered to by those who collect or use the data of First Nations peoples. These principles assert that First Nations have a collective ownership interest in their data, and that they should have control over research and information that impacts them, including access to and control over data and how it is stored, managed, and used. Such data should be stewarded in line with First Nations’ worldviews, traditional knowledge, and cultural practices [13]. Researchers and

investigators should be aware of organizations and standards applicable to the Aboriginal peoples whose data they are collecting. This is especially true for investigations involving human remains, which need to be conducted in accordance with Indigenous worldviews, ceremonies, and customs.

The Limitations of Existing CODIS Databases

The CODIS software managed by the FBI is used in many law enforcement DNA databases in the United States and in about 50 countries around the world. CODIS databases are not only used to search for suspects: they have indexes that contain DNA profiles for unidentified human remains, as well as missing persons and relatives of missing persons, and these are separate from profiles obtained from offenders and forensic samples obtained from crime scenes [14]. Familial searching is used to identify matches between human remains and samples provided voluntarily by relatives of missing persons [15].

There are several limitations to using existing CODIS databases to identify human remains that can be mitigated by FIGG. DNA from remains degrades quickly, and the laboratory work to develop usable DNA profiles is expensive and time-consuming [16]. Next-generation SNP sequencing obtains more information from degraded DNA than does traditional forensic STR typing, and this can be determinative in many identifications [15].

SNP testing, when combined with FIGG, can also overcome the second main limitation of CODIS searching: CODIS is limited to identifying only very close familial relationships [17,18]. Less stringent search criteria are used to capture partial matches in CODIS, which is mainly useful in detecting first degree relatives – parents, children, and siblings [19]. CODIS misses many matches even between these close relatives, and this high rate of false negatives – again useful in the criminal context – means that many first-degree relatives are missed in these searches [17]. This is not surprising, since CODIS was designed to uncover exact matches of highly variable and individualizing forensic STR sequences. Dr. Bruce Budowle, the Director of the Center for Human Identification, estimates that there is about a 50-70% failure rate for these searches. With SNP testing, relationships can be inferred to a much greater degree of accuracy than with CODIS searching; even more importantly, we can discover associations with more distant relatives – up to the sixth degree and beyond, according to Dr. Budowle [17].

The above limitations are part of the reason why more decedent identifications have not been made using CODIS. Dr. Budowle estimates that there are about 7,000 DNA profiles of decedents in CODIS awaiting identification [17].

Canada's National DNA Data Bank has identified only 25 human remains since March 6, 2018, which is the date that the humanitarian indexes came into force [20].

The utility of traditional CODIS searches for identifying human remains is therefore limited and depends upon one or more first-degree relatives being located and voluntarily submitting their DNA to a law enforcement database. Such close relatives are not always available, particularly in humanitarian and historical investigations. In times of war or humanitarian crisis it is difficult to locate specific individuals; where exhumations are taking place at schools and institutions or other mass graves decades after atrocities took place, then many first-degree relatives necessary to make the identifications have passed on.

A History of Mistrust

Another reason why relatives of missing and unidentified persons may not want to provide their DNA to law enforcement databases is that they do not trust law enforcement to steward their data appropriately, protect their privacy, and refrain from misusing their genetic data in the future. If this trust cannot be rebuilt, then it will be difficult for relatives of the missing and survivor communities to benefit from advances in forensic DNA science to identify recovered remains and to reconstruct past atrocities.

At the same time, the communities who have suffered from these atrocities are often the very ones who mistrust genetic research, particularly when it is in the hands of law enforcement. Much of the mistrust on the part of African American and Indigenous peoples is based on past abuses by medical researchers and law enforcement. In one well-known example, the Havasupai Tribe of Arizona sued Arizona State University for sharing out and misusing their genetic data in ways that did not benefit the Havasupai and for which they had not provided informed consent. This included sensitive research into their history, lineage, and origins [21]. The case was eventually settled in the tribe's favour [22]. For Indigenous persons, biological materials are sacred, and so informed consent must be obtained from them; their DNA must be used in ways that are beneficial for the community and in line with their cultural worldviews and customs [21]. Many tribes have been reluctant to participate in genetic research – some have banned it outright – due to a lack of trust and an absence of culturally appropriate research practices [21].

African Americans have also been reluctant to provide their DNA for research and biobanks. Buseh et al. found that this mistrust stems from widespread knowledge about exploitative and unethical research practices. These include

the Tuskegee Study on Untreated Syphilis, which allowed nearly 400 African American men to die from syphilis in order to collect information about the progression of the disease [23]. Many also cited the exploitation of Henrietta Lacks – an African American woman who passed away from cervical cancer and who had her cervical tissues taken without her knowledge by researchers who then used them to produce a highly profitable immortal cell line [24]. Buseh found that, “Well-known abuses of African Americans at the hands of researchers generations ago are still very much alive in the African American consciousness” [23]. Apart from these notorious cases, participants were aware of recent abuses by health researchers in their local communities, and this contributed to their perceptions that researchers are not trustworthy and do not respect their rights and interests [23].

At the same time, Buseh found that research on genetics, health, and family history were seen in a positive light by the African Americans they interviewed, and they very much wanted to benefit from genetic research in these ways. Advances in genetics cannot benefit survivor communities until past abuses and present fears over privacy, control, and future uses of their data have been adequately addressed [23]. To do so, Garrison states that we need to resolve issues surrounding power, control, access to data, and culture and worldviews. We must “remain mindful of the diverse views of research participants and work harder to ensure that just and equitable research practices encourage communication and inclusion of minorities in research in order to break down the barriers of distrust” [21]. For example, DNA samples are treated as the property of the institutions who possess them – not the person to whom they belong – who then control the uses to which those samples are put [21].

The problems of access and control, privacy, responsibility and collective benefit are precisely those that data sovereignty is intended to address. Practitioners of FIGG should therefore take account of data sovereignty principles to build back some of this lost trust, and better ensure that these communities can benefit from these advances in forensic DNA science.

Data Sovereignty in Existing FIGG Investigations

The issues raised by African American and Indigenous peoples – to control, autonomy, privacy, well-being and the care and concern of the broader community – are widely shared. All humanitarian investigations can therefore benefit from the implementation of data sovereignty principles. It should come as no surprise, then, that there are several attempts at present to address concerns over privacy, control, and access to the genetic data used in FIGG.

One example is the current effort to exhume and identify the children who died at the Bon Secours Mother and Baby Home in Tuam, Ireland [6]. Excavations are currently underway to recover the children who died in this institution and were buried in unmarked graves on the site between 1925 and 1961 [6].

The scandal was brought to light by historian and genealogist Catherine Corless, who discovered 796 death certificates for children at the Tuam home who had no burial records [25]. Her work on the history of the lost children and their unknown graves led the Government of Ireland to establish the Mother and Baby Homes Commission in 2015 [26]. The Commission stated in its final report in 2021 that Ireland's mother and baby homes discriminated against and imposed harsh living conditions on women and illegitimate children, which contributed to very high rates of infant mortality at these institutions [26]. The Irish government made a commitment to exhume and identify each set of remains, and to hand them over to their surviving families for burial [6].

This investigation is being led by Forensic Science Ireland (FSI), a government-run laboratory. It appears that FSI plans to use SNP testing and genetic genealogy, which is not surprising given the limitations of traditional CODIS databases. What is novel is that FSI appears to have constructed a private, stand-alone database for relative matching. The entire process is expected to cost about €13 million. Gallagher describes the DNA database as one that will allow for more distant relatives to be matched, and that can be cross-referenced with the commercial D2C databases currently being used for FIGG [6]. This affords relatives and survivors a greater degree of privacy than they would receive if their DNA profiles were uploaded to the D2C databases and opted-in to law enforcement matching.

Another recent initiative has been launched by the DNA Justice Foundation, a non-profit organization established by several prominent FIGG practitioners. The DNA Justice database, in contrast to the commercial D2C databases, will not be made available to the public, and it will charge low or no fees to law enforcement agencies for FIGG investigations. The genetic and familial data uploaded to DNA Justice is to be used exclusively for law enforcement investigations, and DNA Justice seeks explicit consent from users for this purpose [27]. The terms and conditions of the DNA Justice database gives users a greater measure of privacy, control over their data, and transparency regarding its use [28].

These are promising advancements for the field. True data sovereignty in forensic investigations also means thinking in terms of data governance more broadly. Data sovereignty might be imagined in humanitarian

investigations as a private bioinformatics database that is dedicated to a particular investigation. The genetic and familial data provided by survivors and other members of the public will be used only to advance that investigation and will not be shared on the more open platforms currently in use for FIGG. The profiles of the decedents may be uploaded to these D2C databases, and thus have a greater chance of a successful resolution, if this is what the survivor community chooses.

To best achieve the principles of data sovereignty, the data should be stewarded by the survivor community itself. A written charter can be established for this purpose with the input of the relevant stakeholders, and then administered by an organization that is responsible to the community. This could take a number of forms, such as a tribal or First Nations government, a designated non-profit organization, or an academic or forensic institution. It might include a law enforcement organization, or a coroner or medical examiner, if the community has control over the terms of its use.

Some of the essential elements of data sovereignty that should be included in charters governing humanitarian FIGG investigations include the principles that those who provide the data are the ones who decide its conditions of use, including the ability to remove their data from the database, and the ability to destroy the database after the investigation is complete. Those who hold the data are responsible to the community, and they act as fiduciaries of the data which they administer in the best interests of those whose data it is and the survivor community as a whole. Charters to govern FIGG databases should ideally include an independent dispute resolution mechanism that is freely available to survivor communities. These measures might help to rebuild trust with survivor communities and enable them to take advantages of FIGG to reconstruct past atrocities and to return victims to their families for proper burial.

Conclusion: Bringing them Home

One of the first uses of genetic and genealogical data in an investigation into mass atrocities was in Argentina in the 1980s. At that time, survivors were calling upon the government to investigate crimes committed by the military dictatorship in Argentina's 'dirty war' – including widespread torture, imprisonment, and the 'disappearance' of about 30,000 suspected dissidents. As many as 450 children may have been removed from their families and placed in the care of regime members. The identification of the disappeared and the return of their kidnapped children was led by the Abuelas de Plaza de Mayo. In the early 1980s, the Abuelas sought assistance from the scientific community to locate and identify their grandchildren through genetic testing [29]. Dr. Mary-Claire King and the American Association

for the Advancement of Science answered the call, and they developed an index of grandpaternity, testing HLA markers in children with their potential grandparents to determine their degree of relatedness [30]. This was groundbreaking work in the genetics of kinship testing, and it paved the way for the present state of the field. It is also a landmark in the modern movement for truth and reconciliation in the aftermath of atrocity. The work of the Abuelas was seen by many survivors of the dirty war (although not all Argentinians) as an act of reconciliation-of truth-telling and healing not only for the families and children, but for Argentina as well [29].

Dr. Alondra Nelson has stated that part of the ‘social life’ of genomics is its use in reconciliation projects, in which “genetic analysis is used to contribute to community cohesion, collective memory, or social transformation” [31]. Today, the use of bioinformatics and genetic databases has given us an enormous power to identify human remains, to reconstruct past atrocities, and to effect the kind of social transformation that can build collective memory and engender healing. At the time of writing, the City of Tulsa announced that they had identified possible surnames of six of the individuals exhumed from Oaklawn Cemetery [32]. One of the survivors, Brenda Nails Alford said, “The information we have received today is absolutely historical” [32]. However, this potential may not be fully utilized without the ethical governance of these technologies – which cannot be separated from the power, the authority, and the past abuse of these communities that have given rise to their need in the first place. The lessons of data sovereignty, with its emphasis on ethics, on autonomy, and the use of data for the collective benefit of survivor communities, is one way to move forward with the use of these technologies in humanitarian investigations. It puts control over the investigations, and the sensitive data on which they depend, back into the hands of survivor communities themselves – so that they can bring their loved ones home, on their own terms.

References

- Debbie E, Marisa P (2021) A century after the race massacre, Tulsa confronts its bloody past. NPR.
- (1921) City of Tulsa 1921 Graves Investigation. City of Tulsa.
- Polansky C (2022) Massacre descendants’ group: don’t submit your DNA to Tulsa graves investigators. Public Radio Tulsa.
- Dowdeswell TL (2022) Forensic genetic genealogy: a profile of cases solved. *Forensic Sci Int Genet* 1(58): 102679.
- Dowdeswell TL (2023) Forensic genetic genealogy project v. 2022. Mendeley Data version 1.
- Gallagher C (2023) Advanced DNA technology to aid matching of infants buried at Tuam to relatives. *The Irish Times*.
- Bombay A, Matheson K, Anisman H (2014) The intergenerational effects of Indian residential schools: implications for the concept of historical trauma. *Transcult Psychiatry* 51(3): 320-338.
- (2022) Independent special interlocutor for missing children and unmarked graves and burial sites associated with Indian residential schools. Department of Justice Canada.
- Madeleine C (2022) Shirley Soosay’s remains being returned to Samson Cree Nation, decades after her murder. *CBC News*.
- Michael V (2023) Yakima County Coroner’s Office teams with Othram to identify 2008 Jane Doe. *DNASolves.com*.
- Tahu K, John T (2016) *Indigenous data sovereignty: toward an agenda*. Canberra: Australian National University Press.
- Carroll SR, Ibrahim G, Figueroa-Rodríguez OL, Holbrook J, Lovett R, et al. (2020) The CARE principles for Indigenous data governance. *Data Science Journal* 19(43): 1-12.
- (2023) The First Nations principles of OCAP®. First Nations Information Governance Centre.
- Osborn-Gustavson AE, McMahon T, Josserand M, Spammer BJ (2018) The utilization of databases for the identification of human remains. *New perspectives in forensic human skeletal identification* pp: 129-140.
- Alvarez-Cubero MJ, Saiz M, Martinez-Gonzalez LJ, Alvarez JC, Eisenberg AJ, et al. (2012) Genetic identification of missing persons: DNA analysis of human remains and compromised samples. *Pathobiol* 79(5): 228-38.
- Turingan RS, Brown J, Kaplun L, Smith J, Watson J, et al. (2020) Identification of human remains using Rapid DNA analysis. *International Journal of Legal Medicine* 134: 863-872.
- Mittelman D (2021) Heating up cold cases: an interview with Bruce Budowle on human identification. *Forensic Genomics* 1(1): 7-10.
- Debus-Sherrill S, Field MB (2019) Familial DNA searching: an emerging forensic investigative tool. *Sci Justice* 59(1): 20-28.

19. Bieber FR, Brenner CH, Lazer D (2006) Finding criminals through DNA of their relatives. *Science* 312(5778): 1315-1316.
20. (2022) The National DNA Databank of Canada – annual report 2020-2021. Royal Canadian Mounted Police.
21. Garrison NA (2013) Genomic justice for Native Americans: impact of the Havasupai case on genetic research. *Sci Technol Human Values* 38(2): 201-223.
22. (2008) Havasupai Tribe of the Havasupai Reservation v. Arizona Board of Regents and Therese Ann Markow. Arizona Court of Appeals, No. 1 CA-CV 07-0454 and 1 CA-CV 07-0801.
23. Buseh AG, Stevens PE, Millon-Underwood S, Townsend L, Kelber ST (2013) Community leaders' perspectives on engaging African Americans in biobanks and other human genetics initiatives. *J Community Genet* 4(4): 483-494.
24. Skloot R (2010) *The immortal life of Henrietta Lacks*. New York City.
25. McGarry P (2021) Catherine Corless: quest for truth about children buried at Tuam mother and baby home. *The Irish Times*.
26. (2021) Final report of the commission of investigation into mother and baby homes. Department of Children, Equality, Disability, Integration and Youth. Government of Ireland.
27. (2023) DNA Justice Foundation. Terms of service. DNA Justice Foundation.
28. (2023) DNA Justice Foundation. Privacy notice. DNA Justice Foundation.
29. Oren L (2001) Righting child custody wrongs: the children of the disappeared in Argentina. *Harvard Human Rights J* 14: 123-195.
30. Di Lonardo AM, Darlu P, Baur M, Orrego C, King MC (1984) Human genetics and human rights. Identifying the families of kidnapped children. *Am J Forensic Med Pathol* 5(4): 339-347.
31. Nelson A (2016) *The social life of DNA: race, reparations and reconciliation after the genome*. Boston: Beacon Press.
32. Codutti A (2023) Major scientific breakthrough' made in DNA investigation. *Tulsa World*.

