

Article

Genealogy Research and Higher Odds of Family Health History Confidence: A Cross-Sectional Study of US Adults Affected by Cancer

Lynette Hammond Gerido

Department of Bioethics, Case Western Reserve University School of Medicine, Cleveland, OH 44106, USA; lhgerido@case.edu

Abstract

Family health history (FHH) is an important tool for cancer risk assessment, yet its clinical utility is undermined by incomplete data, particularly among historically marginalized communities whose family records have been systematically disrupted by slavery, forced displacement, and undocumented status. I conducted a cross-sectional online survey of 1885 US adults affected by cancer to examine associations between genealogy engagement and FHH confidence and to assess variation across sociodemographic subgroups. Genealogy research was the single strongest modifiable predictor of FHH confidence, with participants reporting genealogy research more than twice as likely to report high confidence compared to those who did not (OR = 2.07, 95% CI 1.72–2.50), a relationship that persisted after full adjustment for age, sex, race, income, and education. Despite comparable rates of genealogy engagement, Black and Hispanic respondents reported substantially lower FHH confidence than White respondents. These findings suggest that while interest in genealogy is already widespread and self-motivated in the population, current clinical tools for FHH collection are inadequate to leverage this interest equitably. Next-generation FHH tools should be designed to reflect the dynamic, collaborative features of genealogy platforms rather than static medical forms, and their development must be grounded in community-centered design principles that prioritize populations for whom incomplete family history poses an added burden.

Keywords: family health history; cancer; genetic genealogy

1. Introduction

Family health history (FHH) is among the most clinically powerful and cost-effective tools (Valdez et al. 2010) available for assessing inherited disease risk (Lin et al. 2019) for rare diseases (i.e., Sickle Cell Disease) (Manswell Butty et al. 2012; Gilpin-Macfoy et al. 2023), cancer (Lowery et al. 2016), cardiovascular disease and diabetes (Onyenobi et al. 2026). Clinically FHH is used to refer at-risk patients for targeted screening, earlier initiation of surveillance and monitoring, and to specialized care (Ginsburg et al. 2019). Also, FHH informs genomic risk assessments and improves polygenic risk score calculations, with smaller population assessments achieving accuracy similar to large-scale polygenic risk scores (Truong et al. 2020; Mars et al. 2022). Despite these benefits of having a clinically documented FHH, many individuals with these conditions may not feel confident in their ability to share their FHH with their health care team (Beadles et al. 2014).

In addition to the cost advantages of FHH over other genomic tools, the conversations within families serve as patient education, improves FHH collection, and provides efficient



Received: 6 April 2026

Revised: 26 May 2026

Accepted: 14 June 2026

Published: 1 July 2026

Copyright: © 2026 by the author.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\) license](#).

identification of at-risk relatives. Yet millions of Americans, particularly those from historically marginalized communities, face documentation challenges and structural barriers (Chavez-Yenter et al. 2022; Lin et al. 2018). Slavery, forced displacement, and decades of structural exclusion has fractured family records in ways that continue to undermine both FHH completeness and equitable access to precision medicine today (Gerido 2022). As more people are interested in documenting their family histories and finding their relatives, genetic genealogy and direct-to-consumer genetic testing are emerging as an opportunity for intervention and patient education.

By using genetic genealogy, families are creating documentation of their biological relationships among their relatives (Steen n.d.). The genetic test results can confirm the degrees and types of biological relationships and potentially address gaps or discrepancies in documentation. Consumer platforms like African Ancestry, Inc., AncestryDNA, and 23andMe have made genetic ancestry testing accessible at scale. Communities are increasingly leveraging knowledge of their genetic relatives not only to reconstruct family trees but to demand accountability and justice. Genetic genealogical research may constitute a pathway toward greater confidence in the FHH individuals are able to report.

Despite this promise, little is known about whether engagement with genealogy, through ancestry testing or family tree research, is associated with greater confidence in an individual's ability to report FHH during clinical encounters, and whether confidence levels vary across Black and Hispanic populations already experiencing disparate burdens of cancer. If genealogy engagement is associated with greater confidence in FHH knowledge, then we can use those findings to study the subsequent effects on clinical reporting of FHH. The broader impacts of this research may lend itself to novel, family-centered opportunities for reducing health disparities in FHH documentation. The purpose of this study is to explore the relationship between genealogy engagement and FHH confidence in a national sample of US adults whose lives have been affected by cancer.

2. Methods

This is a cross-sectional, semi-structured online survey of US adults whose lives have been affected by cancer, either through a personal cancer diagnosis or having a relative with cancer. FHH functions as a clinically relevant risk indicator, reflecting not only shared genetic susceptibility but also shared environmental exposures, behavioral patterns, and social determinants of health that often aggregate within families. Cancer-affected individuals were therefore selected as the study population because FHH is clinically actionable for this group regardless of whether their cancer history reflects hereditary, environmental, or multifactorial etiology.

Data collection occurred in two recruitment waves using a third-party online recruitment platform, Centiment LLC (Centiment). The survey was administered via Qualtrics between 26 April 2023 and 1 June 2023. Centiment recruited participants from its registry and directed eligible individuals to the Qualtrics survey via the Centiment platform. Recruitment occurred in two sequential waves. In the first wave, eligibility criteria were purposively restricted to respondents who self-identified as Black in order to increase representation of a population that experiences a disproportionate burden of cancer mortality (Siegel et al. 2025). In the second wave, recruitment targeted a sample intended to more closely reflect the demographic distribution of the US adult population.

Before beginning survey questions, participants were provided with study information describing the Connected Families study purpose, procedures, voluntary participation, and the nature of the data collected. Individuals indicated consent to participate prior to proceeding. The student research team and I did not have access to personally identifying information at any time and the dataset used for analysis was deidentified. Participant

incentives were administered by Centiment through its platform in accordance with its standard participant compensation procedures.

Data quality was addressed through a two-stage process. In the first stage, Centiment LLC and I conducted quality control during the recruitment phase, refining recruitment strategy and filtering suspected bots and fraudulent accounts prior to data collection. In the second stage, the student research team and I reviewed all remaining responses and excluded those with incomplete data or identifiable patterns of nonsensical or gibberish responses in open-ended items. Of the 5188 total responses collected across both recruitment waves, 3303 were excluded through these combined procedures, achieving a final analytic sample size of $N = 1885$.

2.1. Measures

Primary outcome: Confidence in family health history (FHH) knowledge was assessed with a 5-point Likert item, *‘How confident are you that you could complete a summary of your family health history on a medical form?’* and analyzed as an ordinal outcome (1 = Not confident at all, 2 = A little confident, 3 = Somewhat confident, 4 = Very confident, 5 = Completely confident). For descriptive summaries, a “high confidence” indicator was derived corresponding to the upper two categories (Very confident/Completely confident).

Key predictors: Primary predictors reflected engagement with genealogy and genetic testing, including (1) self-reported genealogy research, assessed with the item *“Have you ever done research on your family tree?”* (Yes/No) and (2) having taken a genetic ancestry test, assessed with the item *“Did you take an ancestry test?”* (Yes/No). The ancestry test item was presented within a dedicated genetic testing section of the survey, preceded by brief descriptions of three types of direct-to-consumer genetic tests to orient respondents to the specific test type being referenced: paternity testing (‘used to determine biological fatherhood’), ancestry testing (‘used to explore clues about the origins of ancestors’), and clinical testing (‘used by health care providers to medically assess disease risk’).

Sociodemographic covariates: Age was converted from text to numeric and categorized into three groups: 18–49, 50–64, and 65 or older. Sex was analyzed as a binary variable (Male/Female). For race and ethnicity, the analytic categories of White, Black, and Hispanic were defined as single-race and single-ethnicity responses, including only respondents who selected that category alone and no others. Respondents who selected more than one racial or ethnic category were retained in descriptive analyses as a separate multiracial group. Categories with very small cell sizes, including Native Hawaiian/Pacific Islander and Middle Eastern North African (MENA), were retained for descriptive characterization but excluded from inferential analyses to reduce sparse-cell bias and protect interpretability. Household income was categorized into four levels ($\leq \$19,999$; $\$20,000$ – $\$39,999$; $\$40,000$ – $\$59,999$; $\geq \$60,000$) and education into five levels (less than high school; high school; post-high school training or some college; college graduate; postgraduate).

2.2. Data Preparation and Statistical Analysis

Survey data were cleaned and recoded in IBM SPSS Statistics version 30, with string responses converted to numeric variables and “Prefer not to answer” responses treated as missing. Race categories with very small cell counts were excluded or collapsed prior to inferential analyses. All analyses were conducted on the complete-case analytic dataset ($N = 1885$) using two-sided tests at $\alpha = 0.05$.

Analyses proceeded in three stages. First, I described sample characteristics using frequencies and proportions, and evaluated bivariate associations between FHH confidence and key predictors using cross-tabulations and Pearson chi-square tests, with Cramer’s V reported as an effect size measure. Second, I fit proportional-odds ordinal logistic regression

models to estimate adjusted associations with FHH confidence, with primary predictors of genealogy research and genetic ancestry test-taking and covariate adjustment for age, sex, race, income, and education. I verified the proportional odds assumption using the Test of Parallel Lines and reported adjusted effects as odds ratios with 95% confidence intervals. Third, to improve interpretability, I derived model-based predicted probabilities of high FHH confidence, defined as P(Very confident/Completely confident), from the fully adjusted model and summarized these across key subgroups.

3. Results

3.1. Sample Population Characteristics

Table A1 summarizes the analytic sample (N = 1885). The sample was roughly evenly split across age groups, with approximately half of respondents ages 18–49 (50.6%), one quarter ages 50–64 (25.9%), and nearly one quarter ages 65 or older (23.6%). Nearly half reported conducting genealogy research (49.3%), while fewer than one in four reported having taken a genetic ancestry test (23.1%). The racial composition of the analytic sample was 49.5% Black, 37.1% White, 6.0% multiracial, 4.5% Hispanic, and 2.8% Asian, with very small representation from Indigenous populations. Racial and ethnic categories smaller than 2% were removed from inferential analyses. Women comprised 59.6% of the sample and men 40.4%. Income and education distributions for the complete-case modeling sample are shown in Table A1.

3.2. Bivariate Associations with FHH Confidence

In bivariate analyses (Table A2), genealogy research showed the strongest association with FHH confidence. Respondents reporting genealogy research were substantially more likely to report high confidence (Very confident/Completely confident: 57.9%) than those not reporting genealogy research (36.0%), $\chi^2(4) = 110.20$, $p < 0.001$, $V = 0.24$. Having taken a genetic ancestry test was also associated with higher confidence (high confidence: 54.8% among test-takers vs. 44.4% among non-test-takers), $\chi^2(4) = 28.15$, $p < 0.001$, $V = 0.12$. Confidence varied across demographic subgroups, increasing with age (40.8% in ages 18–49; 49.1% in 50–64; 57.2% in 65+; $\chi^2(8) = 41.26$, $p < 0.001$, $V = 0.11$) and differing by race/ethnicity ($\chi^2(16) = 62.12$, $p < 0.001$, $V = 0.09$), with the highest prevalence of high confidence among White respondents (56.2%) and lower prevalence among Black (40.6%) and Hispanic respondents (34.1%). Socioeconomic gradients were evident for both income (33.4% high confidence in $\leq \$19,999$ vs. 56.5% in $\$60,000+$; $\chi^2(12) = 70.71$, $p < 0.001$, $V = 0.11$) and education (25.5% in <High School vs. 63.9% in Postgraduate; $\chi^2(16) = 76.42$, $p < 0.001$, $V = 0.10$). Sex differences were comparatively small (48.4% high confidence among females vs. 44.5% among males) and not statistically significant in the omnibus chi-square test ($\chi^2(4) = 7.90$, $p = 0.095$).

3.3. Multivariable Ordinal Regression Predicting FHH Confidence (Odds Ratios)

The proportional-odds ordinal logistic regression models (logit link) predicting higher FHH confidence were adjusted for genealogy research, genetic ancestry test-taking, age group, sex, race, income, and education. The proportional odds assumption was met (Test of Parallel Lines $p = 0.476$). For genetic ancestry test-taking and genealogy research, the modeled reference categories were oriented such that $OR > 1$ indicated higher confidence, requiring no re-expression; the corresponding odds ratios from the fully adjusted model are reported in Table A3. Table A4 presents the complete fully adjusted model, with all remaining predictors re-expressed so that $OR > 1$ uniformly reflects higher confidence.

In the fully adjusted model (Table A4), genealogy research remained strongly associated with higher FHH confidence (Yes vs. No $OR = 2.07$, 95% CI 1.72–2.50, $p < 0.001$). Race

disparities persisted after adjustment: compared with White respondents, Black respondents (OR = 0.60, 95% CI 0.50–0.73, $p < 0.001$) and Hispanic respondents (OR = 0.52, 95% CI 0.34–0.78, $p = 0.002$) had lower odds of higher confidence. Respondents ages 18–49 had lower odds of higher confidence than respondents ages 65+ (OR = 0.77, 95% CI 0.62–0.96, $p = 0.020$). Sex differences were modest but statistically significant after full adjustment (Male vs. Female OR = 0.82, 95% CI 0.69–0.97, $p = 0.023$).

Socioeconomic measures were strongly associated with FHH confidence. Relative to respondents with household income $\geq \$60,000$, those with income $\leq \$19,999$ (OR = 0.56, 95% CI 0.43–0.72, $p < 0.001$) and \$20,000–\$39,999 (OR = 0.64, 95% CI 0.51–0.80, $p < 0.001$) had lower odds of higher confidence. Relative to respondents with postgraduate education, respondents with less than high school education (OR = 0.39, 95% CI 0.22–0.68, $p = 0.001$), high school education (OR = 0.67, 95% CI 0.49–0.93, $p = 0.018$), and post-high school training/some college (OR = 0.65, 95% CI 0.48–0.88, $p = 0.005$) had lower odds of higher confidence. Having taken a genetic ancestry test showed a smaller and borderline association with FHH confidence after full adjustment (Yes vs. No OR = 1.22, 95% CI 0.99–1.52, $p = 0.067$).

3.4. Model-Adjusted Predicted Probabilities of High FHH Confidence

To improve interpretability, we derived model-based predicted probabilities of high confidence, defined as P(Very confident/Completely confident) from the fully adjusted ordinal model, and summarized marginal predicted probabilities across subgroups (Table A5). Predicted high confidence was substantially higher among respondents reporting genealogy research (57.1%) compared with those not reporting genealogy research (36.0%). Predicted high confidence varied by race, with the highest probabilities among White respondents (55.7%) and lower probabilities among Black (40.5%) and Hispanic respondents (34.9%). Predicted high confidence increased with age (42.2% for ages 18–49; 47.3% for 50–64; 54.3% for 65+), income (34.2% for $\leq \$19,999$ vs. 56.2% for \$60,000+), and education (26.8% for less than high school vs. 62.1% for postgraduate). Sex differences were modest (47.7% for females vs. 44.4% for males).

4. Discussion

This study examined associations between genealogy engagement and confidence in family health history (FHH) knowledge among US adults affected by cancer. Our most striking finding was that genealogy research was strongly and independently associated with higher FHH confidence. Adults who reported conducting genealogy research were more than twice as likely to report high FHH confidence compared to those who did not (OR = 2.07), a relationship that persisted after adjustment for age, sex, race, income, and education. This finding is consistent with emerging evidence that engagement with family history cultivates structured knowledge of kinship networks, health patterns, and intergenerational relationships that may support greater capacity to report FHH in clinical settings (Mitchell and Kim 2024; Cramm et al. 2025). The act of genealogical inquiry may help individuals to organize family information in ways that make health patterns more visible, more communicable, and more clinically useful. FHH can be “passed on to future generations, providing a foundational risk profile for immediate family and their progeny for generations to come” (Ginsburg et al. 2019).

A secondary but theoretically important finding concerns the differential associations of genealogy research and direct-to-consumer (DTC) genetic ancestry testing with FHH confidence. In bivariate analyses, both predictors were associated with higher confidence. However, in the fully adjusted model, genealogy research remained a strong and statistically robust predictor (OR = 2.07, $p < 0.001$), while having taken a genetic ancestry test showed

a weaker association that did not reach conventional significance ($OR = 1.22$, $p = 0.067$). This pattern is consistent with the interpretation that it is the active process of genealogical inquiry, seeking out records, reconstructing kinship networks, and engaging deliberately with family history, that is most strongly associated with the kind of organized family knowledge that supports FHH confidence. This distinction has practical implications for how clinicians and public health practitioners think about leveraging genetic genealogy tools for FHH improvement: simply encouraging patients to take ancestry tests may be insufficient if the goal is to build actionable family health knowledge. These interpretations will require future longitudinal research directly comparing active genealogical research with passive DTC test-taking on FHH outcomes to test this hypothesis rigorously.

Despite comparable rates of genealogy engagement, Black and Hispanic respondents reported substantially lower FHH confidence than White respondents, and these disparities persisted in fully adjusted models. This pattern cannot be explained by individual behavior or motivation alone. For communities whose family records were systematically suppressed or never created, the structural conditions that produced these gaps have deep historical roots and present-day ramifications. The legal exclusion of enslaved people from records that affirmed kinship and personhood was enforced through law, labor discipline, and the deliberate fragmentation of family units. It created archival absences that genealogical research cannot easily overcome (Rierson 2019; Müller 2018). These historical patterns find a contemporary parallel in the experiences of undocumented immigrant families, for whom engagement with institutional record-keeping systems carries significant legal risk and whose family members may span multiple countries, languages, and documentation regimes (Rierson 2019; Segarra and Prasad 2024). In both historical and contemporary contexts, the state of being undocumented is not an accident or an oversight. It impacts generations and is a persistent and compounding deficit in the family records that both genealogical research and FHH collection depend upon.

FHH remains an important tool for cancer risk assessment, informing referral decisions, eligibility for genetic counseling, and insurance coverage for clinical genetic testing (Cramm et al. 2025; Pandya et al. 2025). It is important to distinguish, however, between the construct measured in this study, self-reported confidence in one's ability to report FHH, and the clinical completeness and accuracy of FHH that is documented in the medical record. This study establishes that genealogy engagement is associated with greater confidence but it does not establish that greater confidence produces more complete or accurate clinical FHH reporting. That relationship is a plausible hypothesis, but it requires direct empirical testing before causal claims can be made. Nevertheless, the limitation of relying on self-reported FHH without independent verification of accuracy is recognized across the literature; Ricks-Santi et al. (2016) identified the same constraint in their study of self-reported family health history of breast cancer among minority populations, noting that family histories carefully and systematically obtained have demonstrated a sensitivity of 75% and moderately good agreement with cancer registry records (Ricks-Santi et al. 2016). This evidence suggests that the quality of self-reported FHH is not fixed. Future research should prioritize the examination of genealogy-influenced confidence and whether it produces greater clinical accuracy and utility. FHH collection is underscored by mounting evidence that current clinical tools are systematically failing to capture complete and accurate family health information. Studies using genetically defined kinship data have demonstrated that cancer family histories recorded in electronic health records are substantially incomplete, with significant implications for identifying patients who carry hereditary cancer risk (Kiser et al. 2025). This incompleteness has measurable downstream consequences: patients with incomplete family histories are less likely to meet algorithmic criteria for referral to genetic evaluation for hereditary cancer, even when their true family history would qualify them (Harris et al. 2025). The processes by which family

cancer history is collected in clinical settings have been identified as a significant point of failure, with workflow limitations, time constraints, and the inadequacy of standard intake instruments all contributing to gaps that may go undetected and uncorrected (Andrusko and Paradiso 2022; Allen et al. 2023).

Additionally, the populations most overrepresented in forensic genetic databases are frequently the same populations underrepresented in clinical genetic services and most likely to report low FHH confidence in this study. The acceleration of forensic applications without a parallel development of health-oriented tools for these communities continues a disconcerting pattern in which genetic methods are more readily deployed in contexts of law enforcement than in contexts of care, prevention, and community empowerment. This pattern has historical antecedents that the genetic genealogy research community cannot afford to ignore. The collection of biological and familial data, from Black, Hispanic, and other marginalized communities, has not always served the interests of those communities, and the memory of that history. From the exploitation of Henrietta Lacks, which left enduring ethical questions about research autonomy and consent (Wolinetz and Collins 2020) to the use of genetic data in immigration enforcement (Zaretsky 2021), shapes how these communities reasonably approach institutions asking for their biological information. Trust, once broken at a structural level, is not repaired by institutional innovation alone (Creary and Gerido 2023).

Addressing this challenge requires a commitment to community-centered solutions that position affected people as architects of the interventions being built. The community-centered genealogical projects that have emerged in response to histories of oppression and displacement offer instructive models for what this kind of community-centered approach can look like in practice. The Georgetown University 272 project, the 1921 Tulsa Race Massacre Identification Project, the 10 Million Names, and countless genetic family reunification programs (Gelatt 2024) each demonstrate that communities whose family records were most severely disrupted are also among the most motivated and organized to reconstruct them. When genealogical methods are deployed in service of community-defined goals, they can become powerful instruments of identity, healing, and justice. These projects also illustrate the scale of the methodological and relational infrastructure required to do this work responsibly. The genetic genealogy research community is uniquely positioned to help translate these lessons into the design of health-oriented tools.

Limitations and Future Directions

This study has several limitations that warrant consideration. The cross-sectional design precludes causal inference; the associations observed between genealogy research and FHH confidence reflect relationships in the data but cannot establish directionality or rule out unmeasured or residual confounding, which is not accounted for by the covariates included in the fully adjusted model. Genealogy engagement and genetic genealogy uptake were assessed with two single binary items, which capture whether respondents have engaged in any genealogical inquiry or have taken an ancestry test but does not distinguish depth, recency, or outcomes; future studies should employ multi-item measures that can differentiate casual from sustained genealogical activity and assess whether dose–response relationships exist with FHH confidence. Longitudinal and retrospective designs will be necessary to determine whether genealogy engagement precedes and produces gains in FHH confidence or whether individuals with greater existing family knowledge are more likely to pursue genealogical research. Additionally, this study measured self-reported genealogy engagement without assessing the quality or accuracy of the family information participants obtained through that research. This study measured self-reported confidence in FHH knowledge but did not assess the reliability of the information on which that

confidence is based. Family health histories are subject to multiple, compounding sources of inaccuracy. These include: historical misdiagnosis or non-recognition of cancer in previous generations, particularly for cancers that were poorly understood or underdiagnosed in earlier eras; the transmission of imprecise or incomplete information through family oral histories and lore, where diagnoses may be misremembered, generalized, or lost across generations; the use of shared, user-generated genealogical databases, which may contain substantial numbers of unverified entries and errors in documented relationships. A participant may report high confidence in FHH that is itself incomplete or inaccurate for any of these reasons. Future research should therefore assess not only FHH confidence but the concordance between self-reported family histories and independently verified records, and should develop tools capable of capturing the full contextual complexity that comprehensive FHH collection requires.

The analytic sample was demographically diverse but failed to represent the diversity the rich diversity that exists within and between each of the groups that have been singularly described (i.e., Asian, Black, Hispanic, White). Future research may employ frameworks to better reflect the presence of biracial, multiracial, and mixed individuals and families. Also, the sample was drawn entirely from adults whose lives have been affected by cancer. This shared characteristic limits generalizability to the broader population, and future studies should include individuals without cancer experience to establish whether the observed associations hold across a wider range of FHH engagement contexts. Finally, recruitment through an online survey platform likely introduced selection bias toward individuals comfortable completing surveys in digital environments, potentially underrepresenting populations who would be more reachable through telephone or paper-based survey modalities. Future research should employ mixed-mode recruitment strategies to ensure adequate representation of populations least likely to engage through online platforms.

5. Conclusions

Genealogy research is strongly associated with greater confidence in FHH knowledge and represents a largely untapped resource that warrants consideration as a potential lever for improving clinical FHH collection, yet its benefits remain unevenly distributed across the populations that need them most. Addressing these disparities will require more than encouraging individuals to research their family histories. It will require building the tools, systems, and community partnerships that make that work possible, accessible, and safe. Because these communities have a legitimate and historically grounded mistrust toward institutions misusing their biological and familial data, safety must be more than a technical requirement or a procedural standard. As discussed above, the populations most likely to benefit from genealogical tools for FHH are frequently the same populations whose health and family data have historically been weaponized for surveillance and exploitation rather than care. Building tools that are genuinely safe for these communities means earning trust through transparent data governance, community ownership of research processes, and a demonstrated commitment to deploying genetic methods in service of community-defined health goals. By integrating the expertise of the genetic genealogy research community with health informatics, clinical practice, and the communities most affected by incomplete family records, we can build next-generation tools that are co-designed to serve the populations that need them most.

Funding: This research was funded by National Cancer Institute grant number L60CA274876.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and the study protocol (HUM00215047) was approved by the Institutional

Review Board (IRB) of the University of Michigan in Ann Arbor, MI, USA (protocol code IRB00000246 and date of approval: 20 September 2022).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: I gratefully acknowledge the African American Genealogical Society, Cleveland, Ohio (AAGS) and the Afro-American Historical and Genealogical Society, Inc. (AAHGS) for their commitment to preserving family histories and their broader contributions to the field of genealogical research that informed the development of this work.

Conflicts of Interest: The author declares no conflict of interest.

Appendix A

Table A1. Sample population characteristics.

Characteristic	n	%
Total analytic sample	1885	100.0
Genealogy research		
No	955	50.7
Yes	930	49.3
Has taken genetic ancestry test		
No	1449	76.9
Yes	436	23.1
Age group		
18–49	953	50.6
50–64	488	25.9
65+	444	23.6
Race †		
More than one	113	6.0
Asian	52	2.8
Black	926	49.5
Hispanic	85	4.5
White	694	37.1
Sex ‡		
Male	753	40.4
Female	1113	59.6
Household income §		
≤\$19,999	348	19.4
\$20,000–\$39,999	477	26.6
\$40,000–\$59,999	367	20.4
\$60,000+	604	33.6
Characteristic	n	%
Education (edugrp5) §		
Less than High School	52	2.9
High School	385	21.4

Table A1. *Cont.*

Post-HS Training/Some College	640	35.6
College Graduate	510	28.4
Postgraduate	209	11.6

† Race denominator: N = 1870 (analytic race variable excludes small-count categories). Percentages are within N = 1870. ‡ Sex denominator: N = 1866 (binary analytic sex excludes non-binary/prefer-not categories). Percentages are within N = 1866. § Income/Education denominator: N = 1796 (complete-case set used for the fully adjusted ordinal model and predicted probabilities). Percentages are within N = 1796.

Appendix B

Table A2. Bivariate associations between FHH confidence and key predictors/covariates. *Outcome:* fhhconfidence_n (5 levels). *High confidence* defined as Very confident/Completely confident.

Predictor	N (Valid)	Pearson χ^2 (df)	p-Value	Cramer's V	High Confidence (Very Confident/Completely Confident), % by Group
Genealogy research (Yes/No)	1885	110.20 (4)	<0.001	0.24	No: 36.0%; Yes: 57.9%
Has taken genetic ancestry test (Yes/No)	1885	28.15 (4)	<0.001	0.12	No: 44.4%; Yes: 54.8%
Age group (18–49, 50–64, 65+)	1885	41.26 (8)	<0.001	0.11	18–49: 40.8%; 50–64: 49.1%; 65+: 57.2%
Sex (Male/Female; binary analytic)	1881	7.90 (4)	0.095	0.07	Male: 44.5%; Female: 48.4%
Race (analytic categories; excludes Native American, MENA, NHPI)	1870	62.12 (16)	<0.001	0.09	White: 56.2%; Black: 40.6%; Hispanic: 34.1%; Asian: 44.2%; More than one: 48.7%
Household income (4 categories)	1816	70.71 (12)	<0.001	0.11	≤\$19,999: 33.4%; \$20–39,999: 44.8%; \$40–59,999: 45.3%; \$60,000+: 56.5%
Education (5 categories)	1880	76.42 (16)	<0.001	0.10	<HS: 25.5%; HS: 42.3%; Some college: 41.9%; College grad: 51.5%; Postgrad: 63.9%

Notes. χ^2 values are Pearson chi-square tests from cross-tabulations of each predictor with fhhconfidence_n. Race categories were restricted as noted to support adequate expected counts for inference. Cramer's V is reported as an effect size.

Appendix C

Table A3. Key predictors in the “positive direction” (higher confidence).

Predictor (Comparison)	OR	95% CI	p
Has taken genetic ancestry test: Yes vs. No	1.22	0.99–1.52	0.067
Genealogy research: Yes vs. No	2.07	1.72–2.50	<0.001

Table A4. Full model re-expressed so OR > 1 indicates higher confidence.

Predictor (Comparison)	OR	95% CI	p
Has taken genetic ancestry test: Yes vs. No	1.22	0.99–1.52	0.067
Genealogy research: Yes vs. No	2.07	1.72–2.50	<0.001
Age: 65+ vs. 18–49	1.30	1.04–1.61	0.020
Age: 65+ vs. 50–64	1.11	0.87–1.41	0.414
Sex: Female vs. Male	1.22	1.03–1.45	0.023
Race: White vs. More than one	1.18	0.81–1.72	0.377

Table A4. *Cont.*

Predictor (Comparison)	OR	95% CI	<i>p</i>
Race: White vs. Asian	1.59	0.94–2.63	0.082
Race: White vs. Black	1.67	1.37–2.00	<0.001
Race: White vs. Hispanic	1.92	1.28–2.94	0.002
Income: \$60,000+ vs. ≤\$19,999	1.79	1.39–2.33	<0.001
Income: \$60,000+ vs. \$20,000–\$39,999	1.56	1.25–1.96	<0.001
Income: \$60,000+ vs. \$40,000–\$59,999	1.25	0.98–1.59	0.066
Education: Postgraduate vs. <High School	2.56	1.47–4.55	0.001
Education: Postgraduate vs. High School	1.49	1.08–2.04	0.018
Education: Postgraduate vs. Post-HS/Some College	1.54	1.14–2.08	0.005
Education: Postgraduate vs. College Graduate	1.22	0.90–1.64	0.195

Appendix D

Table A5. Adjusted predicted probability of high FHH confidence (Very confident/Completely confident). *Model-adjusted predicted probability computed as $P(\text{Very confident/Completely confident}) = \text{EST4}_1 + \text{EST5}_1$ from the fully adjusted ordinal logistic regression. Values shown as proportions (and %).*

Domain	Category	Adjusted P(Very Confident/Completely Confident)	N
Genealogy research	No	0.3597 (36.0%)	908
	Yes	0.5705 (57.1%)	888
Race (excludes Native Amer., MENA, NHPI)	More than one	0.4776 (47.8%)	107
	Asian	0.4316 (43.2%)	51
	Black	0.4050 (40.5%)	887
	Hispanic	0.3488 (34.9%)	84
	White	0.5571 (55.7%)	667
Age group	18–49	0.4224 (42.2%)	912
	50–64	0.4726 (47.3%)	454
	65+	0.5429 (54.3%)	430
Household income	≤\$19,999	0.3418 (34.2%)	348
	\$20,000–\$39,999	0.4301 (43.0%)	477
	\$40,000–\$59,999	0.4630 (46.3%)	367
	\$60,000+	0.5617 (56.2%)	604
Education	Less than High School	0.2677 (26.8%)	52
	High School	0.4100 (41.0%)	385
	Post-HS Training/Some College	0.4199 (42.0%)	640
	College Graduate	0.5154 (51.5%)	510
	Postgraduate	0.6215 (62.1%)	209
Sex	Male	0.4443 (44.4%)	732
	Female	0.4775 (47.7%)	1064
Total		0.4639 (46.4%)	1796

References

- Allen, Caitlin G., Ridgely Fisk Green, Nicole F. Dowling, Temeika L. Fairley, and Muin J. Khoury. 2023. Understanding the Process of Family Cancer History Collection and Health Information Seeking. *Health Education & Behavior* 50: 572–85. [\[CrossRef\]](#) [\[PubMed\]](#)
- Andrusko, Danielle, and Catherine Paradiso. 2022. Establishing a Process to Improve the Collection of Family Health History. *The Nurse Practitioner* 47: 32. [\[CrossRef\]](#) [\[PubMed\]](#)
- Beadles, Chris A., R. Ryanne Wu, Tiffany Himmel, Adam H. Buchanan, Karen P. Powell, Elizabeth Hauser, Vincent C. Henrich, Geoffrey S. Ginsburg, and Lori A. Orlando. 2014. Providing Patient Education: Impact on Quantity and Quality of Family Health History Collection. *Familial Cancer* 13: 325–32. [\[CrossRef\]](#) [\[PubMed\]](#)
- Chavez-Yenter, Daniel, Melody S. Goodman, Yuyu Chen, Xiangying Chu, Richard L. Bradshaw, Rachelle Lorenz Chambers, Priscilla A. Chan, Brianne M. Daly, Michael Flynn, Amanda Gammon, and et al. 2022. Association of Disparities in Family History and Family Cancer History in the Electronic Health Record With Sex, Race, Hispanic or Latino Ethnicity, and Language Preference in 2 Large US Health Care Systems. *JAMA Network Open* 5: e2234574. [\[CrossRef\]](#) [\[PubMed\]](#)
- Cramm, Jessica, Pamela Flodman, Jacqueline Kim, Katherine Hall, and John Jay Gargus. 2025. O46: Knowledge of Family Health History and Its Association with Social Determinants of Health and Personal Medical History. *Genetics in Medicine Open* 3: 102380. [\[CrossRef\]](#)
- Creary, Melissa, and Lynette Hammond Gerido. 2023. The Public Performativity of Trust. *Hastings Center Report* 53: S76–S85. [\[CrossRef\]](#) [\[PubMed\]](#)
- Gelatt, Julia. 2024. Family Reunification and Separation: How U.S. Immigration Policy Shapes Family Structures. In *Immigration Policy and Immigrant Families*. Edited by Jennifer Van Hook and Valarie King. Cham: Springer Nature. [\[CrossRef\]](#)
- Gerido, Lynette Hammond. 2022. Uprooting Racial Health Disparities: Genealogy as a Community Health Library Service. Proceedings of the ALISE Annual Conference, October 20. Available online: <https://www.ideals.illinois.edu/items/126978> (accessed on 15 March 2026). [\[CrossRef\]](#)
- Gilpin-Macfoy, Fiona, Mindy J. Perilla, and Laura M. Koehly. 2023. Variability in Sickle Cell Knowledge by Sickle Cell Status. *Journal of Genetic Counseling* 32: 916–25. [\[CrossRef\]](#) [\[PubMed\]](#)
- Ginsburg, Geoffrey S., Rebekah. Ryanne Wu, and Lori A. Orlando. 2019. Family Health History: Underused for Actionable Risk Assessment. *The Lancet* 394: 596–603. [\[CrossRef\]](#) [\[PubMed\]](#)
- Harris, Adrian, Jemar R. Bather, Richard L. Bradshaw, Kensaku Kawamoto, Guilherme Del Fiol, Wendy K. Kohlmann, Daniel Chavez-Yenter, Rachel Monahan, Rachelle L. Chambers, Meenakshi Sigireddi, and et al. 2025. Incomplete Family History and Meeting Algorithmic Criteria for Genetic Evaluation of Hereditary Cancer. *JAMA Network Open* 8: e2539870. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kiser, Daniel, Gai Elhanan, Karen A. Schlauch, Robert Read, and Joseph J. Grzymalski. 2025. Assessing Completeness of Cancer Family History in EHRs Using Genetically Defined Kinships: A Cross-Sectional Study. *Genetics in Medicine Open* 3: 103450. [\[CrossRef\]](#) [\[PubMed\]](#)
- Lin, Jielu, Christopher S. Marcum, Melanie F. Myers, and Laura M. Koehly. 2018. Racial Differences in Family Health History Knowledge of Type 2 Diabetes: Exploring the Role of Interpersonal Mechanisms. *Translational Behavioral Medicine* 8: 4. [\[CrossRef\]](#) [\[PubMed\]](#)
- Lin, Jielu, Melanie F. Myers, Laura M. Koehly, and Christopher Steven Marcum. 2019. A Bayesian Hierarchical Logistic Regression Model of Multiple Informant Family Health Histories. *BMC Medical Research Methodology* 19: 56. [\[CrossRef\]](#) [\[PubMed\]](#)
- Lowery, Jan T., Dennis J. Ahnen, Paul C. Schroy, III, Heather Hampel, Nancy Baxter, C. Richard Boland, Randall W. Burt, Lynn Butterly, Megan Doerr, Mary Doroshenko, and et al. 2016. Understanding the Contribution of Family History to Colorectal Cancer Risk and Its Clinical Implications: A State-of-the-Science Review. *Cancer* 122: 2633–45. [\[CrossRef\]](#) [\[PubMed\]](#)
- Manswell Butty, Jo-Anne, Finie Richardson, Charles P. Mouton, Charmaine D. M. Royal, Rodney D. Green, and Kerry-Ann Munroe. 2012. Evaluation Findings from Genetics and Family Health History Community-Based Workshops for African Americans. *Journal of Community Genetics* 3: 1–12. [\[CrossRef\]](#) [\[PubMed\]](#)
- Mars, Nina, Joni V. Lindbohm, Pietro della Briotta Parolo, Elisabeth Widén, Jaakko Kaprio, Aarno Palotie, and Samuli Ripatti. 2022. Systematic Comparison of Family History and Polygenic Risk across 24 Common Diseases. *The American Journal of Human Genetics* 109: 2152–62. [\[CrossRef\]](#) [\[PubMed\]](#)
- Mitchell, Barbara A., and Boah Kim. 2024. Amateur Family Genealogists Researching Their Family History: A Scoping Review of Motivations and Psychosocial Impacts. *Genealogy* 8: 3. [\[CrossRef\]](#)
- Müller, Alex. 2018. Beyond 'Invisibility': Queer Intelligibility and Symbolic Annihilation in Healthcare. *Culture, Health & Sexuality* 20: 1. [\[CrossRef\]](#) [\[PubMed\]](#)
- Onyenobi, Ebuka, Knightess Oyibo, Michael Zhong, and Sally N. Adebamowo. 2026. Evaluating the Impact of Family History and Polygenic Risk Scores on Cardiometabolic Disease Risk. *BMC Medicine* 24: 140. [\[CrossRef\]](#) [\[PubMed\]](#)
- Pandya, Yasheel, Angela Zhu, Ethan Mondell, Sophia Choi, Gina Kim, Valentina Vanos, Siddharth Venkatraman, Panagis Galiatsatos, and Joann Bodurtha. 2025. Perceptions of Family Health History in an East Baltimore Community. *Journal of Community Genetics* 16: 581–87. [\[CrossRef\]](#) [\[PubMed\]](#)

- Ricks-Santi, Luisel J., Nicole Thompson, Altovise Ewing, Barbara Harrison, Kimberly Higginbotham, Cherie Spencer, Adeyinka Laiyemo, Robert DeWitty, Lori Wilson, Sara Horton, and et al. 2016. Predictors of Self-Reported Family Health History of Breast Cancer. *Journal of Immigrant and Minority Health* 18: 5. [CrossRef] [PubMed]
- Rierson, Sandra L. 2019. Fugitive Slaves and Undocumented Immigrants: Testing the Boundaries of Our Federalism. *University of Miami Law Review* 74: 598–710.
- Segarra, Paulina, and Ajnesh Prasad. 2024. Undocumented Immigrants at Work: Invisibility, Hypervisibility, and the Making of the Modern Slave. *Humanities and Social Sciences Communications* 11: 45. [CrossRef]
- Siegel, Rebecca L., Tyler B. Kratzer, Angela N. Giaquinto, Hyuna Sung, and Ahmedin Jemal. 2025. Cancer Statistics, 2025. *CA: A Cancer Journal for Clinicians* 75: 10–45. [CrossRef] [PubMed]
- Steen, Tomoko. n.d. Genetic Genealogy: DNA and Family History: Introduction. Research Guides, Library of Congress. Available online: <https://guides.loc.gov/genetic-genealogy/introduction> (accessed on 21 March 2026).
- Truong, Buu, Xuan Zhou, Jisu Shin, Jiuyong Li, Julius H. J. van der Werf, Thuc D. Le, and S. Hong Lee. 2020. Efficient Polygenic Risk Scores for Biobank Scale Data by Exploiting Phenotypes from Inferred Relatives. *Nature Communications* 11: 3074. [CrossRef] [PubMed]
- Valdez, Rodolfo, Paula W. Yoon, Nadeem Qureshi, Ridgely Fisk Green, and Muin J. Khoury. 2010. Family History in Public Health Practice: A Genomic Tool for Disease Prevention and Health Promotion. *Annual Review of Public Health* 31: 69–87. [CrossRef] [PubMed]
- Wolinetz, Carrie D., and Francis S. Collins. 2020. Recognition of Research Participants' Need for Autonomy: Remembering the Legacy of Henrietta Lacks. *JAMA* 324: 1027–28. [CrossRef] [PubMed]
- Zaretsky, Alexandra. 2021. DNA Collection in Immigration Custody and the Threat of Genetic Surveillance. *California Law Review* 109: 317–54.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.