A comparison of users' and non-users' perceptions of health and ancestry at-home DNA testing

Khadija Baig
khadija.baig@carleton.ca
Carleton University
Ottawa, Canada

Daniela Napoli
daniela.napoli@carleton.ca
Carleton University
Ottawa, Canada

Sonia Chiasson
chiasson@scs.carleton.ca
Carleton University
Ottawa, Canada

ABSTRACT

Direct-to-consumer (at-home) DNA testing allows users to gain ancestry and health information. Previous research has found users to be unconcerned about privacy relating to at-home DNA testing, with incomplete understanding of the process. The shared nature of DNA amongst biological relatives is not always considered by users, and the consequences may be underestimated. We ask: how do individuals perceive DNA ownership, and what do they see as plausible consequences of sharing this data? Are individuals equally concerned about at-home ancestry and health DNA testing? Do these perceptions differ between users and non-users? To investigate these questions, we conducted an online survey with 310 participants. Through statistical analysis, we compare responses for ancestry and health data, and for users and non-users. We discuss our results, their implications, and highlight areas for potential future research.

CCS CONCEPTS

• **Security and privacy → Human and societal aspects of security and privacy**.

KEYWORDS

at-home DNA testing, genes, privacy, human-computer interaction, HCI

1 INTRODUCTION

At-home DNA testing services, like those provided by Ancestry DNA or 23AndMe, provide users the ability to learn more about their ancestry or health [1, 6]. The use of these services has grown exponentially since 2013 [46]. Common reasons disclosed by users for taking these at-home DNA tests include seeking a connection with biological relatives, seeking to understanding more about personal health, and general curiosity of their ancestry [5, 9, 48].

However, the identifiable nature of DNA, combined with unclear privacy policies from at-home DNA testing companies [41], present a series of privacy risks. Cases of misuse have surfaced in the past (e.g., discrimination against someone with certain health markers in their DNA [59]), and this sensitive data continues to be shared on social media [40]. Furthermore, third parties (e.g., insurance companies [40]) with whom the data is shared may also have incentive to use this data in ways that are detrimental to the user. The potential for privacy violations can also extend far beyond the individual taking the test, possibly revealing secrets from ancestors long deceased, with varying degrees of accuracy. By sharing their own DNA, individuals are also sharing partial DNA of their biological relatives (including descendants yet to be born) [14], which creates unique challenges in networked genetic privacy [21].

Given the privacy risks related to at-home DNA testing, we seek to understand more about individuals' perceptions of the topic. Research indicates that users find many benefits in this service [7, 24, 27], with some privacy concerns revolving around control of their data, possible misuse, and inaccuracy of test results [26, 32, 35]. Many users do not consider privacy in their decision to complete an at-home DNA test [48], and may have an incomplete and inconsistent understanding of genetic privacy [7, 48]. Non-users' understanding may also vary highly depending on their degree of interest in the topic [27]. Literature comparing user and non-user perceptions is, however, limited. Additionally, the information provided by ancestry DNA testing differs greatly from that of health DNA testing. Despite this, we find no literature addressing whether individuals distinguish between ancestry and health data in their perceptions of genetic privacy. We formulate our two research questions based on this gap in the literature: ***RQ1:*** *How do privacy perceptions of at-home DNA testing differ based on whether it is for ancestry or health purposes?* and ***RQ2:*** *How do users' and non-users' privacy perceptions of at-home DNA testing differ?*

We conducted a 2×2 between-subjects study, using a 15-minute survey to explore more fine-grained perceptions of genetic privacy. We surveyed users and non-users (defined as individuals who have explicitly chosen not to take an at-home DNA test), thereby getting opinions from individuals who had thought about the tests and taken a decision one way or the other. We compare whether participants perceive risks differently for ancestry and health-related DNA testing, and whether users and non-users respond differently. Our research finds few differences between perceptions of ancestry and health data, but many differences between users and non-users. We discuss participant misconceptions highlighted by our data and

associated risks. Furthermore, we propose privacy-enhancing mitigations to address gaps in usable privacy for users trying to learn how their data is used, as well as for non-users seeking knowledge about the DNA testing process.

2 BACKGROUND

Prior to at-home DNA testing kits, completing a DNA test was rare and usually only prescribed by medical professionals [4]. Traditionally done using a saliva sample, at-home *ancestry* testing provides users the opportunity to meet biological relatives, trace their past [6], and access genealogical databases [38]. At-home *health* testing claims to assess disposition to certain diseases [1, 4, 29]. Sales from private DNA testing companies declined in mid-2019 [18], and throughout 2020 [19, 20]. Many major at-home DNA testing companies pledged to follow guidelines designed by privacy advocacy groups [39] amidst rising privacy concerns during this time. Even so, the at-home DNA testing market is projected to exceed $10 billion USD by 2028 [47].

2.1 User Attitudes

Users perceive many benefits to at-home DNA testing [7, 24, 48], and are interested in expanded testing services [23]. Some users complete these tests to contribute their DNA for research purposes [11, 56, 57], while others perceive a more recreational use for it [5, 7, 56].

Ancestry tests are often used to unite adopted children with their biological relatives [9], reunite families separated over generations or as refugees [16], and make progress in solving staggered cases of violent crimes [31]. It has also been considered for, and used in, supporting evidence that a user belongs to an ethnic minority group [30, 60]. Many are also interested in *health-related* testing to obtain information about their traits and likelihood of developing certain diseases or health conditions [48]. Many complete a test with the intent of sharing their results with a physician [22].

Users have expressed concerns about at-home DNA testing related to privacy, unwanted information, and accuracy [7, 13, 24, 26, 32, 35]. However, in a study conducted by Roberts et al., 40% of participants had not considered the negative effects of taking an at-home DNA test prior to completing one [48]. Furthermore, users may not understand the sensitivity of their DNA data [7, 24]. Many also trust the protections of DNA-testing companies and current legislature to protect them (and their data) from misuse [7, 24, 34]. The literature also suggests feelings of resignation. Users have indicated a desire for control over their data [7, 23], but they are also cynical on whether there is any privacy at all [26].

2.2 Potential Risks

The benefits associated with at-home DNA testing are countered with several privacy risks. Furthermore, these tests could possibly be misleading [12]. The accuracy behind health-related at-home DNA tests is questionable [52], with moderate to high risk of false positives [54]. Studies also show that test results do not motivate behavioural changes in users [25].

Additionally, at-home DNA testing may result in incorrect genetic inferences for a population [58]. DNA can also reveal information about individuals' *phenotype* [29] (physical traits observable

to the human eye, for example: eye, hair, and skin colour), possibly making them identifiable in real life. This also extends to facial features, which can be predicted by genome data with as few as six relevant genes [44].

The potential for privacy violations can extend far beyond the individual taking the test. By sharing their own DNA, individuals are also sharing the DNA of their past biological relatives, those who are living, and those yet to be born [14]. This creates unique challenges in networked genetic privacy [21].

Limited research exists on users' understanding of the networked nature of DNA. Grandhi et al. [24] found that some users acknowledged that doing a DNA test could put their family at risk from law enforcement, or from a future, currently non-existent, risk [24]. In contrast, others expressed entitlement to others' data for the greater good (for example, in the case of family-tree building) or had not considered their family members' privacy [7].

2.3 Non-user Attitudes

Based on the existing literature, non-user attitudes vary depending on how interested they are in at-home DNA testing. Our review uncovered one study exploring the specific perspectives of non-users regarding at-home DNA testing. Using focus groups with 62 participants with varied interest, Hazel et al. [27] found a wide range of opinions. Some non-users thought of ancestry tests as quick and fun, while others acknowledged them as valuable tools [27]. Many participants did, however, have concerns regarding control and misuse of data that might harm them or their families [27]. Ruhl et al. [49] surveyed 1000 participants to understand perceptions of at-home DNA testing. Most (80%) were non-users. No comparisons were made between users and non-users. Around half of all participants surveyed by Ruhl et al. [49] had privacy concerns surrounding genetic testing.

Grandhi et al. [24] compared the privacy concerns of users, non-users who were willing to take an at-home DNA test, and non-users who were unwilling to complete an at-home DNA test. Those who are unwilling to take a DNA test had the highest average level of privacy and security concerns, while existing users had the lowest level. Some non-users did, however, trust in the DNA testing companies' privacy protections and existing legal protections. Even so, many more expressed concerns about data misuse, storage, and sharing.

3 METHODOLOGY

In this 2×2 between-subjects study, we examine the differences in privacy attitudes and perceptions between users and non-users and between health and ancestry data. Our study was cleared by our university's Research Ethics Board. We formulated our survey using an iterative design process and pilot tested the completed survey with two participants per condition to finalise its design.

3.1 Recruitment and Participants

We recruited participants using the online crowdsourcing platform Prolific [43], which is dedicated specifically to research studies. The surveys themselves were hosted on Qualtrics Survey Software [45]. Participants completed a 1-minute pre-screener to assess eligibility for the main survey. To be eligible for the pre-screener, participants

Table 1: Number of participants who completed each version of the survey (after cleaning of the data).

Study Condition			# of Participants
UA	User	Ancestry	76
NUA	Non-User	Ancestry	85
UH	User	Health	64
NUH	Non-User	Health	85
Total			310

needed to reside in Canada and be over 18. Participants were paid according to the rate recommended by Prolific at the time: 0.13 GBP for the pre-screener, and 1.88 GBP for the main survey.

In total, 998 participants were pre-screened. The pre-screener asked participants to identify whether they had completed at-home DNA testing for ancestry or for health purposes. After cleaning the pre-screening data, 190 pre-screening participants had previously used at-home DNA testing and all were invited to complete the appropriate version of our survey. Health users (UH) completed the health version, ancestry users (UA) completed the ancestry version; anyone indicating both were assigned to health because we had fewer health respondents. A random selection of 200 non-users from the pre-screened pool were pseudo-randomly assigned to either survey condition to ensure 100 participants per condition. From the 390 (i.e., 190 users + 200 non-users) invitees, a total of 328 participants completed the main survey.

We excluded 18 responses, 13 because participants failed one of the attention checks, 5 because there were other indicators of unreliable data (e.g., straight-lining). While the process of data exclusion was subjective, the number of surveys removed was small enough that we were able to err on the side of caution. Table 1 summarizes participants per condition in the main survey after data cleaning.

A detailed breakdown of participant demographics for our main survey can be seen in Appendix A. 56% of participants were men, 42% were women and 1.3% were of other genders. Approximately half of our participants were aged 25 – 34, but ages ranged from 18 – 74. Similarly, approximately half of participants had undergraduate degrees, and the most common occupations were in Science, Engineering (20% of all participants).

3.2 Survey Design

Each survey began with a consent form, and all questions contained a "prefer not to answer" option. At the end of each survey, participants were once again provided the opportunity to withdraw their data. As is the norm on Prolific, only participants who completed the surveys were paid. Survey materials are available in Appendix C.

Pre-screener: In the pre-screener we asked participants about their general geographic location, and whether they had completed and received the results of an at-home DNA test. If so, they were asked whether the test was for ancestry or health purposes (or both), and to indicate the DNA testing company from a provided list. If not, we asked about their comfort with completing an at-home DNA test. Those who indicated some level of discomfort were asked whether they had explicitly chosen not to complete a DNA

test; those who chose not to complete a test were included in our study. The pre-screener took approximately 1 minute to complete.

Main survey: Our survey consisted of 145 questions (excluding attention checks), which draw from existing literature [7, 24, 26, 32]. The questions were organised in 11 blocks as described in Table 2.

Non-demographic questions were 4-point Likert scale questions, with the exception of the Effects (3-point scale) and Scenarios (binary scale) blocks. The order of questions within each block was randomised. Prompts in the Scenarios block were based on actual incidents (e.g. [10, 15, 37, 42, 59]). The main survey took approximately 15 minutes to complete.

We used inferential statistics to look for differences between users and non-users (*UA versus NUA, UH versus NUH*) and between ancestry or health-related DNA data (*UA versus UH, NUA versus NUH*). We perform between-subject comparisons (unpaired groups). We do not assume normality, and so conduct non-parametric tests. The Likert-scale options *within* each question block were identical but differed *between* blocks. Therefore, we used the *Holm-Bonferroni* family-wise adjustment per block (family) of questions rather than experiment-wise.

Blocks 2 to 10 consisted of Likert scales, so we conducted *Mann-Whitney U tests* ($p < 0.05$, two-tailed) on the individual questions, given the ordinal nature of their data. The Scenario block consisted of binomial data (*Yes, No*), so we performed *Fisher's exact test*.

4 RESULTS

We have grouped our question blocks into four themes. For each theme, we present the relevant question blocks, discuss any overall trends in the data across conditions, and then we identify the main results relating to our two research questions: RQ1 – comparing ancestry and health conditions, and RQ2 – comparing user and non-user conditions. Detailed statistical results are available in the accompanying figures, but for readability, we focus on the main results in the text. The Likert scale options were 4-point unless otherwise indicated.

4.1 The shared value of DNA

We first discuss how participants perceive the value of their data: what it can do for them and for others. This section includes results from the following question blocks:

- In the *Data Sharing* block, participants rated how comfortable they felt sharing their data with thirteen entities which could be viewed as third-parties.
- In the *Data Usage* block, participants rated their comfort with their data being used for seven different purposes. Most use cases involved various entities completing research using DNA data.
- In the *Contributions* block, participants rated how much they would like *other* DNA testing users to volunteer their own data for these same seven use cases.
- In the *Effects* block, participants rated how their own DNA test would affect each of ten entities on a 3-point scale: positively, negatively, or no effect. The entities included individuals in their personal life (e.g.: biological relatives, colleagues) and entities from other parts of their life (e.g.: insurance rates).

Table 2: Survey organization. N = number of questions.

	Question Block	Description	N
1.	Demographic Info	Age, gender, geographic location	6
2.	Control	Who controls or manages their submitted DNA data?	8
3.	Access	Who has access to their submitted DNA data?	13
4.	Deletion	What happens if they delete their DNA data?	7
5.	Concerns	How concerning are given risks surrounding testing?	9
6.	Data sharing	With whom do participants feel comfortable sharing their DNA data?	13
7.	Data usage	For which reasons are participants comfortable with their DNA data being used?	7
8.	Contributions	Under which circumstances should *other* users of at-home DNA tests share their DNA data?	7
9.	General attitudes	What are participants' general perceptions of genetic privacy (including DNA ownership)?	11
10.	Effects	Who does their at-home DNA test affect?	10
11.	Scenarios	How plausible and likely are given scenarios involving at-home DNA testing? [21 scenarios, 3 questions each]	63

Table 3 presents medians per condition and results from the inferential statistics. Overall, participants prefer sharing for altruistic purposes. Across conditions, participants were generally comfortable sharing their data with the *DNA company*, *university researchers*, and *non-profit researchers*, but not other entities. Many participants were somewhat comfortable with *helping improve the DNA testing company's services*, or with *academic* or *non-profit research* but generally uncomfortable with their data being used for *government* and *law-enforcement research* (and *investigations*).

RQ1 [Shared value]: Health users are more likely than ancestry users to want others to volunteer their data, and perceive a positive effect on their government benefits.

Contributions. We found statistically significant differences showing that health users (UH) were more likely than ancestry users (UA) to want others to contribute their data for *police investigations*, *government research*, *academic research*, and *research by the police*.

Effects. Only one question was statistically significant. Health users (UH) were more likely than ancestry users (UA) to perceive positive effects on *their government benefits*.

RQ2 [Shared value]: Users are more at ease sharing their data, more likely to think others should share too, and perceive more positive effects of sharing.

Data Sharing. Analysis shows statistically significant differences where users (UA, UH) were more comfortable sharing their data with *all thirteen listed entities* than non-users (NUA, NUH). Statistically significant differences were also found between users and non-users in relation to data usage. Health users (UH) were more comfortable than non-users (NUH) with their data being used for every listed purpose. Ancestry users (UA) were significantly more comfortable with their data being used: *by the DNA-testing company for research*, *by law enforcement for police investigations*, and by *pharmaceutical companies for research*.

Contribution of others. Analysis shows statistically significant differences between users and non-users, with users voicing stronger agreement that others should volunteer their data for various purposes. Non-users had little to no interest in this type of sharing by others. We found statistically significant differences between health users (UH) and non-users (NUH) for each of the seven purposes. In each case, users (UH) were more likely to want others to contribute their data than non-users (NUH). Only two of seven

questions showed statistically significant differences between ancestry users (UA) and non-users (NUA), with users reporting higher agreement.

Effects. Overall, participants perceived very little effect of their DNA test on most entities. Where statistically significant differences occurred, users were more likely to perceive positive effects and non-users were more likely to perceive negative effects. Health users (UH) perceived significantly more positive effects than non-users (UH) on *themselves*, *their parents*, *their future children*, and no effects (compared to negative) on their *their employment prospects*, *their insurance rates*, and *their government benefits*. Ancestry users (UA) also perceived more positive effects than non-users (NUA) from using DNA tests on *themselves*, *friends*, *future children*, and *their colleagues*), as well as on their *insurance rates*.

4.2 Potential for negative consequences

In the following section, we discuss participants' perspectives towards potential negative consequences resulting from their test. We include the following blocks of questions:

- In the *Access* block, participants rated the likelihood that each of 13 entities had access to their data.
- In the *Concerns* block, participants rated their level of concern for each of nine potential at-home DNA testing related concerns identified from previous research. Several concerns presumed that certain entities from the *Access* block had partial or complete access to ancestry and/or health data.
- In the *Scenarios* block, participants were presented twelve different scenarios that were either based on a real situation that had occurred within Canada, or based on existing concerns noted in relating literature. Participants answered whether the scenario was: (1) possible, (2) likely to happen in Canada, and (3) likely to happen anywhere in general. Figure 1 displays a heatmap summarizing responses and Figure 5 summarizes results of the inferential statistics tests.

Table 4 summarizes the medians and statistical results relating to the *Data access* and *Concerns* blocks. Across conditions, the majority of participants thought that the *DNA-testing company* had access to users' DNA data. Participants further thought it likely that many other entities also had access to their data. Many participants expressed apprehension over the listed concerns, with over half

Data sharing: How comfortable are you sharing your ancestry/health data with:	Medians				UA-UH		NUA-NUH		UA-NUA		UH-NUH	
	UA	NUA	UH	NUH	U	p	U	p	U	p	U	p
The DNA company itself	2	3	2	3	2225	1.000	3236	1.000	*4795*	*<0.001*	*4091*	*<0.001*
The general public	4	4	4	4	2116	1.000	3397	1.000	*4215*	*<0.001*	*3368*	*0.003*
Other users of the at-home DNA testing company	3	4	3	4	2044	1.000	3336	1.000	*4388*	*<0.001*	*3703*	*<0.001*
Your employer	4	4	3	4	2500	1.000	3787	1.000	*4258*	*<0.001*	*3634*	*<0.001*
The government	4	4	3	4	2743	1.000	3588	1.000	*3767*	*0.038*	*3562*	*0.002*
Law enforcement	4	4	3	4	2590	1.000	3711	1.000	*3856*	*0.035*	*3376*	*0.007*
Pharmaceutical companies	3	4	3	4	2858	0.497	3857	1.000	*3922*	*0.024*	*3588*	*0.002*
Insurance companies	4	4	3	4	3029	0.086	3873	1.000	*3985*	*0.007*	*3852*	*<0.001*
Advertisers	4	4	4	4	2332	1.000	3646	1.000	*4159*	*<0.001*	*3423*	*0.001*
For-profit companies	4	4	3	4	2454	1.000	3459	1.000	*4228*	*<0.001*	*3867*	*<0.001*
University researchers	2	3	2	3	2639	1.000	3758	1.000	*3895*	*0.038*	*3431*	*0.007*
Non-profit researchers	2	3	2	3	2688	1.000	3880	1.000	*3992*	*0.022*	*3475*	*0.007*
Investors in the DNA testing company	4	4	3	4	2741	1.000	3408	1.000	*3980*	*0.007*	*3953*	*<0.001*

Data usage: How comfortable are you with your data being used:	Medians				UA-UH		NUA-NUH		UA-NUA		UH-NUH	
	UA	NUA	UH	NUH	U	p	U	p	U	p	U	p
By the DNA-testing company to improve their services.	2	3	2	3	2592	1.000	3322	1.000	*4241*	*0.001*	*3975*	*<0.001*
By the government for research.	3	3	2	3	2794	0.690	3807	1.000	3927	0.054	*3623*	*0.001*
By law-enforcement for research.	3	4	2	4	2887	0.337	3396	1.000	3799	0.084	*3863*	*<0.001*
By law-enforcement for police investigations.	3	4	3	4	2683	1.000	3750	1.000	*4071*	*0.011*	*3643*	*0.001*
By pharmaceutical companies for research.	2	3	2	3	2611	1.000	3431	1.000	*3977*	*0.041*	*3576*	*0.001*
By academic institutions for research.	2	2	2	2	2733	0.851	3477	1.000	3743	0.084	*3542*	*0.002*
By non-profit organisations for research.	2	2	2	2	2483	1.000	3545	1.000	3847	0.084	*3269*	*0.019*

Contributions: Rate how much you would like others who have also completed the test to contribute their ancestry/health-related DNA data for the following purposes:	Medians				UA-UH		NUA-NUH		UA-NUA		UH-NUH	
	UA	NUA	UH	NUH	U	p	U	p	U	p	U	p
For research by the DNA-testing company to improve their services.	2	3	2	3	2397	0.877	3721	1.000	*4812*	*<0.001*	*3910*	*<0.001*
For research by the government.	3	3	2	4	*3098*	*0.021*	3461	1.000	3373	0.417	*3718*	*<0.001*
For research by the police.	4	4	3	4	*2998*	*0.046*	3695	1.000	3750	0.121	*3673*	*<0.001*
For police investigations.	4	4	3	4	*3070*	*0.023*	3457	1.000	3574	0.369	*3772*	*<0.001*
For research by pharmaceutical companies.	3	3	2	3	2872	0.115	3741	1.000	*4107*	*0.011*	*3684*	*<0.001*
For research by academic institutions.	2	2	1	2	*3152*	*0.010*	3945	1.000	3910	0.085	*3761*	*<0.001*
For research by non-profit organisations.	2	3	2	3	2921	0.101	4111	0.743	3889	0.085	*3402*	*0.007*

Effects: Rate the overall effect your ancestry/health genetic test results would have on each of the following entities:	Medians				UA-UH		NUA-NUH		UA-NUA		UH-NUH	
	UA	NUA	UH	NUH	U	p	U	p	U	p	U	p
Yourself	1	2	1	1	2631	1.000	3302	1.000	*4107*	*0.002*	*3572*	*<0.001*
Your employment prospects	2	2	2	2	2501	1.000	2859	0.083	3513	0.090	*3661*	*<0.001*
Your siblings	2	2	1	2	2578	1.000	3601	1.000	3560	0.290	2924	0.396
Your friends	2	2	2	2	2335	1.000	3686	1.000	*3747*	*0.011*	2999	0.377
Your colleagues	2	2	2	2	2322	1.000	3518	1.000	*3573*	*0.031*	2881	0.396
Your insurance rates	2	2	2	3	2784	0.197	*2604*	*0.021*	*3815*	*0.011*	*4118*	*<0.001*
Your parents	2	2	2	2	2345	1.000	3049	1.000	3596	0.290	*3237*	*0.028*
Your potential government benefits	2	2	2	2	*2758*	*0.047*	3572	1.000	3539	0.290	*3258*	*0.012*
Current, existing children	2	2	1.5	2	2358	1.000	3625	1.000	3403	0.290	2507	0.513
Future children	1	2	1	1	2800	0.197	3887	0.661	*3938*	*0.031*	*3130*	*0.018*

Table 3: *Data sharing, Data usage, Contributions* (4-point Likert scale, 1: most positive, 4: most negative), and *Effects* (3-point scale) question blocks. Medians per condition and the Holm-Bonferroni adjusted statistical results of Mann-Whitney U tests. Green cells with italics indicate statistically significant results.

Data access: Who has access to your most recent ancestry/health data?	Medians				UA-UH		NUA-NUH		UA-NUA		UH-NUH	
	UA	NUA	UH	NUH	U	p	U	p	U	p	U	p
The DNA company itself	1	1	1	1	2104	0.634	3459	1.000	3209	1.000	2409	1.000
The general public	4	4	4	4	2426	1.000	3003	0.279	2977	0.945	2951	1.000
Other users of the at-home DNA testing company	3	4	4	4	1948	0.410	2976	0.401	3128	1.000	2703	1.000
Your employer	4	4	4	4	2821	0.356	3832	1.000	2854	0.803	2785	1.000
The government	3	3	3	2	2886	0.410	3330	1.000	2484	0.079	2747	1.000
Law enforcement	3	3	3	2	2710	0.800	3351	1.000	*2159*	*0.003*	2323	1.000
Pharmaceutical companies	3	3	3	2	2893	0.410	3873	1.000	*2381*	*0.026*	2286	1.000
Insurance companies	3	3	3	3	2990	0.177	3685	1.000	*2358*	*0.021*	2581	1.000
Advertisers	3	3	3	3	2574	1.000	3602	1.000	2719	0.397	2435	1.000
For-profit companies	3	3	3	3	2486	1.000	3067	1.000	2649	0.318	2645	1.000
University researchers	3	2.5	2.5	2	2950	0.269	3605	1.000	*2160*	*0.002*	2413	1.000
Non-profit researchers	3	2	2	3	2827	0.569	3547	1.000	2650	0.318	2723	1.000
Investors in the DNA testing company	3	3	3	3	2419	1.000	3233	1.000	2839	0.850	2799	1.000

Concerns: In your own life, please rate your level of concern with each of the following:	Medians				UA-UH		NUA-NUH		UA-NUA		UH-NUH	
	UA	NUA	UH	NUH	U	p	U	p	U	p	U	p
Your ancestry/health results influencing your job prospects.	3	3	3	2	2699	0.512	4155	0.309	2671	0.291	2096	0.192
Finding out unwanted ancestry/health information	3	3	2	2	*3161*	*0.006*	4359	0.088	2897	0.587	2600	1.000
Finding unwanted family secrets as a result of an ancestry DNA test.	3	3	2	3	2944	0.062	3838	1.000	3185	0.981	3086	0.385
Surveillance due to an ancestry/health DNA test.	2	2	2	2	2412	0.994	3951	1.000	2689	0.291	*1929*	*0.017*
Accuracy of ancestry/health at-home DNA tests.	2	2	2	2	2921	0.062	3526	1.000	2536	0.124	2633	1.000
Genetic discrimination as a result of an ancestry/health DNA test.	3	2	2	2	2869	0.062	3707	1.000	*1990*	*<0.001*	2195	0.385
Ancestry/health information from a DNA test relevant to your immediate or extended family being public	2	2	2	2	2602	0.994	3524	1.000	*2453*	*0.047*	2323	0.446
Adjusted insurance rates due to an ancestry/health DNA test.	2	2	2	2	2614	0.994	3645	1.000	*2287*	*0.018*	*1891*	*0.011*
Your ancestry/health data being used for profit.	2	1	2	1	2820	0.266	3222	1.000	*1738*	*<0.001*	2250	0.385

Table 4: *Data access* and *Concerns* question blocks (4-point Likert scale, 1: most positive, 4: most negative). Medians per condition and the Holm-Bonferroni adjusted statistical results of Mann-Whitney U tests. Green cells with italics indicate statistically significant results.

of participants choosing *concerned* or *very concerned* on nearly all questions.

As shown in Figure 1 and Table 5, participants across all conditions believed most provided scenarios were plausible. While many scenarios were perceived likely to happen in general, many thought they were less likely to happen in Canada.

RQ1 [Consequences]: Health participants are more likely than ancestry participants to be concerned about, and perceive the possibility of, some negative consequences. We found only a few statistically significant differences between health and ancestry conditions. Where differences occurred, participants in the health conditions (UH, NUH) were generally more aware of, concerned about, and more likely to believe the likelihood of some negative consequences than participants in the ancestry conditions (UA, NUA).

Concerns. Statistical analysis showed that health users (UH) were significantly more concerned than ancestry users (UA) about finding out unwanted information from a DNA test.
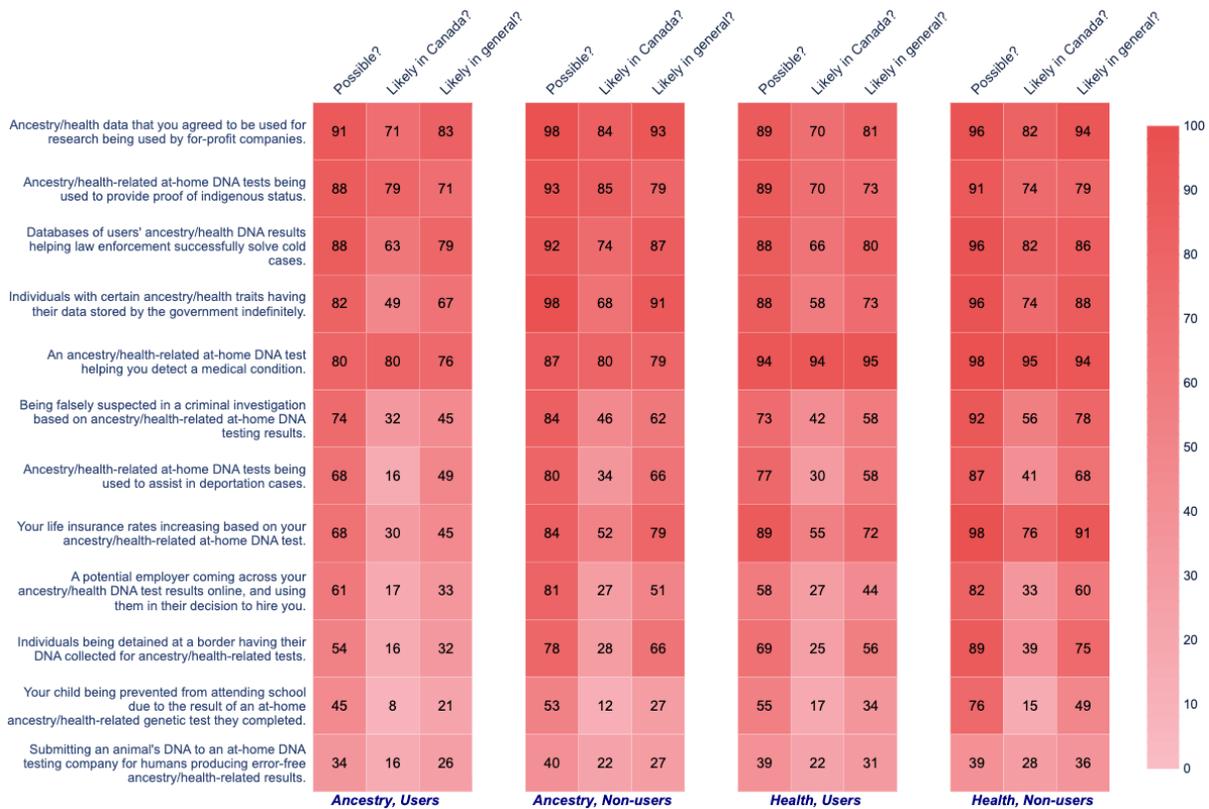
Scenarios. Where statistically significant differences exist, participants in the health conditions (UH, NUH) were more likely to

perceive negative scenarios as either possible or likely. Participants in the health conditions (UH, NUH) believe that it is possible, and likely, that *life insurance rates would increase based on their at-home DNA test.* Health users (UH) were also more likely than ancestry users (UA) to believe that it is likely that *individuals being detained at a border have their DNA collected for health tests.*

Additionally, significantly more health non-users (NUH) than ancestry non-users (NUA) believe that *their child could be prevented from attending school due to results of an at-home genetic test.* They also think it more likely that *those with certain health traits would have their data stored by the government indefinitely.*

Statistically significant results also show that health non-users (NUH) were more likely than ancestry non-users (NUA) to perceive that *at-home DNA tests can be used to provide proof of Indigenous status*, that their DNA test could *help them detect a medical condition*, and that *submitting animal DNA to an at-home DNA testing company for humans would produce error-free results* within Canada.

RQ2 [Consequences]: Non-users are more concerned than users about some negative consequences and think these are more likely.

Scenario	Ancestry, Users			Ancestry, Non-users			Health, Users			Health, Non-users		
	Possible?	Likely in Canada?	Likely in general?	Possible?	Likely in Canada?	Likely in general?	Possible?	Likely in Canada?	Likely in general?	Possible?	Likely in Canada?	Likely in general?
Ancestry/health data that you agreed to be used for research being used by for-profit companies.	91	71	83	98	84	93	89	70	81	96	82	94
Ancestry/health-related at-home DNA tests being used to provide proof of indigenous status.	88	79	71	93	85	79	89	70	73	91	74	79
Databases of users' ancestry/health DNA results helping law enforcement successfully solve cold cases.	88	63	79	92	74	87	88	66	80	96	82	86
Individuals with certain ancestry/health traits having their data stored by the government indefinitely.	82	49	67	98	68	91	88	58	73	96	74	88
An ancestry/health-related at-home DNA test helping you detect a medical condition.	80	80	76	87	80	79	94	94	95	98	95	94
Being falsely suspected in a criminal investigation based on ancestry/health-related at-home DNA testing results.	74	32	45	84	46	62	73	42	58	92	56	78
Ancestry/health-related at-home DNA tests being used to assist in deportation cases.	68	16	49	80	34	66	77	30	58	87	41	68
Your life insurance rates increasing based on your ancestry/health-related at-home DNA test.	68	30	45	84	52	79	89	55	72	98	76	91
A potential employer coming across your ancestry/health DNA test results online, and using them in their decision to hire you.	61	17	33	81	27	51	58	27	44	82	33	60
Individuals being detained at a border having their DNA collected for ancestry/health-related tests.	54	16	32	78	28	66	69	25	56	89	39	75
Your child being prevented from attending school due to the result of an at-home ancestry/health-related genetic test they completed.	45	8	21	53	12	27	55	17	34	76	15	49
Submitting an animal's DNA to an at-home DNA testing company for humans producing error-free ancestry/health-related results.	34	16	26	40	22	27	39	22	31	39	28	36

Figure 1: A heatmap with the percentages of participants who responded "yes" to whether each scenario was (a) possible, (b) likely in Canada, and (c) likely in general. Darker cells represent higher agreement.

Access. We found statistically significant differences for four (out of 13) entities who may have access to DNA data: ancestry non-users (NUA) were significantly more likely than ancestry users (UA) to think that *law enforcement, pharmaceutical companies, insurance companies* and *university researchers* had access to their data.

Concern. Analysis shows statistically significant differences between users and non-users for several listed concerns. Ancestry non-users (NUA) were more concerned than ancestry users (UA) about *discrimination*, their *family data being public*, the adjustment of *insurance rates*, and the use of their data for *profit*. Health non-users (NUH) were more concerned than users (UH) about *surveillance*, and the adjustment of *insurance rates*.

Scenarios. We found statistically significant differences between users and non-users for only a few scenarios; in each case, non-users were more likely to think that the given scenarios were possible and likely to occur.

Considering whether scenarios were possible, we found statistically significant differences where ancestry non-users (NUA) were more likely than users (UA) to believe that *individuals being detained at a border could have their DNA collected* and that *individuals with certain traits could have their data stored by the government indefinitely*. Statistical analysis also showed that health non-users (NUH) were significantly more likely than users (UH) to believe that *individuals being detained at a border could have their DNA collected* for health-related tests, that *a potential employer could come across a candidate's health DNA test results online, and use them in their hiring decision*, and that it was possible to be *falsely suspected in a criminal investigation* based on health at-home DNA testing results.

Considering the likelihood of scenarios happening in Canada, we found statistically significant differences where ancestry non-users (NUA) were more likely than users (UA) to think that at-home ancestry DNA tests could be used to *provide proof of Indigenous status*, that *those with certain ancestry traits could have their data stored by the government indefinitely*, that one could be *falsely suspected in a criminal investigation* based on DNA test results, and that *submitting an animal's DNA to an at-home DNA testing company for humans would produce error-free ancestry results*.

Considering the likelihood of scenarios happening in general, we found only four instances where there were statistically significant differences. Ancestry non-users (NUA) were more likely to believe that *individuals detained at a border have their DNA collected for ancestry tests*, that an ancestry test would *increase their life insurance rates*, and that *individuals with certain ancestry traits could have their data stored by the government indefinitely*. Health non-users (NUH) were more likely than users (UH) to think that *life insurance rates increase based on your health DNA test results*.

Scenarios	Is this possible?				Is this likely in Canada?				Is this likely in general?			
	UA-UH	NUA-NUH	UA-NUA	UH-NUH	UA-UH	NUA-NUH	UA-NUA	UH-NUH	UA-UH	NUA-NUH	UA-NUA	UH-NUH
Individuals being detained at a border having their DNA collected for ancestry/health tests.	0.843	0.552	*0.027*	*0.032*	1.000	1.000	0.349	0.733	*0.038*	1.000	*<0.001*	0.194
Ancestry/health at-home DNA tests being used to assist in deportation cases.	1.000	1.000	0.736	0.507	0.656	1.000	0.075	1.000	1.000	1.000	0.262	1.000
A potential employer coming across your ancestry/health DNA test results online, and using them in their decision to hire you.	1.000	1.000	0.050	*0.019*	1.000	1.000	0.553	1.000	1.000	1.000	0.235	0.475
Your life insurance rates increasing based on your ancestry/health at-home DNA test.	*0.048*	*0.030*	0.243	0.311	0.068	*0.012*	0.053	0.095	*0.021*	0.483	*<0.001*	*0.050*
Your child being prevented from attending school due to the result of an at-home ancestry/health genetic test they completed.	1.000	*0.026*	1.000	0.071	1.000	1.000	0.886	1.000	0.796	0.052	1.000	0.565
Ancestry/health data that you agreed to be used for research being used by for-profit companies.	1.000	1.000	0.683	0.500	1.000	1.000	0.309	0.799	1.000	1.000	0.326	0.189
Ancestry/health at-home DNA tests being used to provide proof of Indigenous status.	1.000	1.000	1.000	1.000	1.000	*<0.001*	*<0.001*	1.000	1.000	1.000	1.000	1.000
An ancestry/health at-home DNA test helping you detect a medical condition.	0.279	0.177	1.000	1.000	0.279	0.306	0.886	1.000	*0.021*	0.067	1.000	1.000
Databases of users' ancestry/health DNA results helping law enforcement successfully solve cold cases.	1.000	1.000	1.000	0.391	1.000	1.000	0.133	0.247	1.000	1.000	1.000	1.000
Submitting an animal's DNA to an at-home DNA testing company for humans producing error-free ancestry/health results.	1.000	1.000	1.000	1.000	1.000	*<0.001*	*<0.001*	1.000	1.000	1.000	1.000	1.000
Individuals with certain ancestry/health traits having their data stored by the government indefinitely.	1.000	1.000	*0.011*	0.391	1.000	*<0.001*	0.005	0.521	1.000	1.000	*0.003*	0.244
Being falsely suspected in a criminal investigation based on ancestry/health at-home DNA testing results.	1.000	1.000	1.000	*0.033*	1.000	1.000	*<0.001*	0.793	1.000	0.440	0.235	0.134

Table 5: *Scenarios* block: results of Fisher's Exact tests (1: Yes, 2: No). Green cells with italics indicate statistically significant results.

4.3 General attitudes towards DNA data

In this subsection, we discuss the *General Attitudes* block of questions. Participants rated their agreement with each of ten statements about DNA ownership, DNA sensitivity, and privacy risks. In general, participants across all conditions perceived high individual ownership over their data (Table 6). They were generally unconcerned with how others chose to share their data, but our results indicate that they considered their own data to be highly sensitive.

RQ1 [Attitudes]: We found no statistically significant differences between ancestry and health conditions for general attitudes.

RQ2 [Attitudes]: Where differences occurred, non-users expressed more privacy-conscious attitudes towards DNA data than users. Analysis showed statistically significant differences between users and non-users on six out of ten questions in this block, with non-users expressing significantly more privacy-conscious attitudes than users. Non-users (NUA and NUH) were more likely to think that their DNA data *was their own* and *was unlike other types of data*, and less likely to be *annoyed if others hid their DNA results*. Non-users were also less interested in *finding biological family* and saw little value in *contributing their data to the community*. Health non-users (NUH) were also more likely than users (UH) to think that they could *be identified through their DNA data*.

4.4 DNA data management

In this section, we consider participants' understanding of data permanence in relation to their DNA data and their understanding of how much control various entities have in managing their DNA data.

- In the *Deletion block*, we asked participants to assume that they deleted an account affiliated with at-home DNA testing. Participants then answered six questions related to what happens to their data in these circumstances.
- In the *Control* block, participants rated how each of eight entities were involved in controlling their DNA data.

Table 7 details findings for this section. Overall, participants across all conditions thought at least some of their data was likely to stay on the platform despite user actions to delete an account. Responses for the *Control* block were generally similar across conditions. Participants perceived the *at-home DNA-testing company* and the *owner of the DNA* have *some* or *all* control over the DNA data. Conversely, participants perceived *non-profit partners* or *other users* to have *little* to *no* control over the DNA data.

RQ1 [Management]: Health users are more likely than ancestry users to believe their DNA data will be deleted. Statistical analysis revealed only one statistically significant difference between the ancestry and health conditions. Health users (UH) were more likely than ancestry users (UA) to believe that *their DNA test results are deleted* with their account.

General attitudes: Rate how strongly you agree or disagree with each of the following statements	Medians				UA-UH		NUA-NUH		UA-NUA		UH-NUH	
	UA	NUA	UH	NUH	U	p	U	p	U	p	U	p
I can be identified by the DNA sample I provided for ancestry testing.	2	2	2	1	2350	1.000	3982	0.576	2679	0.521	*1879*	*0.006*
My family might find out things they didn't want to know.	2	2	3	2	2212	1.000	3297	1.000	2686	0.521	2081	0.105
My DNA is my information and my information alone.	2	1	2	1	2120	1.000	3278	1.000	*2297*	*0.009*	*1646*	*<0.001*
What I do with my DNA is my business.	1	1	2	1	2140	1.000	3211	1.000	2567	0.073	2117	0.105
My family have a right to be concerned about my ancestry DNA test.	2	2	2	2	2939	0.171	3609	1.000	3003	0.792	3021	0.264
I am interested in finding biological family with a DNA test.	2	3	2	3	2360	1.000	3351	1.000	*4274*	*<0.001*	*3825*	*<0.001*
I do not need anyone's consent to take an ancestry DNA test.	1	2	1.5	2	2199	1.000	3297	1.000	3423	0.521	2756	0.576
Making my ancestry DNA test results public adds to the community.	3	3	3	3	2087	1.000	3472	1.000	*4103*	*0.002*	*3342*	*0.006*
People who hide their ancestry DNA results annoy me.	3	4	3	4	2466	1.000	3841	1.000	*4275*	*<0.001*	*3422*	*0.002*
Ancestry DNA data is just like any other data.	3	3	3	3	2223	1.000	3706	1.000	*4309*	*<0.001*	*3407*	*0.014*

Table 6: *General attitudes* question block (4-point Likert scale, 1: most positive, 4: most negative). Medians per condition and the Holm-Bonferroni adjusted statistical results of Mann-Whitney U tests. Green cells with italics indicate statistically significant results.

Deletion: When you delete your account, how likely is it that:	Medians				UA-UH		NUA-NUH		UA-NUA		UH-NUH	
	UA	NUA	UH	NUH	U	p	U	p	U	p	U	p
Your raw DNA is deleted.	3	4	3	4	2887	0.126	3402	1.000	3760	0.092	*3756*	*<0.001*
Your personally identifiable information is deleted (name, email address, mailing address, etc).	3	3	2	3	2832	0.202	3621	1.000	*3978*	*0.019*	*3694*	*<0.001*
Your account is inaccessible indefinitely.	3	3	3	3	2477	0.599	3327	1.000	3570	0.157	*3440*	*0.002*
Your account can be reactivated.	2	2	2	2	1943	0.202	3593	1.000	2753	0.157	*1986*	*0.008*
Your ancestry/health DNA test results are deleted.	3	4	3	4	*3015*	*0.016*	3780	1.000	*4076*	*0.009*	*3860*	*<0.001*
People can still find your account on the website.	3	2	3	3	2194	0.599	3293	1.000	*2352*	*0.013*	*2052*	*0.011*

Control: How involved would each entity be in deciding how your health related DNA data is managed? That is, who controls your DNA data?	Medians				UA-UH		NUA-NUH		UA-NUA		UH-NUH	
	UA	NUA	UH	NUH	U	p	U	p	U	p	U	p
The DNA company itself	1	1	1	1	2206	1.000	3401	1.000	3155	1.000	2576	1.000
The general public	4	4	4	4	2296	1.000	*3465*	*1.000*	*3913*	*0.037*	3150	0.105
Other users of the at-home DNA testing company	3	4	3	4	2213	1.000	3454	1.000	3728	0.413	3011	0.789
The government	3	2	3	2	2670	1.000	3557	1.000	*2367*	*0.025*	2270	0.678
For-profit partners of the company	3	2	3	2	2088	1.000	3616	1.000	3082	1.000	2265	0.789
Non-profit partners of the company	3	3	3	3	2374	1.000	3486	1.000	2945	1.000	2513	1.000
Investors	3	3	3	3	2299	1.000	3314	1.000	2930	1.000	2585	1.000
Yourself (the person completing the DNA test)	2	2	1.5	2	2713	1.000	3478	1.000	3086	1.000	3034	0.789

Table 7: *Deletion* and *Control* question blocks (4-point Likert scales, 1: most positive, 4: most negative). Medians per condition and the Holm-Bonferroni adjusted statistical results of Mann-Whitney U tests. Green cells with italics indicate statistically significant results.

RQ2 [Management]: Non-users are more skeptical than users of data deletion. Analysis revealed statistically significant differences between users and non-users in all questions of the *Deletion* block, with non-users being more skeptical of deletion, but few statistically significant differences relating to control.

Deletion. We found statistically significant differences for all questions between health users (UH) and non-users (NUH), and for half of questions between ancestry users (UA) and non-users (NUA). Non-users were more likely think that data would be retained in several forms and that information from a deleted account could still be visible to others.

Control. We found statistically significant differences on two questions within the *Control* block. Ancestry users (UA) were more likely than non-users (NUA) to believe that *the general public* had some control over users' DNA data, and were less likely to believe that *the government* had control over DNA data.

5 DISCUSSION

5.1 Research questions

Our research questions focus on the differences between participants' privacy perceptions of ancestry and health DNA data, as well as those between users and non-users of at-home DNA testing.

Addressing RQ1, we find few significant differences between participants in the ancestry and health DNA test conditions. Addressing RQ2, we find more differences between the privacy perceptions of users and non-users.

RQ1: Ancestry versus Health. We found no statistically significant differences between ancestry and health conditions for most survey questions. However, we note four areas where most statistically significant differences occurred.

Participants thought health data should be shared for purposes leading to the greater good, but not necessarily ancestry data; participants were significantly more likely to believe that health data should be contributed by others for research by the government and police, to academic institutions, and for police investigations. Similarly, users were significantly more likely to believe health testing (as opposed to ancestry) would positively affect users' potential government benefits.

Users in the *health* data condition (UH) were also significantly more concerned about discovering unpleasant information than ancestry users (UA). Unpleasant news related to health can be perceived as a risk to one's life, potentially with serious medical implications to the user or their biological family members. In comparison, ancestry tests may not have such severe negative consequences.

Significantly more participants in the health conditions believed some of the privacy-compromising scenarios to be likely within Canada or in general. Such scenarios include being tracked by the government, or receiving an increase in life insurance rates.

Finally, significantly more health users (UH) believed their DNA test results would be deleted upon account deletion. These participants may (incorrectly) believe that their DNA results would be treated with as much sensitivity (and possibly regulation) as medical information [8].

Our findings support the limited related literature in the field. For example, in Baig et al.'s study [7], users perceived health data to be valuable for research and development of potentially life-saving drugs/treatments. Participants have also expressed concerns about increased health and life insurance rates in the past [7, 24], noting that insurance premiums are tied to an individual's health. Many also feel uneasy about negative consequences of health testing, and finding out they may potentially have a life-altering condition [7].

RQ2: Users versus Non-users. We found that users of at-home DNA testing services associated numerous benefits with DNA testing, while non-users considered these tests to have more risks and fewer benefits.

Several users many benefits with DNA testing: lower insurance rates, improvements to government benefits, and better employment prospects. Users were also significantly more likely to believe DNA tests would have positive effects on their family members. They were more comfortable sharing their data with third-parties, and wanted others to share their data as well. Additionally, users generally believed that finding family members was a benefit of at-home DNA testing.

In contrast, non-users were not interested in finding family members, and did not associate other benefits with the tests. In fact, significantly more non-users perceived risks associated with at-home DNA testing, such as an increase in insurance rates or negative consequences from an employer accessing results online. Non-user participants were generally more privacy conscious, more protective of their personal data, and less trusting of third-parties. Significantly more non-users thought genetic data was different from other data.

Our study results aligns with existing literature; several studies have shown that users perceive benefits much more strongly than non-users, and perceived more control over their data [24, 49]. Many users had an incomplete understanding of privacy and risk [24], or had not considered risks in their decision [48]. Non-users, conversely, had concerns about control, and perceived several more privacy risks [24, 27], even though some considered the process to be fun [27]. Additionally, non-users in other work [27] have similarly expressed concerns over protecting family members' privacy in the case of a criminal investigation or previously hidden information.

5.2 Misconceptions

Our survey results suggest that many participants have an incomplete understanding of risks associated with at-home DNA testing. We highlight participant misconceptions, discuss these risks, and explain any resulting implications below.

Re-identification. In our survey, users were significantly less likely than non-users to believe that it was possible to be traced back through their DNA sample. However, not only can DNA be used to trace the individuals themselves [28, 33, 50], but it can also be used to identify their relatives, including those who have not taken a DNA test [17, 50]. This is a serious gap in users' understanding of at-home DNA tests considering that more than half of individuals with Northern European heritage in the United States can currently be identified using user uploaded DNA [50].

Data misuse. As part of our survey, we aimed to understand the degree to which participants perceived the plausibility and likelihood of their data being misused in various circumstances. We investigated these perceptions using scenarios based on either real life events, concerns, or benefits of genetic testing. While most participants considered all scenarios plausible, significant differences were found between users' and non-users' perceptions.

Non-users were significantly more likely to believe a variety of negative scenarios to be plausible, and likely to happen in Canada and in general. Users considered several negative scenarios unlikely but many of our scenarios have in fact already happened [10, 15, 59] or are technically feasible [42]. This suggests another serious gap in users' knowledge. Many instances of misuse have been presented in the media over the years. In one case, discrimination based on health-related data occurred [59] on the basis of markers for a disease that had not manifested in the user. In this incident, the DNA information was also shared with unauthorised individuals by the organisation who had received the information in confidence. This demonstrates that users' health-related DNA data, even with its perceived sensitivity, may not always be treated with the expected confidentiality.

It is also plausible that such information may inadvertently be revealed to unintended parties. Several entities, for example,

use social media to conduct research on individuals, such as web-scraping to investigate suspects in criminal investigations, or Facebook searches by hiring managers to research a potential future employee [51, 55]. Having DNA information online may publicise it beyond the user's intended network, and could present further negative consequences such as negatively affecting their job prospects [3, 53].

Permanence of data. Although participants perceived control of their data, intricate details of the associated privacy policies may indicate otherwise [41]. In our survey, significantly more users than non-users demonstrated the belief that DNA data would be deleted with their account. This is not always the case. An at-home DNA testing company experienced an error while merging information into their main database and deleted information temporarily reappeared in users' accounts [36]. This suggests it had remained stored somewhere on their system. The company claims that the data was permanently deleted after the merge [36] but the incident reveals that data might not be deleted as some believe.

Implications. Based on our data, and existing literature [7], users are more likely than non-users to consider DNA data to be similar to any other form of data, such as web tracking cookies used for advertising or personalization purposes, or financial information such as credit card numbers. While these carry privacy risks, the stakes are much higher for DNA data and users seem indifferent to these risks. DNA is inherently identifiable, and is permanently linked to its owner [41]. If the data is leaked or misused, it cannot be replaced or changed by the user (compared to a password or credit card number). In addition, the consequences associated with misuse are arguably more severe with DNA data, such as when it is used to discriminate against users in various scenarios.

DNA data presents networked, relatively unique privacy risks to which most users pay little attention. A user's DNA data is linked to their immediate and extended biological family members, past, present, and future, who may also suffer negative consequences as a result [14, 21]. Family members have no control or say over what happens to this data (i.e., their own partial DNA)[21].

DNA test results cannot be easily rescinded even by users; it is unclear how long data exists on these platforms or what happens to it beyond the lifespan of the initial testing company (e.g., company mergers or dissolutions) [41]. Furthermore, family members of users have no real mechanism in place to request removal of their partial data and they are unlikely to even be informed in case of data breaches or misuse.

As such, our study results are concerning. Users generally consider the effects of their DNA test to be positive on family members, while non-users perceive there to be no effect. Users also seem to have weaker understanding of the consequences of a data breach: why their DNA is unique and the risks they present. This may be exacerbated by perceived control over their data: that is, being able to delete it permanently gives them a way to "take it back" [7]. They also underestimate the likelihood of negative scenarios in their own country of residence. In comparison, our non-users presented more privacy-conscious attitudes, which may have factored into their decision to not take an at-home DNA test.

5.3 Proposed privacy-enhancing mitigations

Dedicated dashboard. Users currently have no efficient method of viewing with whom their data is shared or who has already accessed it. For example, DNA testing companies may provide a list of academic publications resulting from the use of their consumers' data [2] but may not provide any indication to the user about whether their data was actually shared for any given study. Furthermore, users may wish to be informed if new study results are relevant to them.

Providing a usable dashboard for users to view their data, view (or, ideally, control) who has access to it, and see the outcomes of data sharing would be a step towards informed data use. Such a dashboard could, at the very least, list the entities who have received this data, what the data has been used for, and what the outcomes are (if any) of the related research. In addition, providing opportunities for users to filter based on study results, look for certain keywords, and perhaps introducing tools for organisation would make this information manageable.

We recognise that this may be counter to some business models that rely on data sharing despite privacy consequences to end-users; recent legislation (e.g., GDPR) signals increased interest in user rights. As such, regulations could require such dashboards to be provided by vendors.

Awareness campaign. We support the development of informational materials to clarify common misconceptions surrounding at-home DNA testing, and provide knowledge of the process itself. We suggest including information about accuracy and limitations of at-home DNA testing, discussion of common misconceptions relating to re-identification, misuse, and potential consequences, and highlighting the potential effects on immediate and extended biological family members' privacy. This information could be shared by neutral third parties and/or DNA-testing companies could be required to supply this information to individuals before selling DNA tests.

5.4 Limitations and Future Work

Participants may have self-selection bias given their interest in the topic. Given the length of the survey, questions may have primed participants to respond in a way geared towards privacy. Due to the nature of the topic, the social desirability effect may have caused participants to respond how they think researchers expect them to respond. Finally, we assigned non-users randomly to the health or ancestry condition. An alternative could be explicitly asking about their interest in either and basing assignment on this interest, but it is unclear how this may have affected responses.

Future research on genetic privacy could investigate perceptions relating to biological family members and related privacy considerations when deciding to take a test. The design of an interface that helps users manage their DNA data more efficiently would help add to the literature on privacy. Also, limited research explores how individuals develop their understanding of at-home DNA testing. Conducting a study with individuals who explicitly chose not to take an at-home DNA-test may also reveal insights relating to privacy. Finally, based on our results, a follow-up survey could be designed to confirm the more general concepts identified in this paper.

6 CONCLUSION

Despite privacy concerns by individuals, the monetary value of the at-home DNA testing market is expected to reach a record high within the next decade. While benefits are perceived by many, there continue to be many privacy risks associated with such services. We explored whether users and non-users perceive the privacy of such services differently through a survey with 310 participants. We gathered participants' perceptions of the benefits and risks of genetic testing, their understanding of data access and permanence, and the effects such tests have on their biological family members. We further compare whether these perceptions differ across ancestry and health data. We found users to be significantly less concerned about privacy than non-users, perceiving significantly more positive effects to themselves and others. In contrast, non-users displayed several privacy concerns, and were less comfortable sharing their data. In addition, health data was implied to be more valuable to the greater good, and potentially more sensitive, though there were fewer statistically significant differences between the health and ancestry conditions. We discuss common misconceptions based on our findings, and highlight its relevance to the current literature.

ACKNOWLEDGMENTS

REFERENCES

[1] Inc. 23andMe. 2022. 23andMe DNA Testing Kit for Health + Ancestry. https://www.23andme.com/en-ca/dna-health-ancestry/.

[2] Inc. 23andMe. 2023. Publications. https://research.23andme.com/publications/.

[3] Alessandro Acquisti and Christina Fong. 2020. An experiment in hiring discrimination via online social networks. Management Science 66, 3 (2020), 1005–1024.

[4] Megan A Allyse, David H Robinson, Matthew J Ferber, and Richard R Sharp. 2018. Direct-to-consumer testing 2.0: emerging models of direct-to-consumer genetic testing. In Mayo clinic proceedings, Vol. 93. Elsevier, Elsevier, Amsterdam, Netherlands, 113–120. Issue 1.

[5] Kate Amos. 2022. Gifted Genes: An Investigation of College Student Perceptions of Direct-To-Consumer Genetic Kits. WWU Honors College Senior Projects. WWU Honors College.

[6] AncestryDNA. 2022. Uncover your past with AncestryDNA. https://www.ancestry.ca/dna/.

[7] Khadija Baig, Reham Mohamed, Anna-Lena Theus, and Sonia Chiasson. 2020. " I'm hoping they're an ethical company that won't do anything that I'll regret" Users Perceptions of At-home DNA Testing Companies. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, United States, 1–13.

[8] Khadija Nawal Baig. 2023. Privacy, Biological Relatives, and at-Home DNA Testing. Ph. D. Dissertation. Carleton University.

[9] Natalie M Baptista, Kurt D Christensen, Deanna Alexis Carere, Simon A Broadley, J Scott Roberts, and Robert C Green. 2016. Adopting genetics: motivations and outcomes of personal genomic testing in adult adoptees. Genetics in Medicine 18, 9 (2016), 924–932.

[10] Jorge Barrera and Tiffany Foxcroft. 2018. Heredity or hoax? How dog DNA helped uncover a suspected Indian status scam. https://newsinteractives.cbc.ca/longform/dna-ancestry-test.

[11] Angela Bearth and Michael Siegrist. 2020. Psychological factors that determine people's willingness-to-share genetic data for research. Clinical Genetics 97, 3 (2020), 483–491.

[12] Mwenza Blell and MA Hunter. 2019. Direct-to-Consumer Genetic Testing's Red Herring: "Genetic Ancestry" and Personalized Medicine. Frontiers in medicine 6 (2019), 48.

[13] Juli Murphy Bollinger, Robert C Green, and David Kaufman. 2013. Attitudes about regulation among direct-to-consumer genetic testing customers. Genetic testing and molecular biomarkers 17, 5 (2013), 424–428.

[14] Danah Boyd. 2012. Networked privacy. Surveillance & Society 10, 3/4 (2012), 348.

[15] Emanuela Campanella. 2018. Government using ancestry websites to deport immigrants: lawyer, court documents. https://globalnews.ca/news/4616715/ancestry-dna-test-deportation-cbsa/.

[16] LaKisha Tawanda David and Leia Jones. 2019. Genetic Genealogical Methods Used to Identify African American Diaspora Relatives in the Study of Family Identity among Ghanaian Members of the Kassena Ethnic Group. https://doi.org/10.1101/833996 arXiv:https://www.biorxiv.org/content/early/2019/11/07/833996.full.pdf

[17] Yaniv Erlich, Tal Shor, Itsik Pe'er, and Shai Carmi. 2018. Identity inference of genomic data using long-range familial searches. Science 362, 6415 (2018), 690–694.

[18] Christina Farr. 2019. Consumer DNA testing has hit a lull — here's how it could capture the next wave of users. https://www.cnbc.com/2019/08/25/dna-tests-from-companies-like-23andme-ancestry-see-sales-slowdown.html.

[19] Christina Farr. 2020. 23andMe lays off 100 people as DNA test sales decline, CEO says she was 'surprised' to see market turn. https://www.cnbc.com/2020/01/23/23andme-lays-off-100-people-ceo-anne-wojcicki-explains-why.html.

[20] Christina Farr. 2020. Ancestry to lay off 6% of workforce because of a slowdown in the consumer DNA-testing market. https://www.cnbc.com/2020/02/05/ancestry-layoffs-of-6percent-100-people-amid-dna-test-slowdown.html.

[21] Julie Frizzo-Barker, Peter A Chow-White, Anita Charters, and Dung Ha. 2016. Genomic big data and privacy: challenges and opportunities for precision medicine. Computer Supported Cooperative Work (CSCW) 25, 2-3 (2016), 115–136.

[22] Sarah E Gollust, Erynn S Gordon, C Zayac, Georgia Griffin, MF Christman, RE Pyeritz, Lisa Wawak, and Barbara A Bernhardt. 2012. Motivations and perceptions of early adopters of personalized genomics: perspectives from research participants. Public health genomics 15, 1 (2012), 22–30.

[23] Sarah E Gollust, Stacy W Gray, Deanna Alexis Carere, Barbara A Koenig, Lisa Soleymani Lehmann, Amy L McGuire, Richard R Sharp, Kayte Spector-bagdady, NA Wang, Robert C Green, et al. 2017. Consumer perspectives on access to direct-to-consumer genetic testing: role of demographic factors and the testing experience. The Milbank Quarterly 95, 2 (2017), 291–318.

[24] Sukeshini A Grandhi and Linda Plotnick. 2022. Do I Spit or Do I Pass? Perceived Privacy and Security Concerns of Direct-to-Consumer Genetic Testing. Proceedings of the ACM on Human-Computer Interaction 6, GROUP (2022), 1–26.

[25] Stacy W Gray, Sarah E Gollust, Deanna Alexis Carere, Clara A Chen, Angel Cronin, Sarah S Kalia, Huma Q Rana, Mack T Ruffin IV, Catharine Wang, J Scott Roberts, et al. 2017. Personal genomic testing for cancer risk: results from the impact of personal genomics study. Journal of Clinical Oncology 35, 6 (2017), 636.

[26] Tobias Haeusermann, Bastian Greshake, Alessandro Blasimme, Darja Irdam, Martin Richards, and Effy Vayena. 2017. Open sharing of genomic data: Who does it and why? PLoS One 12, 5 (2017), e0177158.

[27] James W Hazel, Catherine Hammack-Aviran, Kathleen M Brelsford, Bradley A Malin, Laura M Beskow, and Ellen Wright Clayton. 2021. Direct-to-consumer genetic testing: Prospective users' attitudes toward information about ancestry and biological relationships. PloS one 16, 11 (2021), e0260340.

[28] Nils Homer, Szabolcs Szelinger, Margot Redman, David Duggan, Waibhav Tembe, Jill Muehling, John V Pearson, Dietrich A Stephan, Stanley F Nelson, and David W Craig. 2008. Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. PLoS genetics 4, 8 (2008), e1000167.

[29] Mathias Humbert, Kévin Huguenin, Joachim Hugonot, Erman Ayday, and J-P Hubaux. 2015. De-anonymizing genomic databases using phenotypic traits. Proceedings on Privacy Enhancing Technologies 2015, 2 (2015), 99–114.

[30] Robert Karl. 2020. 23andDiverseMe: Using Genetic Ancestry Tests to Establish Minority Status. Health Matrix 30 (2020), 475.

[31] Debbie Kennett. 2019. Using genetic genealogy databases in missing persons cases and to develop suspect leads in violent crimes. Forensic science international 301 (2019), 107–117.

[32] Jennifer King. 2019. " Becoming Part of Something Bigger" Direct to Consumer Genetic Testing, Privacy, and Personal Disclosure. Proceedings of the ACM on Human-Computer Interaction 3, CSCW (2019), 1–33.

[33] Bartha Maria Knoppers. 2010. Consent to 'personal' genomics and privacy: Direct-to-consumer genetic tests and population genome research challenge traditional notions of privacy and consent. EMBO reports 11, 6 (2010), 416–419. https://www.embopress.org/doi/abs/10.1038/embor.2010.69

[34] Marie Løtveit and Anniken Sofie Brauti. 2019. The commercialization of our DNA: a qualitative study of attitudes regarding ancestry DNA tests. Master's thesis. Norwegian School of Economics.

[35] Amy L McGuire, Christina M Diaz, Tao Wang, and Susan G Hilsenbeck. 2009. Social networkers' attitudes toward direct-to-consumer personal genome testing. The American journal of bioethics 9, 6-7 (2009), 3–10.

[36] Emily Mullin. 2021. 'Deleted' DNA Data Just Reappeared on a Popular Database. https://futurehuman.medium.com/deleted-dna-data-just-reappeared-on-a-popular-database-e1f43587f7ec.

[37] Jim Mustian. 2015. New Orleans filmmaker cleared in cold-case murder; false positive highlights limitations of familial DNA searching. https://www.nola.com/article_d58a3d17-c89b-543f-8365-a2619719f6f0.html.

[38] L Nikolova. 2018. Scholarly branches in 21st-century genealogy (Earliest 19th-century Bulgarian immigrants in the US at ancestry. com). *Journal of Historical Archaeology & Anthropological Sciences* 1, 6 (2018), 1–10.

[39] Future of Privacy Forum. 2018. Future of Privacy Forum and Leading Genetic Testing Companies Announce Best Practices to Protect Privacy of Consumer Genetic Data. https://fpf.org/blog/future-of-privacy-forum-and-leading-genetic-testing-companies-announce-best-practices-to-protect-privacy-of-consumer-genetic-data.

[40] Lukasz Olejnik, Agnieszka Kutrowska, and Claude Castelluccia. 2014. I'm 2.8% Neanderthal: The beginning of genetic exhibitionism?

[41] Andelka M Phillips. 2015. Genomic Privacy and Direct-to-Consumer Genetics: Big Consumer Genetic Data–What's in that Contract? , 60–64 pages.

[42] The Associated Press. 2020. U.S. to start collecting DNA from people detained at border. https://www.cbc.ca/news/canada/windsor/us-collect-dna-border-detained-1.5417120.

[43] Prolific. 2022. Researchers. https://www.prolific.co/researchers.

[44] Lu Qiao, Yajun Yang, Pengcheng Fu, Sile Hu, Hang Zhou, Shouneng Peng, Jingze Tan, Yan Lu, Haiyi Lou, Dongsheng Lu, et al. 2018. Genome-wide variants of Eurasian facial shape differentiation and a prospective model of DNA based face prediction. *Journal of Genetics and Genomics* 45, 8 (2018), 419–432.

[45] Qualtrics. 2022. Online Survey Software - Powering +1B Surveys Annually. https://www.qualtrics.com/core-xm/survey-software/.

[46] Antonio Regalado. 2019. More than 26 million people have taken an at-home ancestry test. https://www.technologyreview.com/2019/02/11/103446/more-than-26-million-people-have-taken-an-at-home-ancestry-test/.

[47] Kuick Research. 2022. *Home DNA Test Kit Market Opportunity Forecast 2028*. Kuick Research, Delhi, India.

[48] J Scott Roberts, Michele C Gornick, Deanna Alexis Carere, Wendy R Uhlmann, Mack T Ruffin, and Robert C Green. 2017. Direct-to-consumer genetic testing: user motivations, decision making, and perceived utility of results. *Public Health Genomics* 20, 1 (2017), 36–45.

[49] Grayson L Ruhl, James W Hazel, Ellen Wright Clayton, and Bradley A Malin. 2019. Public attitudes toward direct to consumer genetic testing. In *AMIA Annual Symposium Proceedings*, Vol. 2019. American Medical Informatics Association, Oxford University Press, Elsevier, BMJ, United Kindgom, Netherlands,, 774.

[50] Julian Segert. 2018. Understanding Ownership and Privacy of Genetic Data. http://sitn.hms.harvard.edu/flash/2018/understanding-ownership-privacy-genetic-data/.

[51] William P. Smith and Deborah L. Kidder. 2010. You've been tagged! (Then again, maybe not): Employers and Facebook. *Business Horizons* 53, 5 (2010), 491–499.

[52] Emily G Spencer and Eric J Topol. 2019. Direct to consumer fitness DNA testing. *Clinical chemistry* 65, 1 (2019), 45–47.

[53] Robert Sprague. 2011. Invasion of the Social Networks: Blurring the Line between Personal Life and the Employment Relationship. *University of Louisville Law Review* 50 (02 2011).

[54] Stephany Tandy-Connor, Jenna Guiltinan, Kate Krempely, Holly LaDuca, Patrick Reineke, Stephanie Gutierrez, Phillip Gray, and Brigette Tippin Davis. 2018. False-positive results released by direct-to-consumer genetic tests highlight the importance of clinical confirmation testing for appropriate patient care. *Genetics in Medicine* 20, 12 (2018), 1515.

[55] Steven L. Thomas, Philip C. Rothschild, and Caroline Donegan. 2015. Social Networking, Management Responsibilities, and Employee Rights: The Evolving Role of Social Networking in Employment Decisions. *Employee Responsibilities and Rights Journal* 27, 4 (2015), 307–323.

[56] Effy Vayena, Elli Gourna, Jürg Streuli, Ernst Hafen, and Barbara Prainsack. 2012. Experiences of early users of direct-to-consumer genomics in Switzerland: an exploratory study. *Public Health Genomics* 15, 6 (2012), 352–362.

[57] Effy Vayena, Christian Ineichen, Elia Stoupka, and Ernst Hafen. 2014. Playing a part in research? University students' attitudes to direct-to-consumer genomics. *Public health genomics* 17, 3 (2014), 158–168.

[58] Ritchie Witzig. 1996. The medicalization of race: scientific legitimization of a flawed social construct.

[59] Sarah Zhang. 2016. DNA Got a Kid Kicked Out of School—And It'll Happen Again. https://www.wired.com/2016/02/schools-kicked-boy-based-dna/.

[60] Sarah Zhang. 2018. A man says his DNA test proves he's Black and he's suing. https://www.theatlantic.com/science/archive/2018/09/dna-test-race-lawsuit/570250/#:~:text=So%20as%20proof%20of%20his,4%20percent%20sub%2DSaharan%20African..

A PARTICIPANT DEMOGRAPHICS

Demographic	Number of participants	Percentage of participants
Gender		
Male	173	55.8%
Female	131	42.2%
Non-binary	3	1.0%
Genderfluid	1	0.3%
Prefer not to answer	2	0.6%
Age		
18 – 24	78	25.2%
25 – 34	152	49.0%
35 - 44	55	17.7%
45 - 54	14	4.5%
55 - 64	8	2.6%
65 - 74	1	0.3%
Prefer not to answer	2	0.6%
Highest/current level of education		
Elementary school		0.3%
High school	34	11.0%
College	27	8.7%
Technical, trade school, or apprenticeship	17	5.5%
Undergraduate degree (Bachelor's)	154	49.7%
Graduate degree (Master's, PhD) or professional degree	68	21.9%
Post-graduate certificate or diploma	7	2.3%
Prefer not to answer	2	0.6%
Occupation		
Administrative Support	19	6.1%
Art, Writing, Journalism	5	1.6%
Business, Management, and Financial	39	12.6%
Education	16	5.2%
Legal	8	2.6%
Medical	6	1.9%
Science, Engineering, and IT Professional	61	19.7%
Service	20	6.5%
Skilled Labour	7	2.3%
Student	58	18.7%
Unemployed	21	6.8%
Retired	5	1.6%
Other	38	11.3%
Prefer not to answer	10	3.2%

Table 8: Detailed breakdown of participant demographics: age, gender, province of residence, level of education, and occupation.

B   DESCRIPTIVES (MEAN, STANDARD DEVIATION, MEDIAN)

We provide additional descriptive statistics relating to each question in our survey, organised by block. We present them in the following format: **Mean (Standard Deviation) (Median)**.

All blocks, except the Effects block, follow a 4-point Likert-scale (1: most positive option, 4: most negative option). The Effects block follows a 3-point scale (1: Positive, 2: No effect, 3: Negative).

Deletion: When you delete your account, how likely is it that:

	Users Ancestry (UA)	Non-Users Ancestry (NUA)	Users Health (UH)	Non-Users Health (NUH)
Your raw DNA is deleted.	3.26 (0.79) (3)	3.50 (0.74) (4)	2.95 (0.85) (3)	3.52 (0.83) (4)
Your personally identifiable information is deleted (name, email address, mailing address, etc).	2.68 (0.90) (3)	3.07 (0.94) (3)	2.40 (0.85) (2)	3.04 (0.98) (3)
Your account is inaccessible indefinitely.	2.81 (0.74) (3)	3.01 (0.84) (3)	2.66 (0.79) (3)	3.11 (0.90) (3)
Your account can be reactivated.	2.08 (0.79) (2)	1.89 (0.74) (2)	2.25 (0.62) (2)	1.92 (0.83) (2)
Your ancestry/health DNA test results are deleted.	3.20 (0.78) (3)	3.53 (0.77) (4)	2.77 (0.84) (3)	3.47 (0.78) (4)
People can still find your account on the website.	2.76 (0.83) (3)	2.35 (0.94) (2)	2.89 (0.83) (3)	2.47 (1.01) (3)

Concerns: Please rate your level of concern with each of the following:

	Users Ancestry (UA)	Non-Users Ancestry (NUA)	Users Health (UH)	Non-Users Health (NUH)
Your ancestry/health results influencing your job prospects.	2.76 (1.06) (3)	2.47 (1.04) (3)	2.50 (0.95) (3)	2.14 (1.01) (2)
Finding out unwanted ancestry/health information	2.80 (0.88) (3)	2.62 (1.03) (3)	2.25 (0.88) (2)	2.22 (0.92) (2)
Finding unwanted family secrets as a result of an ancestry DNA test.	2.87 (0.93) (3)	2.85 (1.01) (3)	2.42 (0.98) (2)	2.72 (0.97) (3)
Surveillance due to an ancestry/health DNA test.	2.28 (0.99) (2)	2.01 (1.05) (2)	2.24 (0.95) (2)	1.79 (0.91) (2)
Accuracy of ancestry/health at-home DNA tests.	2.28 (0.84) (2)	1.99 (0.99) (2)	1.90 (0.67) (2)	1.96 (0.92) (2)
Genetic discrimination as a result of an ancestry/health DNA test.	2.78 (1.01) (3)	2.10 (1.03) (2)	2.31 (1.02) (2)	2.01 (0.96) (2)
Ancestry/health information from a DNA test relevant to your immediate or extended family being public	2.41 (0.90) (2)	2.04 (0.92) (2)	2.25 (0.92) (2)	2.05 (0.90) (2)
Adjusted insurance rates due to an ancestry/health DNA test.	2.39 (1.05) (2)	1.90 (1.00) (2)	2.21 (0.83) (2)	1.78 (0.85) (2)
Your ancestry/health data being used for profit.	2.03 (0.86) (2)	1.38 (0.71) (1)	1.78 (0.89) (2)	1.55 (0.84) (1)

Control: Who controls your ancestry/health data?

	Users Ancestry (UA)	Non-Users Ancestry (NUA)	Users Health (UH)	Non-Users Health (NUH)
The DNA company itself	1.37 (0.67) (1)	1.38 (0.77) (1)	1.44 (0.61) (1)	1.46 (0.81) (1)
The general public	3.47 (0.66) (4)	3.74 (0.47) (4)	3.51 (0.69) (4)	3.76 (0.48) (4)
Other users of the at-home DNA testing company	3.09 (0.91) (3)	3.36 (0.74) (4)	3.24 (0.80) (3)	3.39 (0.83) (4)
The government	3.01 (0.87) (3)	2.58 (0.96) (2)	2.84 (0.88) (3)	2.59 (0.95) (2)
For-profit partners of the company	2.67 (0.99) (3)	2.59 (1.07) (2)	2.87 (0.82) (3)	2.59 (1.14) (2)
Non-profit partners of the company	2.87 (0.85) (3)	2.73 (0.88) (3)	2.89 (0.76) (3)	2.79 (0.91) (3)
Investors	2.80 (0.92) (3)	2.64 (1.01) (3)	2.85 (0.81) (3)	2.78 (1.04) (3)
Yourself (the person completing the DNA test)	1.96 (1.03) (2)	1.94 (1.06) (2)	1.69 (0.83) (1.5)	1.94 (1.00) (2)

Data access: Who has access to your most recent ancestry/health DNA data?

	Users Ancestry (UA)	Non-Users Ancestry (NUA)	Users Health (UH)	Non-Users Health (NUH)
The DNA company itself	1.28 (0.64) (1)	1.29 (0.72) (1)	1.44 (0.76) (1)	1.38 (0.85) (1)
The general public	3.61 (0.61) (4)	3.54 (0.57) (4)	3.64 (0.52) (4)	3.73 (0.47) (4)
Other users of the at-home DNA testing company	3.01 (0.94) (3)	2.96 (1.01) (3)	3.33 (0.84) (4)	3.27 (0.96) (4)
Your employer	3.67 (0.64) (4)	3.55 (0.66) (4)	3.42 (0.77) (4)	3.51 (0.61) (4)
The government	2.80 (1.03) (3)	2.39 (1.00) (2)	2.47 (0.99) (3)	2.52 (1.09) (2)
Law enforcement	3.05 (0.92) (3)	2.46 (1.01) (2)	2.84 (0.90) (3)	2.59 (1.11) (2)
Pharmaceutical companies	3.04 (0.86) (3)	2.58 (0.97) (3)	2.73 (0.88) (3)	2.46 (1.06) (2)
Insurance companies	3.22 (0.79) (3)	2.74 (1.00) (3)	2.78 (1.05) (3)	2.72 (0.96) (3)
Advertisers	3.21 (0.88) (3)	2.92 (1.00) (3)	3.11 (0.84) (3)	2.92 (1.03) (3)
For-profit companies	2.93 (0.98) (3)	2.59 (1.07) (3)	2.92 (0.90) (3)	2.86 (1.04) (3)
University researchers	2.95 (0.88) (3)	2.38 (0.84) (2)	2.59 (0.87) (2.5)	2.41 (0.98) (2)
Non-profit researchers	2.88 (0.91) (3)	2.54 (0.93) (3)	2.59 (0.92) (2)	2.59 (0.97) (3)
Investors in the DNA testing company	2.79 (0.90) (3)	2.59 (1.00) (2)	2.75 (0.89) (3)	2.76 (1.05) (3)

Data usage: How comfortable are you with your data being <u>used</u>:

	Users Ancestry (UA)	Non-Users Ancestry (NUA)	Users Health (UH)	Non-Users Health (NUH)
By the DNA-testing company to improve their services.	2.11 (0.87) (2)	2.70 (1.02) (3)	1.98 (0.81) (2)	2.84 (0.99) (3)
By the government for research.	2.68 (1.04) (3)	3.08 (0.90) (3)	2.41 (0.99) (2)	3.00 (0.90) (3)
By law-enforcement for research.	2.91 (1.06) (3)	3.25 (0.92) (4)	2.56 (1.02) (2)	3.34 (0.91) (4)
By law-enforcement for police investigations.	2.97 (1.01) (3)	3.44 (0.85) (4)	2.80 (0.98) (3)	3.38 (0.87) (4)
By pharmaceutical companies for research.	2.43 (1.01) (2)	2.85 (0.92) (3)	2.27 (0.94) (2)	2.92 (1.06) (3)
By academic institutions for research.	2.01 (0.87) (2)	2.32 (1.04) (2)	1.81 (0.77) (2)	2.40 (1.07) (2)
By non-profit organisations for research.	2.17 (0.93) (2)	2.52 (1.02) (2)	2.11 (0.86) (2)	2.54 (1.08) (2)

Data sharing: How comfortable are you <u>sharing</u> your ancestry/health data with:

	Users Ancestry (UA)	Non-Users Ancestry (NUA)	Users Health (UH)	Non-Users Health (NUH)
The DNA company itself	2.04 (0.93) (2)	2.89 (0.9) (3)	2.11 (0.84) (2)	3.06 (0.92) (3)
The general public	3.34 (0.79) (4)	3.78 (0.52) (4)	3.53 (0.71) (4)	3.86 (0.38) (4)

Other users of the at-home DNA testing company	2.96 (0.94) (3)	3.56 (0.73) (4)	3.22 (0.71) (3)	3.69 (0.58) (4)
Your employer	3.25 (0.94) (4)	3.81 (0.48) (4)	3.23 (0.89) (3)	3.76 (0.48) (4)
The government	3.29 (0.88) (4)	3.57 (0.76) (4)	3.08 (0.93) (3)	3.59 (0.70) (4)
Law enforcement	3.33 (0.84) (4)	3.60 (0.79) (4)	3.23 (0.85) (3)	3.59 (0.74) (4)
Pharmaceutical companies	3.09 (0.87) (3)	3.44 (0.85) (4)	2.77 (0.94) (3)	3.31 (0.93) (4)
Insurance companies	3.42 (0.84) (4)	3.74 (0.69) (4)	3.00 (0.99) (3)	3.71 (0.59) (4)
Advertisers	3.50 (0.72) (4)	3.87 (0.46) (4)	3.58 (0.64) (4)	3.89 (0.31) (4)
For-profit companies	3.32 (0.80) (4)	3.75 (0.58) (4)	3.36 (0.63) (3)	3.85 (0.36) (4)
University researchers	2.38 (0.94) (2)	2.75 (1.03) (3)	2.23 (0.79) (2)	2.69 (0.96) (3)
Non-profit researchers	2.57 (0.98) (2)	3.02 (1.09) (3)	2.38 (0.88) (2)	2.88 (1.00) (3)
Investors in the DNA testing company	3.29 (0.88) (4)	3.71 (0.59) (4)	3.08 (0.90) (3)	3.78 (0.52) (4)

Contributions: Rate how much you would like others who have also completed the test to <u>contribute their ancestry/health-related DNA data</u> for the following purposes:

	Users Ancestry (UA)	Non-Users Ancestry (NUA)	Users Health (UH)	Non-Users Health (NUH)
For research by the DNA-testing company to improve their services.	2.01 (0.97) (2)	2.96 (0.99) (3)	2.03 (0.98) (2)	2.89 (1.06) (3)
For research by the government.	3.11 (1.01) (3)	3.28 (0.83) (3)	2.61 (0.99) (2)	3.27 (0.92) (4)
For research by the police.	3.29 (0.91) (4)	3.55 (0.78) (4)	2.89 (0.98) (3)	3.47 (0.92) (4)
For police investigations.	3.28 (0.93) (4)	3.45 (0.89) (4)	2.81 (1.02) (3)	3.47 (0.95) (4)
For research by pharmaceutical companies.	2.63 (1.06) (3)	3.14 (0.87) (3)	2.28 (1.09) (2)	3.00 (1.10) (3)
For research by academic institutions.	2.16 (0.95) (2)	2.56 (1.07) (2)	1.66 (0.78) (1)	2.39 (1.08) (2)
For research by non-profit organisations.	2.49 (1.03) (2)	2.87 (1.02) (3)	2.11 (0.94) (2)	2.60 (1.09) (3)

General attitudes: Rate how you strongly you agree or disagree with each of the following statements:

	Users Ancestry (UA)	Non-Users Ancestry (NUA)	Users Health (UH)	Non-Users Health (NUH)
I can be identified by the DNA sample I provided for ancestry testing.	2.00 (0.75) (2)	1.86 (0.89) (2)	2.06 (0.88) (2)	1.62 (0.80) (1)
My family might find out things they didn't want to know.	2.46 (0.82) (2)	2.28 (0.95) (2)	2.56 (0.74) (3)	2.31 (0.88) (2)
My DNA is my information and my information alone.	1.80 (0.79) (2)	1.46 (0.77) (1)	1.90 (0.72) (2)	1.44 (0.61) (1)
What I do with my DNA is my business.	1.49 (0.60) (1)	1.31 (0.62) (1)	1.58 (0.59) (2)	1.42 (0.70) (1)
My family have a right to be concerned about my ancestry DNA test.	2.56 (0.96) (2)	2.52 (1.08) (2)	2.17 (0.94) (2)	2.40 (0.92) (2)
I am interested in finding biological family with a DNA test.	2.32 (0.97) (2)	2.99 (0.92) (3)	2.34 (0.88) (2)	3.06 (0.88) (3)
I do not need anyone's consent to take an ancestry DNA test.	1.57 (0.76) (1)	1.80 (0.97) (2)	1.75 (0.93) (1.5)	1.79 (0.86) (2)
Making my ancestry DNA test results public adds to the community.	2.73 (0.89) (3)	3.24 (0.83) (3)	2.80 (0.89) (3)	3.29 (0.78) (3)

People who hide their ancestry DNA results annoy me.	3.32 (0.79) (3)	3.81 (0.42) (4)	3.30 (0.69) (3)	3.69 (0.54) (4)
Ancestry DNA data is just like any other data.	2.64 (0.99) (3)	3.29 (0.83) (3)	2.76 (0.89) (3)	3.20 (0.88) (3)

Effects: Rate the overall effect your ancestry/health genetic test results would have on each of the following entities:

	Users Ancestry (UA)	Non-Users Ancestry (NUA)	Users Health (UH)	Non-Users Health (NUH)
Yourself	1.26 (0.53) (1)	1.60 (0.64) (2)	1.17 (0.46) (1)	1.70 (0.80) (1)
Your employment prospects	2.00 (0.41) (2)	2.16 (0.40) (2)	1.92 (0.45) (2)	2.34 (0.59) (2)
Your siblings	1.55 (0.53) (2)	1.69 (0.58) (2)	1.50 (0.56) (1)	1.66 (0.67) (2)
Your friends	1.80 (0.43) (2)	2.00 (0.31) (2)	1.84 (0.41) (2)	1.95 (0.37) (2)
Your colleagues	1.92 (0.36) (2)	2.06 (0.24) (2)	1.95 (0.33) (2)	2.05 (0.41) (2)
Your insurance rates	2.08 (0.46) (2)	2.34 (0.53) (2)	1.87 (0.61) (2)	2.58 (0.59) (3)
Your parents	1.54 (0.55) (2)	1.71 (0.62) (2)	1.56 (0.53) (2)	1.86 (0.65) (2)
Your potential government benefits	2.03 (0.40) (2)	2.16 (0.53) (2)	1.82 (0.43) (2)	2.13 (0.65) (2)
Current, existing children	1.59 (0.55) (2)	1.74 (0.60) (2)	1.53 (0.57) (1.5)	1.61 (0.61) (2)
Future children	1.45 (0.53) (1)	1.75 (0.67) (2)	1.26 (0.48) (1)	1.58 (0.69) (1)

C   SURVEY MATERIALS

There are 4 versions of this survey with slightly modified questions. Each question contains a *Prefer not to answer* option. We specify the user version of each question with **U:** and the non-user version with **NU:** .

■■■

(ancestry version intro) All questions in this survey apply only to ancestry related genetic tests. That is, at-home DNA tests done to understand genetic ethnicity (where your 'ancestors' were from), or see potential DNA matches (others who may be related to you).
(health version intro) All questions in this survey apply only to health related genetic tests. That is, at-home DNA tests done to understand how your DNA might influence your overall health (for example, how likely you are to develop a disease, how your DNA influences overall wellness - like whether you are more/less likely to be affected by caffeine - and whether you are a carrier for certain genes).

Q1. U: Have you given consent to the testing company to have your ancestry/health-related DNA data used for research?
    NU: Would you give consent to the testing company to have your ancestry/health-related DNA data used for research?

Yes / No / Unsure / Prefer not to answer

Q2. U: Think of your most recent DNA test completed for ancestry/health purposes. Who currently has access to your most recent ancestry/health data?
NU: Assume you have completed a DNA test for ancestry/health purposes. Who would have access to your ancestry/health data?
Question block with 4 options (No access, Probably doesn't have access, Probably has access, Definitely has access)
- The DNA company itself   - The general public.   - Other users of the at-home DNA testing company   - Your employer
- The government.   - Law enforcement   - Pharmaceutical companies   - Insurance companies   - Advertisers
- For-profit companies   - University researchers   - Non-profit researchers   - Investors in the DNA testing company

Q3. U- How comfortable are you sharing your most recent ancestry/health data with:
    NU- How comfortable would you be sharing your ancestry/health data with:
Question block with 4 options (Very uncomfortable sharing, Uncomfortable sharing, Comfortable sharing, Very comfortable sharing)
- The DNA company itself   - The general public.   - Other users of the at-home DNA testing company   - Your employer
- The government.   - Law enforcement   - Pharmaceutical companies   - Insurance companies   - Advertisers
- For-profit companies   - University researchers   - Non-profit researchers   - Investors in the DNA testing company

Q4. U: Think back to your most recent ancestry/health at-home DNA test. How involved is each entity in deciding how your ancestry/health-related DNA data is managed? That is, who controls your DNA data?
    NU: Assume you have completed an ancestry/health at-home DNA test. How involved would each entity be in deciding how your ancestry/health-related DNA data is managed? That is, who controls your DNA data?
Question block with 4 options (Not at all involved, Uninvolved, Involved, Very involved)

- The general public   - Yourself (you, the owner of your DNA)   - The DNA testing company itself   - The government
- Other users of the at-home DNA testing company   - For-profit partners of the company   - Investors
- Non-profit partners of the company

Q5. U: Assuming that your ancestry/health-related data has been stripped of identifiers (for example, your name), please indicate how comfortable <u>you</u> are with your data being used:
NU: Assuming you completed a DNA test, and your ancestry/health-related DNA data has been stripped of identifiers (for example, your name), please indicate how comfortable <u>you</u> would be with your data being used:
Question block with 4 options (Very uncomfortable, Uncomfortable, Comfortable, Very uncomfortable)
- By pharmaceutical companies for research   - By the government for research
- By law-enforcement for police investigations - By the DNA testing company to improve their services
- By academic institutions for research   - By non-profit organisations for research
- By law-enforcement for research

Q6. U: Think of your most recent at-home DNA test completed for ancestry/health purposes. Please rate the degree to

- By law-enforcement for police investigations   - By the DNA testing company to improve their services
- By academic institutions for research   - By non-profit organisations for research   - By law-enforcement for research

Q7. U, NU: Please rate how strongly you agree or disagree with each of the following statements:
Question block with 4 options (Strongly disagree, Disagree, Agree, Strongly agree)
- I can be identified by the DNA sample I provided for **ancestry/health** testing
- My family might find out things they didn't want to know   - My DNA is my information and my information alone
- What I do with my **ancestry/health** DNA is my business   - I am interested in finding biological family with a DNA test
- People who hide their **ancestry/health** DNA test results annoy me   - **Ancestry/health** DNA data is just like any other data
- I do not need anyone else's consent to take an **ancestry/health** DNA test
- Making my **ancestry/health** DNA test results public adds to the community
- My family have a right to be concerned about my **ancestry/health** DNA test,

Q8. U, NU: For the following question, if an option doesn't apply to you, please answer hypothetically.
Please rate the *overall* effect your ancestry/health-related genetic test results would have on each of the following entities:
Question block with 3 options (Negatively, No effect, Positively)
- Yourself   - Your siblings   - Your friends   - Your colleagues
- Your parents   - Your potential government benefits   - Your employment prospects   - Future children
- Current, existing children   - Your insurance rates   - Your parents

Q9. U: Think of your most recent ancestry/health-related DNA test. When you delete your account, how likely is it that:
NU: Assume you have completed a DNA test for ancestry/health purposes. If you were to delete your account, how likely is it that:
Question block with 4 options. (Very unlikely, Unlikely, Very likely, Likely)
- Your DNA is deleted.
- Your personally identifiable information is deleted (name, email address, mailing address, etc)
- Your account is inaccessible indefinitely
- Your account can be reactivated
- Your **ancestry/health** DNA test results are deleted
 - People can still find your account on the website.

Q10. U, NU: Below, you are presented with a series of potential scenarios.
Please rate: (1) whether each scenario is possible (2) how likely each scenario is to occur in [country of residence]
(3) how likely each scenario is to occur in general (worldwide).
3 question blocks (Is this possible?, Is this likely in [country of residence]?, Is this likely in general?), with 2 options each (Yes, No)
- Individuals being detained at a border having their DNA collected for **ancestry/health** tests
- **Ancestry/health** DNA tests being used to assist in deportation cases
- A potential employer coming across your **ancestry/health** DNA test results online, and using them in their decision to hire you
- Your life insurance rates increasing based on your **ancestry/health** DNA test
- Your child being prevented from attending school due to the result of an **ancestry/health** genetic test they completed
- **Ancestry/health** data that you agreed to be used for research being used by for-profit companies,
- **Ancestry/health** related at-home DNA tests being used to provide proof of Indigenous status,
- An at-home **ancestry/health** DNA test helping you detect a medical condition,
- Databases of users' **ancestry/health** DNA results helping law enforcement successfully solve cold cases,
- Submitting an animal's DNA to an at-home DNA testing company for humans producing error-free results,
- Individuals with certain **ancestry/health** traits having their data stored by the government indefinitely,
- Being falsely suspected in a criminal investigation based on **ancestry/health** DNA testing results.

Q11. U, NU: In your own life, please rate your level of concern with each of the following:
Question block with 4 options (Very unconcerned, Unconcerned, Concerned, Very concerned)
- Your **ancestry/health** DNA results influencing your job prospects   - Finding out unwanted **ancestry/health** information
- Finding unwanted family secrets due to an **ancestry/health** DNA test   - Surveillance due to an **ancestry/health** DNA test
- Accuracy of **ancestry/health** at-home DNA tests   - Your **ancestry/health** data being used for profit
- Adjusted insurance rates due to an **ancestry/health** DNA test
- Genetic discrimination as a result of an **ancestry/health** DNA test
- **Ancestry/health** information from a DNA test relevant to immediate or extended family being public

Q12. You have reached the end of the questionnaire. If you select "submit" and advance to the next screen, you will no longer be able to withdraw your data.

Submit my responses / I wish to withdraw from the study and have my data removed. I understand that I will not be paid.