



Published in final edited form as:

AJOB Empir Bioeth. 2018 ; 9(4): 222–234. doi:10.1080/23294515.2018.1544177.

A content analysis of the views of genetics professionals on race, ancestry, and genetics

Sarah Catherine Nelson^a, Joon-Ho Yu^b, Jennifer K. Wagner^c, Tanya M. Harrell^b, Charmaine D. Royal^d, and Michael J. Bamshad^{b,e}

^aPublic Health Genetics, University of Washington, Seattle, Washington, USA

^bPediatrics, University of Washington, Seattle, Washington, USA

^cGeisinger Health System, Center for Translational Bioethics & Health Care Policy, Danville, Pennsylvania, USA

^dDepartment of African & African American Studies, Duke University, Durham, North Carolina, USA

^eGenome Sciences, University of Washington, Seattle, Washington, USA

Abstract

Over the past decade, the proliferation of genetic studies on human health and disease has reinvigorated debates about the appropriate role of race and ancestry in research and clinical care. Here we report on the responses of genetics professionals to a survey about their views on race, genetics, and ancestry across the domains of science, medicine, and society. Through a qualitative content analysis of free-text comments from 515 survey respondents, we identified key themes pertaining to multiple meanings of race, the use of race as a proxy for genetic ancestry, and the relevance of race and ancestry to health. Our findings suggest that for many genetics professionals the questions of what race is and what race means remain both professionally and personally contentious. Looking ahead as genomics is translated into the practice of precision medicine and as learning healthcare systems offer continued improvements in care through integrated research, we argue for nuanced considerations of both race and genetic ancestry across research and care settings.

Address for correspondence: Joon-Ho Yu (joonhoyu@u.washington.edu), Pediatrics, University of Washington, 1959 NE Pacific St, Box 356320, Seattle, 98195, Washington, USA.

AUTHOR CONTRIBUTIONS: JHY, JKW, CDR, and MJB conceived of and designed the survey study and along with TMH led data collection efforts. JHY and SCN conceived of and designed the qualitative content analysis study; SCN carried out thematic coding in frequent consultation with JHY. JHY and SCN drafted the manuscript, which was critically revised for important intellectual content by JKW, CDR, and MJB. All authors gave final approval of the version to be published.

CONFLICTS OF INTEREST: None.

ETHICAL APPROVAL: This study was reviewed by the institutional review board(s) at the University of Washington and determined exempt (“Race, Ancestry, and Genetics Survey,” Protocol #44090).

INTRODUCTION

I have somehow managed to hold seemingly mutually-exclusive views that 1) races don't exist and are biologically meaningless and 2) races have a genetic basis and biological influences on health. – Genetic epidemiologist survey respondent

As a physician and public health professional knowing about the geographic/ ancestry/ "ethnic" group /"whatever you want to call it" is useful in providing appropriate care and services to specific populations. – Clinical geneticist survey respondent

For decades, human geneticists, ethicists, social scientists, and healthcare providers have discussed and debated the appropriate role of race and ancestry in biomedical research and clinical care. While there may be consensus in the scientific community that socially-defined races are not discrete taxonomic, biological, or genetic groups, disagreements remain about whether and to what extent race is a useful proxy for genetic or other biological differences between individuals. Recent public discourse further illustrates disagreement within the scientific community about the socially and scientifically appropriate way to discuss race, genetics, and health (Reich 2018; BuzzFeed 2018). The literature reveals a cyclical revisiting of these issues (Burchard et al. 2003; Cooper, Kaufman, and Ward 2003; Sankar and Cho 2002; NHGRI 2005; Caulfield et al. 2009; Bamshad et al. 2004; Peterson 2005; Bamshad 2005) which, in the extreme, calls for the disposal of race as a variable in genetic research and clinical practice (Perez-Rodriguez and de la Fuente 2017). Yudell and colleagues (2016) recently posited the use of race in genetic research has become even more problematic in the post-genomic era, mired in confusion about the degree of biological relevance of race and poised to fuel racist beliefs via popular misinterpretations. Accordingly, they called for the convening of a national committee to issue recommendations on how to “move past” the use of race as a classifying variable, a recommendation similar to one previously proposed by Jacqueline Stevens (2003).

Genetic ancestry has become the main conceptual framework used by genetic researchers to categorize individuals and is often posed as the solution for how genetic research, and subsequent translation into clinical care, can move past problematic uses of race (Burchard et al. 2003; Shields et al. 2005). Genetic ancestry is typically estimated using allele frequency differences that have arisen through geographic separation, historical migrations, selective pressure within and between human populations, and mere chance. In concert with advancements in genetic technologies, statistical methods have been developed to detect and describe these genetic differences (Alexander, Novembre, and Lange 2009; Price et al. 2006), allowing genetic researchers to adjust for genetic ancestry rather than (or in addition to) reliance upon externally or self-ascribed racial or ethnic labels when modelling relationships between genetic variation and disease. However, statistically-estimated genetic ancestry can often be conflated with conventional notions of race, particularly when ancestry is parsed at the continental level (Fujimura and Rajagopalan 2011). Thus, in practice, genetic ancestry might not enable a “move beyond” race (Bonham, Warshauer-Baker, and Collins 2005), but rather might unintentionally reinforce racialized thinking — particularly when it comes to assigning and interpreting group labels (Fullerton et al. 2010). Using the

conceptual framework of genetic ancestry might simply relocate racial thinking from macro-level biological differences to the molecular level of genes.

The complete erasure of race from research and practice has potentially undesirable consequences. For example, it might impede progress to meet calls for increased diversity in genetic studies (Popejoy and Fullerton 2016). Furthermore, race as a social reality can have biological implications that initiate or otherwise exacerbate disease processes. For example, using the construct of “embodiment” Krieger (2005) has proposed health disparities to be the “biological expression” of the social context of racial discrimination. For biomedical research to ignore the inscription of systemic racism and other social contributors to health into the body will lead to incomplete understandings of most health conditions. At the same time, proposing genetic research as the solution to interrogate and rectify health disparities might lead to a national research focus that is misaligned with the realities of social and environmental contributions to such disparities (West, Blacksher, and Burke 2017). While these concerns might in the past have been perceived as academic or of narrow interest, the proposed integration of genomic, behavioral, environmental, and social determinants information into the rapidly developing practice of precision medicine highlights the need to develop an equally nuanced and integrated understanding of race, ancestry, and health (Bonham, Callier, and Royal 2016). Furthermore, efforts to inform and realize precision medicine overlap with—and perhaps depend upon—the successful reformulation of healthcare systems and research institutions into learning healthcare systems that intentionally integrate research and clinical care under one proverbial roof (Institute of Medicine 2012; Institute of Medicine 2015; Faden et al. 2013).

Clearly, a complicated set of issues underlies debates and discussions about the use of race and ancestry in research and clinical care, and the stakes are high. While we support normative and theoretical work to continue robust and thoughtful deliberation, we are concerned that such efforts might not yield desired changes in practice and policy without more empirical research to understand how genetics professionals currently perceive and use the concepts of race and ancestry in their daily work. Specifically, empirical studies can help identify points of intervention where policies and recommendations would be most tractable and efficacious.

We sought to address this need for empirical knowledge by surveying genetics professionals regarding their views on the relationships between race, ancestry, and genetics. Survey items covered a range of topics, including science, medicine, and social issues. Here we report on a qualitative content analysis of open text comment boxes provided throughout the survey, where respondents expanded on their answers to survey items and provided overall feedback. Numerous ideas emerged from this qualitative inquiry, especially given the breadth of survey items and the number and diversity of respondents. However, in this report we focus on three interrelated questions that we think serve as a starting point for discussion about how race and genetic ancestry should be used in an era of genomic research and precision medicine. First, how do genetics professionals define and use the term “race”? Second, what are genetics professionals’ views on using race as a proxy for ancestry? Third, what are the understandings of genetics professionals regarding the influence of race and ancestry on health?

METHODS

Data collection

To explore what genetics professionals know and think about the relationships between race, ancestry, and genetics, we circulated an online survey in 2012 using distribution lists from three professional societies: the American Society of Human Genetics (ASHG) annual meeting attendees, the National Society of Genetic Counselors (NSGC) membership, and the American College of Medical Genetics (ACMG) membership. The survey consisted of 49 statements (see Supplementary Resources) divided into five sections: science (two sections), medicine, society, and common statements about race. For each statement, respondents rated their level of agreement on a five-point scale. The survey included six comment boxes: one at the end of each of the five survey sections and one at the completion of the survey. In these free-text boxes, respondents were encouraged to both clarify and elaborate on their responses to statements and provide more general feedback. The survey also collected demographic information such as educational degree, profession, self-identified race, sex, age, country of residence, and experience with ancestry testing both personally and in research settings. The survey was designed to enable both quantitative and qualitative analyses. A quantitative analysis of responses to the 49 survey statements is reported separately. Here we report a qualitative content analysis of the respondents' comments to the 49 survey statements from the survey of genetics professionals.

The University of Washington IRB determined this study ("Race, Ancestry, and Genetics Survey," Protocol #44090) exempt.

Data analysis

Survey comments were analyzed using content analysis, a form of qualitative inquiry that aims to identify and distill themes, ideas, and topics from text. Our coding process involved a blend of directed and conventional content analysis (Hsieh and Shannon 2005). First, we used directed content analysis based on our a priori research questions to develop a preliminary coding framework from a review of relevant background readings. Next, we expanded the initial coding framework using conventional content analysis: reading through 100 surveys to further derive and refine codes. Completeness of the codebook was established by comparing to another set of codes generated separately by another member of the study team. Additional codes were then added to ensure coverage of this prior codebook. Following this reconciliation process, the codebook comprised 44 thematic codes grouped into nine code families: relationship between concepts, nature of concepts, relationship to health, how concepts function, harms and benefits, responses to concepts, context, population genetics, and survey feedback. In addition to the thematic codes, an additional 49 structural codes were created to track comments directed at each of the 49 survey items. Ultimately, the codebook covered the breadth of topics addressed by the survey items, namely scientific understandings of race, genetics, and ancestry and their appropriate place in biomedical research; the perceived relationships between race, ancestry, and health and their appropriate place in individual healthcare; the societal implications of genetic ancestry testing; and more general reflections about the existence and nature of race. The codebook is available as supplementary material. One primary coder carried out all the coding, in

frequent consultation (weekly, over a three-month period) with a second member of the analysis team.

We analyzed the coded dataset via an iterative process of decontextualization and recontextualization, wherein survey comments were separated and sorted (decontextualized) and then grouped and analyzed into larger thematic units of meaning drawn from multiple respondents (recontextualized) (Ayres, Kavanaugh, and Knafl 2003). Based on our sense of broad themes, we first selected the codes most informative to this particular inquiry, namely to answer the following questions: 1) how do genetics professionals define and use the term “race”, 2) what are genetics professionals’ views on using race as a proxy for ancestry, and 3) what are the understandings of genetics professionals regarding the influence of race and ancestry on health. Second, quotations captured by each code were reviewed and organized into code summaries, a type of summary document useful for condensing and summarizing potentially broad and complex codes into their more salient and coherent subparts. These code summaries involved counting occurrences of both the overall code and each sub-theme contained within the code. We also conducted simple counting queries; co-occurrence queries, to examine relationships between themes; and counting queries stratified by different demographic variables or by responses to specific survey items. In the results section below, we primarily report quotes and occasionally give counts of quotes by code. Due to the structure of the data (open text boxes in a survey, as opposed to transcripts of interviews or focus groups), there are typically not multiple quotes with a given code from a single respondent. Coding and subsequent analyses were carried out using the qualitative data analysis software Atlas.ti, version 7.1 (2013).

RESULTS

The survey was sent to 9,858 individuals via the online tool Survey Monkey (www.surveymonkey.com). A total of 1,779 respondents took the survey (~18% response rate), of which 515 included at least one comment. These comments comprise the dataset for this qualitative content analysis. The demographics of the 515 respondents who commented are shown in Table 1; note percentages in this table are given over non-missing responses. Males and females were roughly equally represented, and the average age was 45 years. The majority of respondents (77%) indicated living in the United States and self-identified as white (83%). A variety of clinical and research professions were represented, and most (75%) reported working in an academic rather than a government or private setting. A doctoral degree was the most common (48%) degree obtained, while only 12% did not have at least one post-baccalaureate degree. Of those without a post-baccalaureate degree, 36% were students or trainees. Approximately one third of respondents indicated a profession in clinical care, specifically clinical and medical genetics (20%), or genetic counseling (14%). Other professions represented suggest respondents were engaged mainly in research (e.g., genetic epidemiologist; population geneticist; ethical, legal, and social implications (ELSI) researcher; etc.), although 18% of respondents selected more than one profession and therefore might perform research and clinical work.

Overall, the subset of respondents who made comments does not differ substantially from those who responded to the survey without comments; however, formal tests reveal some

statistically significant differences (see Table 1). Specifically, those who commented were more likely to have a doctoral degree, were more likely to use ancestry in their own research, and were of a slightly higher mean age compared to those who did not comment. A diverse subset of respondents made comments in each of the six comment boxes, such that a variety of disciplines, educational backgrounds, and age ranges were represented in each survey section. The distribution of comments across survey sections suggests that comments dealt with the range of survey topics rather than being limited to just one or two areas, e.g., primarily just science or medicine (see supplementary Table S1). Thematic codes, which became the focus of this analysis, were also distributed across different types of respondents (see supplementary Table S2). Although we did not perform an exhaustive search of each combination of codes and demographic variables, our evaluation suggests that subsequent analyses were informed by a broad rather than narrow set of respondents.

Multiple meanings of race

Survey responses in open text boxes illustrated that genetics professionals define race in multiple ways. They often identified different components of meaning, namely a biological or genetic component and a social or cultural component. They were further likely to prioritize the cultural meaning while downplaying the genetic meaning. However, despite this separation of meanings, genetics professionals appeared to struggle with the multiple meanings of race, especially as race is a historically political and social issue and as colloquial uses of the term blur the distinctions between meanings.

Two examples demonstrated how some respondents separated genetic from cultural meanings of race and ascribed greater importance and validity to the latter. One respondent largely rejected the biological aspects of race based on her understanding of scientific evidence. She explained:

Race is a concept that is unlikely to vanish in the near future, and in a cultural context it is important and meaningful to people's identities, communities and behaviors, including behaviors and attitudes related to health care. But as a biological/genetic/biomedical concept it has been shown over and over that there are no boundaries or subspecies, there is very little variation to begin with in humans relative to other primates, and what is present is amazingly unstructured given the massive geographic range in which we live. (2440, population geneticist)

In the following example, another respondent described two "types" of race and then effectively ranked them based on their perceived importance.

Where "race" is ambiguous, I have usually assumed both "genetic" and "cultural" race is intended...I think race as a cultural matter exists much more than race as a biological or genetic matter. But I think race as a cultural matter has some correlation with biology, though not a perfect one or, necessarily, a causal one. But our (Western) current views of "races" do correlate with genetic variations (and vice versa though not to the same extent) and hence I have to say they are somewhat biologically meaningful. Kind of. (8914, ELSI researcher)

These examples are representative of the many respondents who emphasized social over biological meanings of race. This trend is evidenced by the 118 quotes coded as “race not biological” versus the 19 quotes coded as “race biological.”

That respondents recognized multiple meanings of race was apparent in their comments on survey items such as “Races don’t exist” (23 quotes) and “Race has no biological basis” (8 quotes). In the following comment, the respondent clarified her responses to these survey items by explaining the “different components” of race:

Race, when used as a term that signifies cultural affiliation, does exist. It also does have some biological basis in that some ostensible physical features tend to associate with the term “Race.” But it is neither exclusively biological nor exclusively cultural and distinguishing between these two factors requires constant vigilance on part of the scientist using that term.... I believe that our task at hand is to attempt to deconstruct the term “race” and identify the different components that constitute this term and this concept. Then, to associate specific facets of “Race” with appropriate medical usage where relevant. Addressing the social connotations of the term “Race” is a whole different issue. (3555, population geneticist and genetic epidemiologist)

Her comment highlights the struggle genetics professionals may face when tasked with identifying, isolating, and appropriately applying different components of race. Further, these tasks might be complicated by colloquial or informal uses of the term “race,” which often blur the boundaries between different meanings. As this respondent reflected:

This is an interesting series of questions. I have somehow managed to hold seemingly mutually-exclusive views that 1) races don’t exist and are biologically meaningless and 2) races have a genetic basis and biological influences on health. The contradiction, [in my opinion], stems from the way we colloquially define race. (6364, genetic epidemiologist)

Race as a proxy

Genetics professionals expressed a range of views on the legitimacy and acceptability of using race as a surrogate or proxy for genetic ancestry. Their comments were grounded in their experiences/observations that at least some — albeit imperfect — overlap or correlation exists between race and genetic ancestry, specifically between traditional race categories and continental origins. Thus race might be used as a placeholder when genetic ancestry information is desired but unavailable. Comments suggested that respondents viewed this practice as balancing scientific accuracy and practical utility (see Figure 1). Specifically, many comments (~20) expressed disagreement with using race as a proxy for ancestry, namely because race is scientifically inaccurate or imprecise. Those commenters would tip the balance to the left in Figure 1, prioritizing scientific accuracy over practical utility and therefore not accepting race as a proxy for ancestry. The following quote illustrates how the decision to use race as a proxy for genetic ancestry involves weighing accuracy and utility.

Race is a short-hand way to classify genetic ancestry. Like most short cuts, it’s not a perfect measure of what it is actually meant to measure. But as long as some

genetic diseases are more common among people of certain races, it is clear that race is correlated with some differences in genetic variation and at least some of these differences affect health. (7876, genetic epidemiologist and genetic counselor)

The respondent acknowledged that race falls short of ideal accuracy (“not a perfect measure”) but also that it conveys some useful information regarding risk of genetic disease and, thus, has some practical utility. Further, whether respondents emphasized accuracy or utility when considering the use of race as a proxy for genetic ancestry depended on both the context (clinical or research) and on the availability of additional information. As this respondent explained:

We don’t use genetic ancestry in medicine, race is our proxy so it’s not really accurate to ask if genetic ancestry should be considered in the diagnosis of disease. (7732, clinical geneticist and ELSI researcher)

Numerous comments (~29) cast using race as shorthand for genetic ancestry in a more favorable light, particularly in clinical settings. Those respondents who viewed race as an imperfect yet functional proxy for ancestry were often motivated by practical considerations, such as the lack of available genetic information — tipping the scale to the right by emphasizing practical utility at the expense of scientific accuracy (Figure 1). The relative weight respondents placed on accuracy versus utility was also influenced by the availability of other information. Specifically, respondents pointed to the unavailability of genetic information (such as genetically-inferred ancestry or genotyping) in most clinical encounters as a reason why race might be an acceptable proxy. The following two quotes exemplify this position:

Because genetic ancestry data are not readily available for most patients, race, as it is currently defined, must be used as a proxy. (9174, genetic counselor)

While race may be an imprecise term, if it’s used as a synonym for ethnic background, this might be considered a proxy for estimating genetic risk of disease. Assuming that we can’t perform genetic testing on every individual because of cost, in this case “race” would be useful in many of the above scenarios. (9207, PhD level student)

As seen here, respondents following this rationale emphasized utility over accuracy, even citing cost as a practical barrier to accessing genetic information. Another respondent explained that time, in addition to cost, is also a factor:

Race is a good starting point for diagnosis of diseases, at times race does correspond with certain biological aspects of genetic ancestry. If there is [a] problem that requires immediate attention then race may be a better option than waiting for confirming genetic ancestry. This may also be expensive if it is not covered under insurance. Genetic ancestry is the best way of diagnosing differences in certain diseases and conditions, but race can sometimes be a good proxy or at least more informative than treating disease similarly for all patients. (9118, human geneticist and student)

By contrast, in the context of genetic research, respondents more often expressed a negative view of using race as a proxy for genetic ancestry. We observed this in responses to survey items “Race is a useful proxy for ancestry” and “Genetic ancestry, as inferred from genetic markers, is a better proxy for genetic relationships among continental groups.” One respondent wrote:

Ancestry, as inferred from either genetic markers or knowledge of country of origin of parents and grandparents, is the most useful proxy in genetic studies, in my opinion. (5259, medical geneticist)

While some respondents either clearly accepted or rejected the use of race as a proxy for genetic ancestry, other respondents' views fell along a spectrum. Those who did not view race as an acceptable proxy for genetic information appeared to prioritize the desire for accuracy over the practical considerations mentioned above (e.g., cost, time). As illustrated in the following quotes, this viewpoint was also distinguished by an emphasis on knowing the specific genotypes as an important aspect of accuracy.

[R]ace should not be used as a proxy for genetics in the relationship between genetics and health. Genetic markers, not race, should be considered when diagnosing and treating patients. (4399, human geneticist)

Given what we now know about genetic ancestry, classification into “races” is irrelevant. What is important is the individual's genetic makeup, and since that is getting much easier to assess, using “race” as a proxy, certainly in health care, is pretty pointless. (6592, clinical and medical geneticist)

Notably, a few respondents (~4) eschewed race as a genetic proxy and instead argued for race as a proxy for environmental and social factors, which have a critical influence on health. For example:

The consideration of “race” for health care in the US functions mostly a proxy for socio-economic factors, than any meaningful genetic or biological characteristic. (3006, genetic epidemiologist and computer scientist)

These respondents acknowledged that race can be relevant to health care, but only insofar as it conveys information about socioeconomic factors, not “genetic or biological” ones. Another respondent went into more detail on why she sees race as an important proxy for non-genetic information and why race can be a useful tool when measuring health outcomes on a societal level:

I believe that “race” is a social term/construct and that while it may be a poor proxy for ancestry, it may offer clues/suggest relationships to race and disparities in health...I would say that the use of the word “Race” is bad, but by dismissing it or trying to rid of it entirely, we may later realize that race (as a social construct of course) could be a valuable tool to measure health outcomes not necessarily related or related fully to genetic underpinnings but rather related to the health stress/burden of being labeled, for example, a black man in America. (8518, genetic counselor and research coordinator)

Race, Ancestry, and Health

Overall, we found genetics professionals' views on the influence of race or ancestry on health varied by the type of condition (monogenic vs. complex); the importance of environmental factors; and the clinical purpose (e.g., risk assessment, diagnosis, or treatment). The preponderance of comments focused on individual clinical care. We note that respondents did not always clearly distinguish between race and ancestry when commenting on health impacts, which we address further below.

Many respondents thought ancestry had an important influence on health. A total of 82 comments were assigned the code "ancestry health effect" compared to only seven comments coded as "ancestry health no effect." In specifying their position, respondents frequently gave examples of single-gene (monogenic) diseases typically associated with specific populations, such as sickle cell disease and cystic fibrosis, where the causal genetic variants (alleles) vary in frequency due to historical and geographical separation of populations that have experienced different environmental selective pressures. In the following quote, the respondent used diseases typically thought of as "population-specific" to illustrate why he thinks ancestry is important to health. This respondent also drew a parallel between ancestry and race, albeit without elaborating on exactly how and to what extent these two terms are related.

There is overwhelming genetic evidence that relatively or exclusively population-specific alleles can have a very strong influence on biology and health. Sickle cell disease does not exist in Native Americans, nor does hereditary hemochromatosis in East Asians. Those who say biological ancestry (or even "race") has no effect on health and disease have political motivations that are unfounded in the scientific facts. (0259, human geneticist and population geneticist)

This respondent used two monogenic diseases as examples of why ancestry is important to health: sickle cell disease, which is commonly associated with African Americans and West Africans (although notably, this view obscures the fact that sickle cell is common in many other populations (Braun 2002)), and hereditary hemochromatosis, which is more common in populations of European descent (Limdi and Crampton 2004). Another respondent used the example of sickle cell disease in this instance to draw a more explicit parallel between race and ancestry:

Race is a surrogate for genetic ancestry; genetic ancestry can have significant influence on the probability of inheriting alleles that can influence health, e.g., sickle-cell. (6280, population geneticist and genetic epidemiologist)

These examples illustrate the view that (1) ancestry is important to health and (2) race is similar to ancestry, thus (3) race also influences health. Other quotes illustrate agreement with point (1) but disagreement with points (2) and (3). Instead, these respondents explicitly distanced race from ancestry and argued that ancestry, not race, influences health.

"Race" does not exist, but distribution of genetic determinants of health are unevenly distributed. So knowing the ancestry background may influence health: but this is not through the concept of race: to be Ashkenazi raises the risk to have a child with Bloom syndrome, or to be French Canadian from St. Jean lake increases

the likelihood to carry [Maple syrup urine disease]: thus knowing your origin may influence the prevention screening e.g. but this is not linked to the concept of race. (8612, clinical and medical geneticist)

In this next quote, a genetic counselor explained why she asks her patients about their ancestry and not their race:

A person's ancestry may have a biological basis, meaning, genetic basis and influence on health. But it is no means the most important influence. And a person's ancestry may not actually match a specific race. I do not ask my patients to which race they belong. I ask them where their ancestors come from. (5920, genetic counselor)

In contrast to the previous examples, some respondents noted ways in which ancestry might *not* influence health. A few respondents (~4) specified this viewpoint by noting that ancestry is not relevant for common, complex diseases. For example:

[I]f a disease is heritable and it is common in your family, then your recent "ancestry" is very important. In common diseases, it is less clear that genetic ancestry matters much if at all. (9976, human geneticist, population geneticist, and genetic epidemiologist)

While this respondent clearly held ancestry to be relevant for heritable, Mendelian conditions, she was suspect of its role in common conditions.

When considering how ancestry should be used in a clinical context, some respondents (~8) distinguished between three main clinical purposes: risk assessment, diagnosis, and treatment. For example, in the following quote the respondent indicated that ancestry is a useful piece of information when determining a patient's risk of developing certain diseases. The focus on "probabilities" suggested that she was addressing risk assessment:

As a physician and public health professional knowing about the geographic/ ancestry/ "ethnic" group /"whatever you want to call it" is useful in providing appropriate care and services to specific populations. It doesn't define the patient but it's useful info regarding probabilities-- patients of African ancestry are less likely to have [cystic fibrosis] or [phenylketonuria] than those with northern European ancestry, but some do have CF/ PKU. (2966, clinical geneticist)

As in the previous sections, it is notable which diseases she used to illustrate her point: cystic fibrosis and phenylketonuria. These are both monogenic disorders for which risk allele frequencies vary considerably by population and thus where it would make sense to consider population origin when determining disease risk.

Like risk assessment, disease diagnosis might also utilize probabilities in that the likelihood of a given diagnosis can depend on the background likelihood (risk) of developing that disease. In the following quote, the respondent suggested that ancestry might be used to gauge probabilities and thus help guide diagnosis. He also extended the utility of ancestry in the diagnostic context to include "common perceptions of race" as a potentially useful proxy for ancestry. However, he drew a line between diagnosis and treatment. In contrast to disease diagnosis, a patient's treatment should *not* be influenced by race or ancestry.

Differentially diagnosing, speeding the efficiency of identifying a condition, focusing attention on possibilities and probabilities can all be aided by attention to some aspect of genetic ancestry, even if proxied by common perceptions of race, but treatment, interventions, and responding to individuals with x or y condition is independent of race or ancestry. (4796, human geneticist)

Thus respondents' views on when and how ancestry should be incorporated into the healthcare setting depended in part on the clinical task at hand: risk assessment, diagnosis, or treatment.

DISCUSSION

We have surveyed genetics professionals about their views on race, genetics, and ancestry across the domains of science, medicine, and society. We conducted a similar survey among anthropologists (Wagner et al. 2017; Ifekwunigwe et al. 2017). In this qualitative content analysis of free-text comments made by a subset of 515 respondents, we observed that genetics professionals recognize multiple, interrelated definitions of race. Fujimura and Rajagopalan (2011) observed slippage between concepts of genetic ancestry and race in biomedical research and noted the potential for subtle distinctions between the two concepts to become “lost in translation” when consumed by a broader audience (e.g., other researchers, media, members of the public). Similarly, we found that some survey respondents identified different aspects or components of the term “race,” namely a socio-cultural component and a genetic or biological component, and that slippage may occur when these different components of meaning are not well-articulated or recognized. We identified this phenomenon as the *compartmentalization of race*, which involves both the separation of race into different components and the selective use of one (or more) components based on context. Respondents inhabit and traverse different meanings of race and genetic ancestry in various contexts. While most respondents downplayed biological definitions of race in favor of social ones, many acknowledged some overlap or correspondence between race and genetic ancestry. Notably, we found that genetics professionals' flexible use of race and genetic ancestry depended heavily on context rather than subscription to any single definitive biological or social definition. Though study populations differed, similar to the Bonham et al. (2009) study of physicians we observed a greater acceptance of using race in making screening versus treatment decisions. This acceptance was complicated by the type of condition (e.g., monogenic versus multifactorial disease).

In general, respondents were more likely to accept using race as a proxy for genetic ancestry in a clinical versus research context, recognizing a tradeoff between scientific accuracy and practical utility. We have visualized this tradeoff as a balance or see-saw (Figure 1), with accuracy at one end and utility on the other. The acceptability or likelihood of using race as a proxy for ancestry, and therefore which way the balance tips, is often influenced by the availability of genetic ancestry, which historically has been less available in clinical versus research contexts. We therefore locate the potential for definitional slippage (Fujimura and Rajagopalan 2011) at the juncture of genetic research and its clinical translation. This finding suggests the need for future research to better understand how genetic ancestry-

based knowledge is translated into clinical care and the extent to which genetic ancestry estimates or even ancestry-based interpretations transfer.

For the bioethics community, identifying the translation of genomics research into clinical practice as a likely site of definitional slippage between race and ancestry raises troubling questions in this current era of translational genomics. Indeed, Wolf and colleagues have called for a new approach to the “ethics of translational genomics,” shifting away from past ethics frameworks based on the (now dissolving) divide between research and clinical care (Wolf, Burke, and Koenig 2015). A more comprehensive ethics of translational genomics will require knowledge of how geneticists are conceptualizing and deploying concepts of race and ancestry, across the spectrum from research to clinical practice — knowledge we have aimed to help build with this study.

Precision medicine and learning healthcare systems are two major trends in contemporary healthcare that merit reinvigorated attention to issues of race and ancestry, particularly in service of developing an ethics of translational genomics. As commonly envisioned, precision medicine involves integrating genomic, lifestyle, and environmental information into patient care, with the goals of personalizing prevention, diagnosis, and treatment. Unlike many of the clinical settings described by our survey respondents, both race and genetic ancestry information might soon both be commonly available, spurred in part by the continued growth of direct-to-consumer genetic ancestry testing (Regalado 2018). This portends the need for healthcare providers to articulate and examine when and why concepts of race and ancestry are used. Notably, a provider’s own race may affect her level of comfort in discussing race and ancestry with patients, a challenge that will need to be addressed to increase transparency and patient satisfaction in clinical encounters (Bonham et al. 2009). Given increasing opportunities to take clinical action based on genetic information (Manolio et al. 2013), it will be imperative for healthcare providers together with their patients to consider, coincidentally and possibly even synergistically, the role of self-identified race and its underlying social determinants in screening, diagnosis, and treatment decisions.

Furthermore, as healthcare moves towards learning healthcare systems (LHS), complexities of race, ancestry, and genetics will likely intensify (Carney and Kong 2017; D’Alessandro, O’Neil, and LaGatta 2017). LHS mark a dissolving boundary between biomedical research and clinical care. With a focus on informatics, big data, machine learning, and artificial intelligence, LHS may lead to the programmatic encoding of concepts such as race and genetic ancestry. The algorithmic “black boxing” of already complex concepts of race and ancestry will make them less visible and less interrogated at a time when they should perhaps be more visible and more scrutinized (Wagner 2018).

While machine learning is often perceived as objective, it has potential to worsen discriminatory practices across sectors, including health care, by either creating new disparities or exacerbating existing ones. d’Alessandro and colleagues (2017) give the example of an algorithm that explicitly excludes race as a factor, yet unwittingly folds it back in by including correlates of race. This scenario highlights that it is not just the definition of variables such as race and ancestry that merit attention, but also the overlap or correlations between them. Cabitza et al. (2017) further caution that the adoption of

machine-learning in medicine might obscure “qualitative and complex” factors, such as social and organizational issues. Given how the latter are often the root causes of health disparities, it is important for genetics professionals not only to consider how variables and the correlations between them are defined in LHS, but also to examine what variables might be missed altogether. In summary, studying and ultimately reducing health disparities in a LHS context will necessitate having effective and reliable variables, including for concepts of race and genetic ancestry.

Our analysis has several limitations that warrant consideration. The overall survey sample was a self-selecting subset (18% response rate) of those emailed from ASHG, NSGC, and ACMG distribution lists and therefore could be subject to nonresponse bias. The subset of survey respondents who left comments in free-text boxes, is further potentially biased in that individuals with strong views might have been more likely to add comments. We think this tendency toward more decided views is unlikely because comments often articulated more nuanced attitudes where respondents reflected on the interrelatedness between genetics professionals’ understandings of race and ancestry. We also did not observe substantial demographic differences between those who commented and those who did not (see Table 1). Nevertheless, these findings cannot be viewed as representative of views among genetics professionals writ large. The racial diversity of our sample is limited in that the majority of respondents who self-reported race identified as “white.” Indeed, of the 21 respondents whose quotes are presented in this manuscript, 17 self-identified as white, 1 as South Asian, and 3 were missing race information. It is also possible that the order of survey sections influenced respondents’ comments. Relying on a single primary coder in consultation with a second team member may have introduced coding bias into the analysis. Additionally, we must note that we did not ask respondents specifically about precision medicine or LHS, as these topics had not received much attention at the time this survey was conducted in 2012. However, despite the fact this survey was conducted several years ago, the availability of ancestry and genetic information in clinical settings has not changed much in the interim. Furthermore, this survey was designed to explore the relationship between race and ancestry, and we were therefore limited in our ability to explore the parallel concept of ethnicity; however, this area is ripe for future research. In summary, these nuanced qualitative results thus provide a starting point for deliberation—rather than conclusions—about points of apparent cognitive dissonance, such as the claim that geneticists’ unexamined use of race and by extension genetic ancestry contribute to reified understandings of race as a biological construct.

A more seamless integration between research and clinical care, as observed in precision medicine and LHS, is unlikely to resolve the complex relationships among meanings and usages of the terms race, ancestry, and genetics. Instead, if left uninspected, these terms could become algorithmically “black boxed” in data-driven practices, obscuring their meanings and the nature of the relationships between them. However, genetics professionals have an opportunity to examine their own personal and professional practices, in particular to foster awareness and transparency of how context elicits specific uses of race and genetic ancestry. These examinations should occur at both the individual and collective level, such as in the issuing or updating of explicit position statements on race and nondiscrimination from genetics professional societies (Royal et al. 2010; Watson and Greene 2001).

Cognizance of these issues among genetics professionals will help to realize the promise of genomics to improve human health without reinforcing stigmatizing and biologically-reified notions of race. More broadly, we support thoughtful consideration of self-identified race and genetic ancestry-based health information in clinical care so as to maximally benefit all patients. Genetics professionals have an essential role to play in eliminating racial discrimination in all its forms and contexts, including research and healthcare.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS:

The authors thank the survey respondents for their participation.

FUNDING: This work was supported by Grant No. R00HG006446 from the National Human Genome Research Institute (NHGRI) (JKW contribution), Grant No. R00HG007076 from the NHGRI (JHY contribution), and the Greenwall Foundation and Grant No. R01HG006295 from the NHGRI (CDR contribution).

REFERENCES

- Alexander David H, John Novembre, and Kenneth Lange. 2009 “Fast Model-Based Estimation of Ancestry in Unrelated Individuals.” *Genome Research* 19 (9): 1655–64. doi:10.1101/gr.094052.109. [PubMed: 19648217]
- Atlas.ti. 2013 “Version 7.1 [Computer Software].” Berlin, Germany.
- Ayres Lioness, Kavanaugh Karen, and Knafel Kathleen A. 2003 “Within-Case and Across-Case Approaches to Qualitative Data Analysis.” *Qualitative Health Research* 13 (6): 871–83. doi: 10.1177/1049732303255359. [PubMed: 12891720]
- Bamshad Michael. 2005 “Genetic Influences on Health.” *JAMA* 294 (8): 937. doi:10.1001/jama.294.8.937. [PubMed: 16118384]
- Bamshad Michael, Wooding Stephen, Salisbury Benjamin a, and Stephens J Claiborne. 2004 “Deconstructing the Relationship between Genetics and Race.” *Nature Reviews. Genetics* 5 (8): 598–609. doi:10.1038/nrg1401.
- Bonham Vence L., Callier Shawneequa L., and Royal Charmaine D.. 2016 “Will Precision Medicine Move Us beyond Race?” *New England Journal of Medicine* 374 (21). Massachusetts Medical Society: 2003–5. doi:10.1056/NEJMp1511294. [PubMed: 27223144]
- Bonham Vence L, Sellers Sherrill L, Gallagher Thomas H, Frank Danielle, Odunlami Adebola O, Price Eboni G, and Cooper Lisa A. 2009 “Physicians’ Attitudes toward Race, Genetics, and Clinical Medicine.” *Genetics in Medicine* 11 (4): 279–86. doi:10.1097/GIM.0b013e318195aaf4. [PubMed: 19265721]
- Bonham Vence L, Warshauer-Baker Esther, and Collins Francis S. 2005 “Race and Ethnicity in the Genome Era: The Complexity of the Constructs.” *The American Psychologist* 60 (1): 9–15. doi: 10.1037/0003-066X.60.1.9. [PubMed: 15641917]
- Braun Lundy. 2002 “Race, Ethnicity, and Health: Can Genetics Explain Disparities?” *Perspectives in Biology and Medicine* 45 (2): 159–74. doi:10.1353/pbm.2002.0023. [PubMed: 11919376]
- Burchard Esteban González, Ziv Elad, Coyle Natasha, Gomez Scarlett Lin, Tang Hua, Karter Andrew J, Mountain Joanna L, Pérez-Stable Eliseo J, Sheppard Dean, and Risch Neil. 2003 “The Importance of Race and Ethnic Background in Biomedical Research and Clinical Practice.” *New England Journal of Medicine* 328 (12): 1170–75. doi:10.1056/NEJMs025007.
- BuzzFeed. 2018 “How Not To Talk About Race And Genetics,” 3 30 https://www.buzzfeed.com/bfopinion/race-genetics-david-reich?utm_term=.fymKyZQy0#.di42OXNOM.

- Cabitz Federico, Rasoini Raffaele, and Gensini Gian Franco. 2017 “Unintended Consequences of Machine Learning in Medicine” *JAMA* 318 (6). American Medical Association: 517. doi:10.1001/jama.2017.7797. [PubMed: 28727867]
- Carney Timothy Jay, and Kong Amanda Y.. 2017 “Leveraging Health Informatics to Foster a Smart Systems Response to Health Disparities and Health Equity Challenges.” *Journal of Biomedical Informatics* 68: 184–89. doi:10.1016/j.jbi.2017.02.011. [PubMed: 28214562]
- Caulfield Timothy, Fullerton Stephanie M, Ali-Khan Sarah E, Arbour Laura, Burchard Esteban G, Cooper Richard S, Hardy Billie-Jo, et al. 2009 “Race and Ancestry in Biomedical Research: Exploring the Challenges.” *Genome Medicine* 1 (1): 8. doi:10.1186/gm8. [PubMed: 19348695]
- Cooper Richard S, Kaufman Jay S, and Ward Ryk. 2003 “Race and Genomics.” *New England Journal of Medicine* 348 (12): 1166–70. doi:10.1056/NEJMs022863. [PubMed: 12646675]
- D’Alessandro Brian, O’Neil Cathy, and Gatta Tom La. 2017 “Conscientious Classification: A Data Scientist’s Guide to Discrimination-Aware Classification.” *Big Data* 5 (2): 120–34. doi:10.1089/big.2016.0048. [PubMed: 28632437]
- Faden Ruth R, Kass Nancy E, Goodman Steven N, Pronovost Peter, Tunis Sean, and Beauchamp Tom L. 2013 “An Ethics Framework for a Learning Health Care System: A Departure from Traditional Research Ethics and Clinical Ethics.” *The Hastings Center Report Spec No* (February): S16–27. doi:10.1002/hast.134.
- Fujimura JH, and Rajagopalan R. 2011 “Different Differences: The Use of ‘genetic Ancestry’ versus Race in Biomedical Human Genetic Research.” *Social Studies of Science* 41 (1): 1–22. <http://sss.sagepub.com/content/41/1/5.short>.
- Fullerton Stephanie M, Yu Joon-Ho, Crouch Julia, Fryer-Edwards Kelly, and Burke Wylie. 2010 “Population Description and Its Role in the Interpretation of Genetic Association.” *Human Genetics* 127 (5): 563–72. doi:10.1007/s00439-010-0800-0. [PubMed: 20157827]
- Hsiu-Fang Hsieh, and Shannon Sarah E. 2005 “Three Approaches to Qualitative Content Analysis.” *Qualitative Health Research* 15 (9): 1277–88. doi:10.1177/1049732305276687. [PubMed: 16204405]
- Ifekwunigwe Jayne O., Wagner Jennifer K., Yu Joon-Ho, Harrell Tanya M., Bamshad Michael J., and Royal Charmaine D.. 2017 “A Qualitative Analysis of How Anthropologists Interpret the Race Construct.” *American Anthropologist* 119 (3): 422–34. doi:10.1111/aman.12890. [PubMed: 30078844]
- Institute of Medicine. 2012 “Best Care at Lower Cost: The Path to Continuously Learning Health Care in America.” <http://www.nationalacademies.org/hmd/Reports/2012/Best-Care-at-Lower-Cost-The-Path-to-Continuously-Learning-Health-Care-in-America>.
- Institute of Medicine. 2015 “Genomics-Enabled Learning Health Care Systems: Gathering and Using Genomic Information to Improve Patient Care and Research: Workshop Summary.” Washington, D.C.: National Academies Press. doi:10.17226/21707.
- Krieger Nancy. 2005 “Embodiment: A Conceptual Glossary for Epidemiology.” *J Epidemiol Community Health* 59: 350–55. doi:10.1136/jech.2004.024562. [PubMed: 15831681]
- Limdi JKK, and Crampton JRR. 2004 “Hereditary Haemochromatosis.” *Qjm* 97 (6): 315–24. doi:10.1093/qjmed/hch065. [PubMed: 15152104]
- Manolio Teri A, Chisholm Rex L, Ozenberger Brad, Roden Dan M, Williams Marc S, Wilson Richard, Bick David, et al. 2013 “Implementing Genomic Medicine in the Clinic: The Future Is Here.” *Genetics in Medicine: Official Journal of the American College of Medical Genetics* 15 (4): 258–67. doi:10.1038/gim.2012.157. [PubMed: 23306799]
- NHGRI. 2005 “The Use of Racial, Ethnic, and Ancestral Categories in Human Genetics Research.” *American Journal of Human Genetics* 77 (4): 519–32. doi:10.1086/491747. [PubMed: 16175499]
- Perez-Rodriguez Javier, and Fuente Alejandro de la. 2017 “Now Is the Time for a Postracial Medicine: Biomedical Research, the National Institutes of Health, and the Perpetuation of Scientific Racism.” *The American Journal of Bioethics* 17 (9): 36–47. doi:10.1080/15265161.2017.1353165.
- Peterson Susan K. 2005 “The Role of the Family in Genetic Testing: Theoretical Perspectives, Current Knowledge, and Future Directions.” *Health Education & Behavior: The Official Publication of the Society for Public Health Education* 32 (5): 627–39. doi:10.1177/1090198105278751. [PubMed: 16148209]

- Popejoy Alice B., and Fullerton Stephanie M.. 2016 “Genomics Is Failing on Diversity.” *Nature* 538 (7624): 161–64. doi:10.1038/538161a. [PubMed: 27734877]
- Price Alkes L, Patterson Nick J, Plenge Robert M, Weinblatt Michael E, Shadick Nancy A, and Reich David. 2006 “Principal Components Analysis Corrects for Stratification in Genome-Wide Association Studies.” *Nature Genetics* 38 (8): 904–9. doi:10.1038/ng1847. [PubMed: 16862161]
- Regalado Antonio. 2018 “2017 Was the Year Consumer DNA Testing Blew Up.” *MIT Technology Review*, 2 12 <https://www.technologyreview.com/s/610233/2017-was-the-year-consumer-dna-testing-blew-up/>.
- Reich David. 2018 “How Genetics Is Changing Our Understanding of ‘Race.’” *New York Times*, 3 23 <https://www.nytimes.com/2018/03/23/opinion/sunday/genetics-race.html>.
- Royal Charmaine D, Novembre John, Fullerton Stephanie M, Goldstein David B, Long Jeffrey C, Bamshad Michael J, and Clark Andrew G. 2010 “Inferring Genetic Ancestry: Opportunities, Challenges, and Implications” *American Journal of Human Genetics* 86 (5). The American Society of Human Genetics: 661–73. doi:10.1016/j.ajhg.2010.03.011. [PubMed: 20466090]
- Sankar Pamela, and Cho Mildred K. 2002 “Toward a New Vocabulary of Human Genetic Variation.” *Science (New York, N.Y.)* 298 (5597): 1337–38. doi:10.1126/science.1074447.
- Shields Alexandra E, Michael Fortun, Hammonds Evelyn M, King Patricia a, Lerman Caryn, Rapp Rayna, and Sullivan Patrick F. 2005 “The Use of Race Variables in Genetic Studies of Complex Traits and the Goal of Reducing Health Disparities: A Transdisciplinary Perspective.” *The American Psychologist* 60 (1): 77–103. doi:10.1037/0003-066X.60.1.77. [PubMed: 15641924]
- Stevens Jacqueline. 2003 “Racial Meanings and Scientific Methods: Changing Policies for NIH-Sponsored Publications Reporting Human Variation.” *Journal of Health Politics, Policy and Law* 28 (6): 1033–88. doi:10.1215/03616878-28-6-1033.
- Wagner Jennifer K. 2018 “Data Disparities: Discrimination & Equity in Big Data.” In Pacific Symposium on Biocomputing. <http://psb.stanford.edu/previous/psb18/keynote/>.
- Wagner Jennifer K, Joon-Ho Yu, Ifekwunigwe Jayne O, Harrell Tanya M, Bamshad Michael J, and Royal Charmaine D. 2017 “Anthropologists’ Views on Race, Ancestry, and Genetics” *American Journal of Physical Anthropology* 162 (2). Wiley-Blackwell: 318–27. doi:10.1002/ajpa.23120. [PubMed: 27874171]
- Watson Michael S, and Greene Carol L. 2001 “Points to Consider in Preventing Unfair Discrimination Based on Genetic Disease Risk: A Position Statement of the American College of Medical Genetics.” *GENETICS IN MEDICINE* 3 (6). <https://www.nature.com/articles/gim200184.pdf?origin=ppub>.
- West Kathleen McGlone, Blacksher Erika, and Burke Wylie. 2017 “Genomics, Health Disparities, and Missed Opportunities for the Nation’s Research Agenda.” *JAMA* 318(9): 9–10. doi:10.1001/jama.2017.3096.
- Wolf Susan M., Burke Wylie, and Koenig Barbara A.. 2015 “Mapping the Ethics of Translational Genomics: Situating Return of Results and Navigating the Research-Clinical Divide.” *J Law Med Ethics*. doi:10.1002/jmri.23741.Proton.
- Yudell Michael, Roberts Dorothy, DeSalle Rob, and Tishkoff Sarah. 2016 “Taking Race out of Human Genetics.” *Science* 351 (6273): 564–65. doi:10.1126/science.aac4951. [PubMed: 26912690]

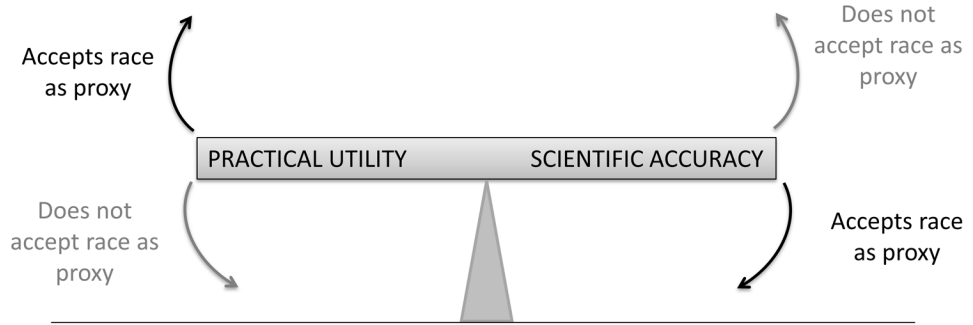


Figure 1. Schematic of the balancing between practical utility and scientific accuracy in determining the acceptability of using race as a proxy for genetic information.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 1.

Demographics of survey respondents, both the overall set of 1,779 who took the survey and the subset of 515 respondents who made comment(s). The percentages are calculated over non-missing data only. Each variable presented here was tested for statistically significant differences between the group of respondents who commented (n=515) versus those that did not comment (n=1,264). Chi-square tests were used for all categorical variables except those with less than five observations in a given category, where Fisher's exact test was used. For the continuous variable "age," the Welch Two-Sample t-test was used.

Characteristic	Overall (n=1,779)		Made comments (n=515)		
	Frequency	Percent	Frequency	Percent	
Sex	Male	622	42%	217	45%
	Female	865	58%	260	55%
	Missing	292	-	38	-
Age*	mean age (range; SD)	43 (20-92; 13.6)		45 (20-89; 13.8)	
	32 and under	377	27%	98	22%
	33-41	349	25%	105	23%
	42-54	360	25%	120	26%
	55 and over	333	23%	131	29%
	missing	360	-	61	-
Lives in US	Yes	1,143	77%	368	77%
	No	335	23%	109	23%
	Missing	301	-	38	-
Self-identified race[†]	Asian	95	8%	26	7%
	Black, African American	15	1%	4	1%
	White	966	84%	313	83%
	Other	47	4%	22	6%
	More than one	25	2%	10	3%
	Missing	631	-	140	-
Profession**	Clinical Geneticist and MD Medical Geneticist	251	17%	64	13%
	Non-MD Medical Geneticist	102	7%	32	7%
	Human Geneticist	187	13%	59	12%
	Population Geneticist	96	6%	41	9%
	Genetic Epidemiologist	113	8%	37	8%
	Genetic Counselor	277	19%	69	14%
	ELSI Researcher	17	1%	8	2%
	Student or Trainee	110	7%	31	6%
	More than one	196	13%	88	18%
	Other	143	10%	51	11%
Missing	287	-	35	-	
Education level**	Baccalaureate or less	221	15%	56	12%
	Masters	304	21%	80	17%
	PhD	606	41%	230	48%
	MD	197	13%	64	13%

Characteristic	Overall (n=1,779)		Made comments (n=515)		
	Frequency	Percent	Frequency	Percent	
	MD/PhD	147	10%	48	10%
	Missing	304	-	37	-
Work environment	Academic	1,080	73%	357	75%
	Government	104	7%	36	8%
	Private	304	20%	84	18%
	Missing	291	-	38	-
Uses genetic ancestry in their research *	Yes	719	48%	256	54%
	No	703	47%	203	43%
	Unsure	67	4%	17	4%
	Missing	290	-	39	-

* p -value<0.05,

** p -value <0.001

† No survey respondents selected “American Indian or Alaska Native” or “Native Hawaiian or Other Pacific Islander,” except in combination with another selected category (i.e. counted under “More than one”).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript