05E 048I

Project Number: SMP-9781 -46

The Ethics of Race in Genetic Research

An Interactive Qualifying Project Report

submitted to the Faculty of

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Bachelor of Science

By

Virginia O'Connell

Date: June 30, 2005

Professor Samuel M. Politz, Advisor

Abstract

Arguments for and against the use of race in genetic research and drug development are discussed. The opinions of prominent researchers in this field were examined and the consequences that large scale projects such as the Human Genome Project could have on ethnic groups were considered. Researchers and ethnic groups need to come to a common understanding to develop solutions, ideally resulting in more accurate treatments for disease.

In recent years, large advances in the fields of genome mapping and genetic research have been made, particularly involving the human genome. Included in the large amounts of data being generated is information that possibly could define useful differences between human populations. This possibility is causing a great deal of debate as to how this information should be ultimately used. Ethical concerns include the social impacts of defining race on a genetic level, in particular, the ways in which genetic information might be used to design race-specific drugs for medical purposes (Feldman, Lewontin, King 2003). Viewpoints concerning the development of such drugs differ, based on interpretations of genetic race studies already completed (Wilson, Weale, Smith 2001). These points of view include designing studies that either recognize traditionally-defined racial groups as valid genetic clusters, or ignore traditional labels in favor of purely genetic characteristics (Aldhous 2002). But the topic of race is not solely limited to the realm of science. As studies are done involving minorities, these groups are worried that genetic information will be used by racists who will then be able to define them as genetically inferior (Wadman 2004). In order to prevent future discrimination, many organizations involved in collecting data from the human genome are discussing its ethical implications prior to the release of the data.

Several major groups are involved in projects designed to determine the differences and similarities between humans of different ethnicities. Three of these major projects are the Human Genome Project (HGP), the International HapMap (haplotype mapping) project, and the Human Genome Diversity Project (HGDP), still in development. Although the Human Genome Project claims to be the first to include an

ethics department in its project budget, the Human Genome Diversity Project has been discussing ethical and social consequences with the groups it plans to include in its study since the early 1990's. Each project is employing a different approach towards collecting and analyzing information. The Human Genome Project (http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml) was completed in 2003, having achieved its goal of determining the sequence of all of human genomic DNA. The sequence is currently being analyzed to identify genes. Ethical, legal and social issues were discussed during the course of the project, including privacy and legalization, social consequences of genetic testing, and the impact genetic research might have on minority groups. The HGP website claims that this ethical, legal and social initiative (ELSI) is the world's largest bioethics program, and "the first large scientific undertaking to address potential ELSI implications arising from project data." The International HapMap Project (http://www.hapmap.org) is a three year project designed to create a database of genetic differences and similarities between human beings, with six countries participating (Japan, United Kingdom, Canada, China, Nigeria, and the United States). The information from this project is intended to be made available in the public domain, and the ultimate goal for the use of the data is to develop individualized drugs based on gene response. The sample groups consist of four populations with members of African, Asian and European ancestry, and Single Nucleotide Repeats, or SNPs, are used to determine the difference between these populations. The Human Genome Diversity Project, or HGDP, (http://www.stanford.edu/group/morrinst/hgdp/faq.html) was conceptualized in the early 90's, but has not yet started – the project organizers are extremely concerned about the

ethical implications of the results. Their goal is to ensure that the world's population is included within their sample database, in order that the genome will not be known purely in terms of a "small sample of people of European origin," which they claim is the current standard. One of the project's goals is to determine what the people of the ethnic groups they are sampling think of the project and to discuss with them potential uses of the results. Language is used as a distinguishing feature between population groups in this study, and it is estimated that there are between 4000 and 8000 of these groups, all of which will eventually be included in the project. Key genome characteristics used as criteria to distinguish racial groups include single nucleotide polymorphisms, or SNPs, which are small genetic changes, generally occurring outside the coding sequences of DNA. They often occur in the same region of DNA among individuals of similar genetic ancestry. (http://www.ncbi.nlm.nih.gov/About/primer/snps.html) Linkage disequilibrium, another distinguishing factor considered, is defined as the tendency for alleles at separate sites to be found together more frequently than would be expected by chance (Goldstein 2001).

The issue of race is fairly recent, in terms of duration of human existence, as an early population of humans would only encounter others that were similar in ethnicity to themselves, due to the limitations of travel (Smedley 2005). The concept of race only developed after these individual population groups came into contact with one another, and a need to develop a social hierarchy arose (Smedley 2005). In order to study race from a biological standpoint, it is important to understand the anthropological and historical perspectives that have developed. The distinction between ethnicity and race is important to keep in mind as well – race refers to the innate biological difference between

humans, while ethnicity covers the social and anthropological differences that have developed over time, and can be learned (Smedley 2005). Historically, several distinguishing characteristics have been used to determine "racial" groups – physical characteristics, language, religion or politics being several. This method of determining groupings is inaccurate, once one considers the rate of travel and intermarriage occurring in the modern world. In the early 18th century, the term 'race' was developed for social categories that signified class structure. This usage has continued, and it is recognized today that inequality is a trait that marks racially defined cultures, in which each race is segregated by physical characteristics and cultural behaviors that are assumed to be universal for that race (Smedley 2005). The unclear definition of the terms 'race' and 'ethnicity' in current research is an issue that needs to be resolved to prevent future confusion, as well as creating a uniformity of methods and results. From a public policy standpoint, this inequality means that necessary resources, such as medical care, would be inaccessible to those most discriminated against (Smedley 2003). Inequality in such a situation would imply that biological genetic differences would not be the only factor in determining a group's susceptibility for a specific disease (Smedley 2005).

The need for discussion involving the individuals and ethnic groups that have the potential to be most affected by the results of genetically based racial studies is a cause of great concern to many within those groups. In the past, many of the ethnic groups that might be included in these studies have been treated badly, or discriminated against. Their concern is that this sort of treatment could happen again, but this time with the support of scientific data. The Human Genome Diversity Project has been encouraging

discussion and eliciting input from minority groups, in order to gain their perspective on the matter and to improve future work.

Frank Dukepoo, a Native American geneticist at Northern Arizona University, highlights the need for education of both the public and groups that may be participating in genetically based ethnic/racial studies (Dukepoo 1999). He highlights the point that the HGDP and the HGP need to be identified as two different projects, as Native Americans have focused more on the possible consequences of the HGDP, still in the planning stages, than any of the other genome projects which are better funded and already underway. The HGDP has established a Model Ethical Protocol, which some Native Americans view as a "device for Western scientists to conduct research *on* rather than *with* indigenous people" (Dukepoo 1999). This viewpoint was further reinforced by the HGDP announcing their intent to research Native American genomics rather than extending an invitation for Native American participation. Even if another project with similar goals were to develop a better method for interacting with ethnic groups used in its study, the memory of poor communication such as this would bias people enough that they would refuse participation. Another danger, Dukepoo points out, is that Native American focus has been almost exclusively on the HGDP, while other programs such as the HGP and studies done by the NIH are also planning to include Native Americans in their sample populations (NIH Conference 1998). The NSF's plant genome initiative leaves many tribes fearful that their traditional medicinal plants will be patented, forcing native healers to cease practice of their art (Dukepoo 1999). Many Native Americans feel as though they are "being researched to death by outsiders" and are not conducive to agree to further research. Racial stereotypes have been enforced by research in the past,

as the emphasis on Native American subjects in alcohol studies has shown (Goldman 1997). Many are offended that the image of the "drunk Indian" has been perpetuated and fear that similar insensitivity will be involved in genome studies. (Dukepoo 1999)

Felix Konotery-Ahulu, an African physician specializing in genetic counseling, feels that the current organizers of genomic studies are overlooking past incidents of "scientific racism" (Konotey-Ahulu 1999). Even if information is false, or inconclusive, the media will not necessarily interpret it as such. This was the case in an AIDS study done in 1987, which had concluded that Africans were predisposed to getting the disease (Konotey-Ahulu 1987). However, the group that the data was gathered from was anthropologically distinct from other African ethnic groups, and could not possibly be used to make such sweeping generalized statements. The authors of the study then withdrew the paper due to "erroneous data," (Eales, Nye, and Pinching 1988) but the media ignored the retraction, and public opinion caused discrimination against Africans (Konotey-Ahulu 1999). Other studies that have led to incorrect popular assumptions include a book by Steve Jones, which says that "the sickle cell gene is normally found only in Africans," a statement which is incorrect (Konotey-Ahulu 1999). In 1971, it was suggested that black travelers be separated from other passengers in airports to be tested for sickle cell anemia (Green, Huntsman, and Serjeant 1971), a statement that drew much outrage, as again, it was based on false assumptions. In countries where there are deep racial divides and strong racial tensions, there is a much higher likelihood of information being misused, or used to reinforce traditional stereotypes and derogatory treatment. Konotery-Ahulu's concern is that the researchers picked to head the research will have to

be able to keep an open mind, and not just use their findings to explain stereotypes or ethnic traditions.

Trefor Jenkins of the South African Institute for Medical Research describes the plans that the HGDP has in the case of scientific discoveries being misused for racist purposes – the HGDP scientists would organize into response teams to inform the public through the media (Jenkins 1999). Jenkins believes that if the researchers were available for answering questions then racist views would be easily countered – his opinion is that in addition to carrying out the research, scientists have a responsibility to stand by their results and defend them from misuse. Since the HGDP is a group of researchers working in a collaborative format, defending and clarifying results would be much easier than if the work were carried out on an individual level, where it would be easier for miscommunication to occur.

Jenkins mentions an incident that occurred in South Africa's apartheid past which highlights the type of misuse of information that many fear – the former head of the country's chemical and biological warfare program was researching a drug that would selectively sterilize only those of a certain ethnicity – in this case the African population (Jenkins 1999). Jenkins suggests that the government should be aware of gene sequences that could potentially be used for biological or genetic terrorism so that in the case of an emergency, such as the situation described above, an antidote or vaccine could be developed quickly. Although Jenkins is in favor of the HGDP, he feels that the inhabitants of Africa "stand to be denied any compensation if they suffer adverse effects from their voluntary participation."

Several different theories have been proposed describing how to structure, interpret and use results that are found when genetic data is analyzed. There is both support for and disagreement with the concept of using self defined ethnic groups as a valid organizational method in scientific research. One controversy is whether information that is uncovered in analyzing genetic data should be used to develop drugs for use by certain ethnic groups who are more prone towards a specific disease, or if medicine should be individualized, assuming individuals cannot be defined in terms of ethnic groups.

Neil Risch is in support of the idea that there is "great validity in racial/ethnic self-categorizations," (Risch, Burchard, and Ziv 2002) which applies both to public policy decisions and research decisions. Race in this instance is defined as a quantitative factor, which can indicate a risk factor for disease, in the same way that blood pressure, gender, or exposure to environmental toxins are used now. Ethnicity would be used as one more indicator used to calculate the overall risk factor of an individual when diagnosing a disease. Risch gives the examples of breast cancer occurring more frequently in women than in men, and lung cancer occurring more frequently in smokers than in non-smokers to illustrate this point. Risch believes that it is unlikely everyone needing to be diagnosed will be able to afford full genetic testing, and instead medical practice must rely on an assessment of their total risk factors, of which ethnicity is one. Since genetic differentiation among various ethnic groups is greatest when members of a particular ethnic group inbreed and are isolated, and this differentiation is reduced when migration and intermating occurs, Risch points out that "genetic differentiation is greatest when defined on a continental basis," since through much of history, the human race did

not have the mobility and ease of travel available today. Using this logic, it can be assumed that races are able to be categorized according to their continent of origin, while ethnicities can be categorized on a self defined basis, including factors such as geographic, social, cultural and religious grounds. Ancestry is defined as the race and ethnicity of an individual's ancestors, whatever the current affiliation of the individual is. This distinction between the three terms makes discussion much clearer. Risch analyzes the data from the study of James Wilson (Wilson, Weale, and Smith 2001) to come to the conclusion that "self-defined race, ethnicity or ancestry are actually more genetically informative that clusters based on analysis of random genetic markers." The differentiation in genetic clusters is useful only in groups whose ancestors diverged many millennia ago, and not useful in situations where the groups have been recently separated, resulting in smaller genetic differentiations. Because of this, the self-defined groupings of race and ethnicity should be more than adequate to differentiate between individuals for medical and diagnostic purposes. Risch does not suggest a 'race-neutral' approach to medical research, but rather says that more research needs to be done, by studying ethnic groups individually to determine the disease risk and drug response that best applies to each of them.

Charles Rotimi, however, brings up the viewpoint that racial labels should be based not on the present location of an individual or group of individuals, but on the ancestry of any one individual (Rotimi 2004). Due to the ease of travel in the modern world, the definition of any one person's ancestry could be varied greatly, even among a group of seemingly ethnically similar persons. Rotimi says, "We must be willing to move beyond old and simplistic interpretations of differential frequencies of disease

variants by poorly defined social proxies of genetic relatedness like 'race.'" His point is made clear when he points out that "observed patterns of geographical differences in genetic information do not correspond with our notion of social identity, including 'race' and 'ethnicity.'" Other of these social identities include tribal background, geopolitical boundaries, language, and other social and behavioral activities. The terms 'race' and 'ethnicity' here are used interchangeably. Furthermore, Rotimi suggests the concept of differentiating ethnic groups and socially identified groups via genetics is as ridiculous a concept as defining all citizens of a country to be genetically similar, without taking into consideration the effects of immigration and emigration. He admits that there are some groups that can be defined genetically which are also similar racially, but points out that there are always exceptions that need to be considered. For example, Rotimi brings up the fact that Tay-Sachs disease, which has been noted more frequently in individuals with Ashkenazi Jewish ancestry, also occurs in individuals with no history of Jewish ancestry (Rotimi 2004). Cystic Fibrosis, as well, can be most often found in those of European ancestry, but also occurs in those with no European ancestor. If the ultimate goal is to use the definitions of race that projects that the HGP and International HapMap project discover to produce drugs better suited to certain ethnic groups, then every individual has to be taken into consideration. As Rotimi says, "variation is continuous, it is discordant with race, and the future categorization of groups for drug development and treatment will probably not correspond to our current sociopolitical group definitions."

In addition to large scale projects designed to collect a great amount of data concerning race and genetics, such as the HGDP and the HapMap project, there is also

work being done on an individual level. The conclusions drawn from this type of work give insight into results that might occur in these large scale projects.

A study by Hinds, Stuve and Nilsen (2005) comes to the conclusion that there is a correlation between SNP alleles and genomic differences, defining groups of European, African and Asian ancestry. A study done by Wilson, Weale and Smith (2001) on variable drug response among genetic clusters of individuals finds that "commonly used ethnic labels are both insufficient and inaccurate representations of the inferred genetic clusters, and that drug-metabolizing profiles ... differ significantly among the clusters." These authors address the fact that there are commonly known inter-ethnic differences among drug responses, but strives to illustrate that genetically inferred clusters can be made without knowing the ethnicity or geographic origin of the individual, and thus are more informative than commonly used ethnic and racial labels. They suggest that since the clinical significance of the difference in drug response is so high, it should be "a clinical priority to assess genetic structures as a routine part of drug evaluation."

The groups of individuals chosen to be tested can skew the data collected to read in the way either side of the debate wishes it to read. Goldstein and Hirschhorn (2004) bring up this concern, bringing up publication bias, the favoring of positive results, might cause skewed or irrelevant results which would then be incorrectly included in further studues. Risch (2002) points out the discrepancies evident in a study such as that of Wilson (2001). Three "ethnically defined" clusters were used - Caucasians (Norwegians, Ashkenazi Jews, and Armenians), Africans (Bantus, Afro-Caribbeans, and Ethopians), and Asians (Chinese and New Guineans). These clusters are based on geography, but the decision to include New Guinians in the same ethnic category as the Chinese is

controversial, as previous population genetic studies have shown New Guinians to be ethnically and racially defined as Pacific Islanders, who are significantly different than the Chinese, being ethnically and racially defined as Asians (Risch 2002). Arranging the ethnic groups in this manner ensures that a significant difference will be found between members of the same "ethnic group," even though the grouping initially was flawed (Risch 2002). Serre and Pääbo (2004) discussed ways in which study design is able to influence conclusions. For example, the decision to organize groups based on geographical location versus ethnically defined populations tends to produce results defining racial groups as "isolated by distance," whereas "population" grouping tends to produce results defining racial groups as being of their "continent of origin" (Serre and Pääbo 2004). Serre and Pääbo conclude that if samples are taken homogeneously worldwide, there are no clusters resulting, rather a gradient of alleles, showing no major genetic differences between races.

The consequences of studies based on defining ethnic groups as genetically similar have already been demonstrated in several instances. Drugs (e.g. BiDil) have been manufactured for a specific ethnic group, and there are cases of misdiagnosis concerning diseases that are commonly associated with members of a certain ethnicity. One example is that of sickle cell anemia, commonly associated with those that are phenotypically black. The debate as to whether to manufacture drugs on an ethnic group level or an individual level involves deciding which tactic will most benefit the people affected.

BiDil is a drug specifically designed to treat heart failure in African Americans (Taylor, Cohn, and Worcel 2002). The initial human trial for the drug was sponsored by

the Association of Black Cardiologists, the (American) National Medical Association, and members of the Congressional Black Caucus. The trial was stopped early, not because the drug was a failure, but because it was so successful. With such resounding support from the ethnic group the drug was designed to help, and the success of the trial, it seems as though this type of drug is a glimpse into the future of medicine. But by specifying a drug such as BiDil for a particular ethnic group, it negates the possibility that another ethnic group, which may have similar positive results in reaction to the drug will ever benefit from it. And if an individual who appears to be of a particular ethnic group but does not have the same ancestral background as those the drug would benefit takes it, it is reasonable to assume that they might not benefit from it (Rotimi 2004).

The study done on BiDil, however, came to the conclusion that "African Americans between the ages of 45 and 64 are 2.5 times more likely to die of heart failure than Caucasians in the same age range" (Nitromed 2005). This seems dramatic, until one considers the group surveyed – this age group is responsible for only 6% of heart failure related deaths, and once the age group of 65 and over is considered, the statistical differences between racial groups are negligible (Duster 2005). The difference in death rate can also be explained from a social viewpoint, as darker skin color in the United States has been correlated with poor medical care (Duster 2005). This sort of manipulation of statistical data is another example of results being made to appear in favor of any argument, similar to the manipulation of experimental design highlighted previously.

The possibility that one might not benefit from an ethnically targeted drug is highlighted in one case involving a boy who appeared to be phenotypically European

(Witzig 1996). The 8 year old was suffering from acute abdominal pain and anemia, and surgery was considered, until a technician discovered that the child in fact tested positive for sickle cell anemia, a diagnosis previously not considered, as the disorder is most common among those who are phenotypically black. The child, however, carried the markers for sickle cell anemia, and his parents were from Grenada, being of Indian, northern European and Mediterranean ancestry. As these are not areas commonly known for producing high frequencies of sickle cell anemics, the diagnosis was not considered.

Another case illustrating this point involves a man who was classified as black during his medical history, who was also experiencing abdominal pain, similar to that of the boy in the previous example. He told doctors that he had been previously informed that he had "sickle cell," but it had never been treated. He was then treated for the abdominal pain as though it were sickle call anemia. The next morning, the man fell into cardiac arrest and could not be resuscitated, ultimately dying due to blood loss from a bleeding peptic ulcer. The sickle cell trait or disease couldn't be confirmed (Witzig 1996).

In order to facilitate the correct treatment of diseases affecting individuals, a consensus needs to be reached among researchers, the medical profession, and the population as a whole. Problems arise when any one point of view is discounted, as seen in the previous examples. In order to solve these problems, both the researchers and their subjects need to recognize their blind spots on the topic. For instance, a member of an ethnic population not in favor of using information for fear that it will affect them negatively doesn't take into account the advantages that genetically tailored individualized health care might bring them. Members of the scientific community, on

the other hand, tend to overemphasize the scientific factors of racially related genetic studies, and not focus enough attention on possible social consequences. The design of experiments involving these data needs to be consistent, in order to prevent confusion and differing results based on differing sampling strategies. Hopefully, in the future, these conflicting methods of study will be resolved.

Bibliography

Aldhous P: Geneticist fears 'race-neutral' studies will fail ethnic groups. *Nature* 2002, 418. www.nature.com/nature

Dukepoo FC: It's More Than the Human Genome Diversity Project. The Human Genome Diversity Project. Politics and the Life Sciences September 1999

Duster T: Race and Reification in Science. *Science* 2005, 307. www.sciencemag.org

Eales, L-J, K.E. Nye, and A.J. Pinching (1988). "Group-Specific Component and AIDS: Erroneous Data." *Lancet* 1:936.

Feldman MW, Lewontin RC, King MC: A genetic melting-pot. *Nature* 2003, 424. www.nature.com/nature

Green, R.L., R.G. Huntsman, and G.R. Serjeant (1971). "The Sickle-Cell and Altitude." *British Medical Journal* 4:593-95.

Goldman, D. et al. (1997). "Linkage and Association of a Functional DRD2 Variant [Ser311Cys] and DRD2 Markers to Alcoholism, Substance Abuse and Schizophrenia in Southwestern American Indians." *American Journal of Medical Genetics (Neuropsychiatric Genetics)* 74:386-94.

Goldstein DB, Hirschhorn JN: In genetic control of disease, does 'race' matter? *Nature Genetics* 2004, 36.

Goldstein DB, Weale ME: Population Genomics: linkage disequilibrium holds the key. *Curr. Biol.* 2001, 11. 576-579

Hinds DA, Stuve LL, Nilsen GB, Halperin E, Eskin E, Ballinger DG, Frazer KA, Cox DR: Whole-Genome patterns of Common DNA Variation in Three Human Populations. *Science* 2005, 307.

Human Genome Diversity Project. Morrison Institute for Population and Resource Studies. http://www.stanford.edu/group/morrinst/hgdp/faq.html

Human Genome Project Information. Wednesday, October 27, 2004. http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml

International HapMap Project. http://www.hapmap.org

Jenkins T: Ethics and The Human Genome Diversity Project: An African Perspective. The Human Genome Diversity Project. Politics and the Life Sciences September 1999

Konotey-Ahulu, F.I.D. (1987). "Group Specific Component and HIV Infection." *Lancet* 1:1267.

Konotey-Ahulu FID: The Human Genome Diversity Project: Cogitations of an African Native. The Human Genome Diversity Project. Politics and the Life Sciences September 1999

NIH Conference. "Involving Diverse Communities in Genetic Research." Natcher Conference Center, National Institutes of Health, November 23 1998.

NitroMed, Inc., "BiDil Named to American Heart Association's 2004 'Top 10 Advances' List; Only Cardiovascular Drug Recognized by AHA for Dramatically Improving Survival in African American Heart Failure Patients," PR Newswire US, 11 January 2005.

Risch N, Buchard E, Ziv E, Tang H: Categorization of humans in biomedical research: genes, race and disease. *Genome Biology* 2002, 3:1–2007.12.

Rotimi CN: Are medical and non-medical uses of large-scale genomic markers conflating genetics and 'race'? *Nature Genetics Supplement* 2004, 36.

Serre D, Pääbo S: Evidence for Gradients of Human Genetic Diversity Within and Among Continents. *Genome Research* 2004, 14:1679–1685.

Smedley A, Smedley BD: Race as Biology Is Fiction, Racism as a Social Problem Is Real: Anthropological and Historical Perspectives on the Social Construction of Race. *American Psychologist* 2005, 60:16–26.

Smedley, B. D., Stith, A. Y., & Nelson, A. R. (Eds.). (2003). Unequal treatment: Confronting racial and ethnic disparities in health care (Committee on Understanding and Eliminating Racial and Ethnic Disparities in Health Care, Institute of Medicine). Washington, DC: National Academies Press.

SNPs: Variations on a Theme. March 29, 2004. http://www.ncbi.nlm.nih.gov/About/primer/snps.html

Taylor AL, Cohn JN, Worcel M, Franciosa JA & A-HeFT Investigators. "The African-American Heart Failure Trial: background, rationale and significance." J. Natl. Med. Assoc. 94, 762-769. 2002.

Wadman M ,Washington: Geneticists struggle towards consensus on place for 'race'. *Nature* 2004, 431. www.nature.com/nature

Wilson JF, Weale ME, Smith, Gratrix F, Fletcher B, Thomas MG, Bradman N, Goldstein DB: Population genetic structure of variable drug response. *Nature Genetics* 2001, 29:265-269

Witzig R The medicalization of race: scientific legalization of a flawed social construct. *Ann. Intern. Med.* 125, 675-679. (1996)