NIH Public Access
Author Manuscript
Hum Genet. Author manuscript; available in PMC 2011 March 1.

Published in final edited form as:
Hum Genet. 2010 March ; 127(5): 563–572. doi:10.1007/s00439-010-0800-0.

Population description and its role in the interpretation of genetic association

Stephanie M. Fullerton,
Department of Bioethics and Humanities, Center for Genomics and Healthcare Equality, University of Washington School of Medicine, Box 357120, Seattle, WA 98195, USA

Joon-Ho Yu,
Department of Bioethics and Humanities, Center for Genomics and Healthcare Equality, University of Washington School of Medicine, Box 357120, Seattle, WA 98195, USA

Julia Crouch,
Department of Bioethics and Humanities, Center for Genomics and Healthcare Equality, University of Washington School of Medicine, Box 357120, Seattle, WA 98195, USA

Kelly Fryer-Edwards, and
Department of Bioethics and Humanities, Center for Genomics and Healthcare Equality, University of Washington School of Medicine, Box 357120, Seattle, WA 98195, USA

Wylie Burke
Department of Bioethics and Humanities, Center for Genomics and Healthcare Equality, University of Washington School of Medicine, Box 357120, Seattle, WA 98195, USA

Stephanie M. Fullerton: smfllrtn@u.washington.edu

Abstract

Despite calls for greater clarity and precision of population description, studies have documented persistent ambiguity in the use of race/ethnicity terms in genetic research. It is unclear why investigators tolerate such ambiguity, or what effect these practices have on the evaluation of reported associations. To explore the way that population description is used to replicate and/or extend previously reported genetic observations, we examined articles describing the association of the peroxisome proliferator-activated receptor-gamma-γ Pro12Ala polymorphism with type 2 diabetes mellitus and related phenotypes, published between 1997 and 2005. The 80 articles identified were subjected to a detailed content analysis to determine (1) how sampled populations were described, (2) whether and how the choice of sample was explained, and (3) how the allele frequency and genetic association findings identified were contextualized and interpreted. In common with previous reports, we observed a variety of sample descriptions and little explanation for the choice of population investigated. Samples of European origin were typically described with greater specificity than samples of other origin. However, findings from European samples were nearly always compared to samples described as “Caucasian” and sometimes generalized to all Caucasians or to all humans. These findings suggest that care with population description, while important, may not fully address analytical concerns regarding the interpretation of variable study outcomes or ethical concerns regarding the attribution of genetic observations to broad social groups. Instead, criteria which help investigators better distinguish justified and unjustified forms of population generalization may be required.

Correspondence to: Stephanie M. Fullerton, smfllrtn@u.washington.edu.

Electronic supplementary material The online version of this article (doi:10.1007/s00439-010-0800-0) contains supplementary material, which is available to authorized users.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Introduction

The use of racial and/or ethnic group descriptors in population-based genetic research has received much recent attention, in both the bioethics and broader scientific literature. Critique has focused on the inconsistent description of population samples by race/ethnicity and the inadequate explanation of how or why research participants are assigned to specific racial or ethnic categories for analysis (Sankar and Cho 2002; Race, Ethnicity, and Genetics Working Group 2005; Shields et al. 2005; Lee et al. 2008; Caulfield et al. 2009). The concern with such practices is that ambiguity of population description can lead to unwarranted generalization of findings to broad social categories (with potential to promote racial stereotyping and stigmatization) (Sankar 2006; Fujimura et al. 2008; Lee 2009) as well as hinder scientific progress by interfering with the assessment of confounding due to population stratification or the replication of population-specific observations (Anonymous 2000; Kaplan and Bennett 2003; Anonymous 2004; Comstock et al. 2004).

Two studies have investigated the use of race/ethnicity terms in the genetic epidemiology literature, documenting the degree to which ambiguity persists despite calls for greater clarity and precision in population analysis and labeling (Shanawani et al. 2006; Sankar et al. 2007). These studies suggest, for example, that only 9% (Sankar et al. 2007) and 28% (Shanawani et al. 2006) of articles explain their methods for assigning race and/or ethnicity to study samples. However, these studies have not explained why so many scientific investigators and journal editors apparently tolerate imprecise and ill-defined population description in reports of genetic research. Interviews with genetic scientists (Smart et al. 2008; Hunt and Megyesi 2008a, b) and/or editors of genetics journals (Outram and Ellison 2006; Smart et al. 2006) are beginning to fill in this gap. Although genetic scientists tend to acknowledge that race/ethnicity categories are imprecise and inconsistent, many report that the use of these proxies nevertheless adds value to a study (Smart et al. 2008; Hunt and Megyesi 2008b). Others suggest that scholarly consideration of the nature of population description is more appropriately “outsourced” to other disciplines, such as anthropology, so that geneticists can remain “free to continue their work without having to concern themselves unduly with classificatory problems outside of their control” (Outram and Ellison 2006).

An alternative explanation is that population description remains imprecise because scientific interpretation relies on under explored, and largely implicit, understandings of population biology (Smart et al. 2006; Hunt and Megyesi 2008a; Foster 2009). When geneticists use population description in studies that replicate or extend a previously reported genetic association, for example, they do so with the aim of confirming a putative association. Early replication attempts ideally involve independent investigators repeating the initially reported analysis in a new sample from the same underlying population, with the expectation of observing the same direction and magnitude of association (Chanock et al. 2007; Ioannidis et al. 2008). The next step is then to extend the association to other populations, understood to differ from the first report or initial replication population (Chanock et al. 2007; McCarthy et al. 2008; Igl et al. 2009). For this reason the NCI-NHGRI Working Group on Replication in Association Studies explicitly recommends “description of replication samples, including source, ascertainment and comparability to initial sample” such that when replication is attempted “a similar population should be studied, and notable differences between the populations studied in the initial and attempted replication studies should be described” (Chanock et al. 2007, p. 658, Box 3). However, what exactly constitutes a “similar” or “different” population is less clear and left to individual interpretation or implicit consensus.

Prior empirical investigations of population description in the genetic literature have not distinguished between the publication practices of different genetic sub-disciplines or

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

attempted to examine the effect of research context (i.e. primary observation, confirmatory study, etc.) on the character or consistency of population description. To explore more carefully the way that population description is used to validate and/or extend previously reported genetic observations, we examined a collection of articles describing the association of the peroxisome proliferator-activated receptor-gamma (*PPARγ*) Pro12Ala polymorphism with type 2 diabetes mellitus (T2DM) and related phenotypes, published between 1997 and 2005. The Pro12Ala polymorphism was first identified in a screen of mutations among a described sample of Caucasian diabetics from Baltimore, MD and shown to be present at varying frequency in other human populations (Yen et al. 1997). Although initial reports of the association with diabetes were inconsistent and controversial, a major study reported in 2000 convincingly demonstrated that the Ala variant was, indeed, inversely associated with T2DM risk in populations of Scandinavian origin (Altshuler et al. 2000). Most reports published since then have attempted to confirm and extend the observation to other populations.

We reviewed the *PPARγ* literature to examine how population samples employed in the validation of a previously reported genetic association are described, justified, and interpreted.

Methods

Sample selection

The peer-reviewed articles chosen for analysis were identified as part of an ongoing HuGE (Human Genome Epidemiology) review of variation in the *PPARγ* gene and associations with T2DM, obesity, and related phenotypes. The initial review that was conducted according to established review guidelines (http://www.medicine.uottawa.ca/public-health-genomics/web/assets/documents/HuGE_Review_Handbook_V1_0.pdf) identified 111 articles published between 1997 and the summer of 2005 which explored association with variants in and around the *PPARγ* gene.

We focused on the 80 articles identified in that review that reported associations with the *PPARγ* Pro12Ala variant specifically, i.e. we excluded from further consideration any article in which an analysis of that polymorphism was not reported. Association with the Pro12Ala polymorphism was the only criteria for including an article in our analysis and no attempt was made to further differentiate articles according to specific health outcomes investigated. Although a wide range of phenotypes were considered, because all reports involved investigation of the effect of a specific single nucleotide polymorphism on traits related to metabolic outcomes, we regard each as an attempt to replicate and/or extend the originally reported association of the polymorphism with body mass index, insulin sensitivity, and T2DM (Deeb et al. 1998). Sample selection was thus based on a defined variant–disease association rather than on the use of racial, ethnic, or population descriptors, dates of publication, or journal impact factors, as in earlier studies (Shanawani et al. 2006; Sankar et al. 2007). In this way, we identified a sample of articles that we believe better reflects “typical” epidemiological and publication practices, at least as employed in a specific disease context.

A list of the 80 articles surveyed is provided here as Supplementary Material.

Data analysis

The 80 articles identified were subjected to a detailed content analysis (Weber 1990; Hsieh and Shannon 2005) in an effort to determine (1) how populations sampled for the purposes of replicating *PPARγ* Pro12Ala associations were described, (2) whether and how the choice of replication sample was explained, and (3) how the allele frequency and genetic association findings identified were contextualized and interpreted. All articles were imported into the software package Atlas.ti (version 5.0) for content analysis and coding.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Content codes were developed to capture each of the above-outlined elements. We coded text that described how the research question was described and how it was justified as important or of interest. In particular, we were interested in whether authors explicitly referred to the nature of the population sample in the framing of their research question. We also captured the rationale, when given, behind authors' use of specific populations in the study design, with the goal of understanding assumptions about the analytical desirability of different types of samples, including references to presumed genetic homogeneity or expected differences with respect to previously investigated samples. We coded any reference to other populations or groups based on the race/ethnicity that were compared to a study's population sample, either with respect to similarities or differences in reported allele frequencies or in terms of observed gene–disease associations. Any attempt to generalize findings beyond the immediate study was also coded, capturing the particular relevance of specific comparison populations, where noted.

Two coders (JY and JC) coded articles in parallel; inter-coder differences were reviewed by other members of the research team (KFE and SMF) and all discrepancies reconciled. Descriptive statistics were used to summarize publication information about our sample, including year of publication, journal impact factor (as reported by Thomas Reuters Journal Citation Reports®) and citations (as reported by the Science Citation Index, accessed via the ISI Web of Science® search engine), and country of origin of first author. Coded data were tabulated to assess the frequency with which specific codes were mentioned and how they were characterized. Population descriptions, which were drawn verbatim from the articles, were classified as representing a race/ethnicity, national, or geographic attribution from the context of their presentation, e.g. the Oji-Cree were described in one article as “self-defined as Oji-Cree on the basis of their language” and classed as an ethnic group (rather than nationality) on that basis. We tabulated different classes of justifications for choice of sample populations and calculated the frequency and type of generalizations (made beyond the immediate study sample).

Results

Articles focused on *PPARγ* Pro12Ala association

Reports of the association of *PPARγ* Pro12Ala with T2DM and related phenotypes increased steadily after the association was first reported by Yen et al. (1997), and then again after the definitive validation of the association by Altshuler et al. (2000) 3 years later. Our survey identified approximately 10–12 articles per year from 2000 onwards, published in a range of journals, many of which focus specifically on diabetes, obesity, endocrinology, and metabolism. The average impact factor at the time of publication was 4.7 and the average number of citations 1 year after completion of the initial review (in August 2006) was 37, although a full third of the articles we examined had been cited fewer than 10 times. These findings are consistent with a body of articles that mostly represents independent attempts at validating a previously reported genetic association. The majority of the studies were conducted by first authors residing in Europe and East Asia; 14 of the 80 studies were conducted by first authors in the United States.

Nature of population description

A total of 101 different population samples were described in the 80 *PPARγ* Pro12Ala reports. Ten of the 80 articles described the analysis of more than one population, with the number of samples investigated ranging from 2 to 9 (average 3.2). Ten of the 101 samples (10%), in 10 reports, were not described with respect to geographic, national, racial/ethnic, or ancestral origin. Instead, readers were directed to additional citations for sample descriptions. In these instances, there was no way to determine, from the immediate published report, in what population(s) the association had been investigated.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

The remaining 91 population samples were described either by race or ethnicity (47% of samples), nationality (33%), via a combined nationality and race/ethnicity description (18%), or by the geographic region of sampling (2%) (Table 1). Three of the populations sampled were described by "ancestry" although no explicit attempt to estimate genetic ancestry was reported. From context, we interpreted these latter descriptions as signaling the racial or ethnic background of study participants rather than nationality. For example, for a sample recruited from a US-based medical center, the authors stated that "All subjects were of Northern European ancestry" with no other reference to the demographic characteristics of the research subjects.

In Tables 1, 2, and 3, we organize our summary of the sampled populations by the biogeographical origin of the referenced populations using columns headed "Origin of Reported Sample." Origin was inferred from the reported geographic location of the population sampled or, for the case of US-based racial/ethnic groups, from accepted understanding of historical population movements. We chose this approach in an effort to examine what effect, if any, the (broadly construed) 'ancestral' background of a study sample might have on reporting and/or interpretative practices. The majority of the population samples employed in these studies of *PPARγ* Pro12Ala variant were of European origin (56 of 91 or 62%), whereas 25% were of Asian, 5% were of African, and 4% were of Native American origin (Table 1). Three other population samples (described as Mexican-American, Arabian, and Israeli) could not be classified on the above basis and hence are marked "Other" in our tables.

Samples of European origin were described using a much broader array of population descriptors than samples of other biogeographic origin (Table 1). The only populations described in terms of geographic location were of European origin, and all 16 of the samples described in terms of combined nationality and race were of European origin; all were identified as being of the Caucasian race. Similarly, nearly two-thirds of the samples described by nationality were European, with 11 nations named, whereas only 3 non-European nations (Japan, Korea and Israel) were used for sample description in the remaining cases. In contrast, the non-European origin samples were nearly always described with regard to racial and/or ethnic identity. For Asian origin populations, sample descriptions based in racial/ethnic identity were distinguished from descriptions of nationality by the context of their use (e.g. "Chinese" could be interpreted as a description of nationality, but was used in two articles to describe the ethnic background of study participants). Unlike with the European origin populations, no Asian samples of specific national origin were further qualified by a racial designation. In addition, unlike a large number of European origin samples described racially as "Caucasian", no Asian origin samples were described using such broad racial identifiers. There was no relationship between the nature of population description and the year or journal of publication.

Justification for choice of study populations

Although a majority of population samples were described with regard to race/ethnicity, national, or geographic origin, only 9 of the 80 *PPARγ* Pro12Ala reports (11%) included any explanation of, or justification for, the choice of study population. Therefore, it was impossible to determine if associations were being investigated in an effort to replicate the initially reported association, extend observations to novel population settings, utilize relevant and readily available case samples, or for other reasons. Where justification was provided for the investigation of the Pro12Ala association in a specific population sample, the explanation came in one of two main forms: either (1) the population sampled was described as having special demographic properties which implied that it was especially appropriate for investigation of genetic association, or (2) the sample investigated was described as likely to differ from previously investigated samples with regard to the Pro12Ala association (Table 2).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

5 of the 80 articles surveyed included explanations which invoked special population properties: two focused on samples of Asian origin, two focused on samples of Native American origin, and one focused on a sample of European origin. In four of these cases, the particular population property highlighted was a presumed genetic homogeneity (e.g. Japanese, Danish Caucasians), or low genetic heterogeneity due to population isolation and/or genetic founder effects (e.g. Canadian aboriginals, Parkataje Indians). However, direct evidence of population homogeneity and/or lack of confounding due to population stratification, was not presented. In another example, a study comparing "Native Japanese" and "Japanese-Americans" suggested that the comparison was appropriate because the samples were likely to be genetically similar, but differ with regard to environmental exposures, providing an opportunity for investigation of gene–environment interaction.

Four other studies justified their population choice based on expected population differences: two focused on samples of Asian origin, one on an African-American sample, and one involved a comparison of Arabian and Scandinavian samples. In three of the four cases, the chosen population samples were described as likely to differ genetically from previously studied Caucasian or European origin populations. A fourth, African-American, sample was justified by the expectation that ethnicity-related genetic background differences might alter the phenotypic expression of the candidate variant. No European or Native American population samples were singled-out on such grounds.

Interpreting observations: allele frequency comparisons

Genetic association studies focused on replication must relate current association findings to previous reports and situate new observations in population context. Two major kinds of observations were reported and discussed in the 80 articles describing associations with the *PPARγ* Pro12Ala variant: allele frequency observations and gene–disease associations. Allele frequency comparisons generally involved reporting the observed frequency of the Pro12Ala variant in the population sample under study and then relating the observed frequency to allele frequencies reported in populations sampled previously and such allele frequency comparisons were included in 26 (32%) of the 80 articles surveyed here (Table 3).

Allele frequency comparisons emphasizing similarities between the reported population sample and population samples described in previous reports were twice as common as comparisons emphasizing differences (Table 3). Samples of European origin were most often included in comparisons emphasizing allele frequency similarities and then nearly always compared to samples described as being of Caucasian race/ethnicity. Similarly, when allele frequency similarities were noted for samples of Asian origin, the population samples noted for comparison were typically also of Asian origin. The one exception in this regard involved a report in which both allele frequency similarities and differences were noted and the reported frequency was compared to the distribution of frequencies initially reported by Yen et al. 1997 (marked as footnote "b" in Table 3). Only one other non-Asian and non-European population sample (a sample described as Mexican-American) had a reported allele frequency of Pro12Ala said to be similar to that reported previously for a Caucasian sample.

Allele frequency comparisons emphasizing differences between the reported population sample and population samples described in previous reports were, in contrast, less common and most often involved populations of Asian origin (Table 3). In these cases, the referent population identified for comparison was nearly always of Caucasian race/ethnicity. One report of a presumed European origin population (the sample was not described, but the study was conducted in Finland) noted differences in allele frequency compared to that reported for a Japanese sample. The emphasis in these instances on Asian-European allele frequency differences is notable because the absolute difference in reported allele frequencies is rather small: approximately 10% in samples of European origin versus 2–5% in samples of Asian

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

origin (rs1801282, dbSNP). A low minor allele frequency is, nevertheless, a concern when samples sizes are modest and power to detect association is reduced.

Interpreting observations: generalization of reported associations

In addition to allele frequency comparisons, the other main type of observation reported and discussed in the 80 articles was the association of the *PPARγ* Pro12Ala polymorphism with T2DM, obesity, and related phenotypes. While discussion of an observed association (or failure to observe an expected association) did sometimes involve making comparisons to previous reports in a manner analogous to the allele frequency comparisons, the majority of reports (61 of 80, 76%) made no explicit attempt to generalize observations beyond the immediate study sample. Thus, most authors appeared content to note (or alternatively, question) the previously reported Pro12Ala association, but not to infer from their findings that the genotype–phenotype association was confirmed or disconfirmed. Indeed a number of authors were explicit about not generalizing their study observations.

In the remaining 19 reports, sample-specific observations were generalized beyond the immediate study, to a broader population unit which, the articles implied, encompassed the studied population (Table 4). In two cases, observations were generalized to a nation or ethnicity, in 5 cases to a race, and in 12 cases to all of humankind. In only four instances did reported associations involve multiple population samples such that generalization to a broader population might represent a defensible interpretation of the available data. In one of the latter cases, a study of Native Japanese and Japanese-Americans, findings were generalized to individuals of Japanese ethnicity; in the other three instances results were inferred to apply to all humans, including a joint study of Scandinavians and French-Canadians. When generalizations were made to a race, all involved generalizing from a single European study population to the broader racial category of “Caucasian.” In other instances, associations reported for a single European study sample were generalized to all humans. In at least three cases, broad interpretative statements were made which implied that results were believed to apply to all humans, even though this was not stated explicitly. We regarded statements referencing the “general population” to mean generalization to all humans, though it also possible that authors meant the general population from which the study sample had been drawn.

Discussion

There has been intense interest over the past several years in population description in genetic research and, especially, the attribution of genetic observations to groups defined with racial or ethnic group labels (Sankar and Cho 2002; Race, Ethnicity, and Genetics Working Group 2005; Shields et al. 2005; Lee et al. 2008). Our study examined population description in a set of 80 research articles focused on the association of a single polymorphism, the Pro12Ala variant of *PPARγ* gene, with a range of diabetes-related metabolic outcomes. In common with previous investigations of genetic research publications (Shanawani et al. 2006; Sankar et al. 2007), we observed a wide variety of population sample descriptions, ranging from the use of geographic descriptors, to national labels, to racial group designations. In addition, in common with previous reports, the assignment of population label was rarely explained nor was the sample chosen for investigation typically justified in methodological terms (Hunt and Megyesi 2008a; Lee 2009). Our observations thus confirm that calls for greater care with respect to population description in genetic research continue to go unheeded in many cases (Sankar et al. 2007; Hunt and Megyesi 2008a, b).

Our study differs from previous analyses of reporting practices in its explicit focus on articles chosen for their consideration of a defined genetic polymorphism, rather than publications which use “race” and/or “ethnicity” as keywords (Shanawani et al. 2006; Lee 2009), or articles

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

published in specific journals (Sankar et al. 2007). We were thus able to examine descriptive practices as they arise in the context of a particular class of genetic investigation (i.e. the replication of a previously reported association) and without regard to the specifics of the population(s) under consideration or the nature of publication venue. In addition to differences in sample selection, our analysis focused on the relationship of population description to result interpretation rather than on methods for assigning race/ethnicity to study participants (Shanawani et al. 2006; Sankar et al. 2007) or the ways that race/ethnicity terms are used in the primary report of a genetic observation, e.g. as a dependent or independent variable (Kaplan and Bennett 2003; Sankar et al. 2007). This difference in emphasis allowed us to examine how results generated in the context of variously defined samples are situated relative to those of other studies or samples, and hence how such results appear to be understood scientifically (Foster 2009), despite the ambiguity of sample assignment or definition.

Most of the studies in our sample were conducted in populations of European origin (i.e. participants sampled from a population currently living in Europe or with an inferred recent ancestral origin within such a population), which may reflect the fact that the association with T2DM was validated in a sample of Scandinavian origin (Altshuler et al. 2000) or could instead be related to the fact that the Pro12Ala variant tends to be found at higher frequency in populations of Caucasian race/ethnicity or European origin (Yen et al. 1997). Only one European sample was described as genetically homogeneous, suggesting that concerns about population stratification were not a major consideration in sample choice. Alternatively, the larger number of European replication attempts could reflect the national and/or geographic origins of the scientists conducting the association studies. Populations of European origin tended to be described using a wider variety of terms, and generally more specifically, than samples of other biogeographic origin.

European samples also served more often as the sample of reference in allele frequency comparisons. In such comparisons, the polymorphism frequency observed in European samples was more likely to be described as similar to that reported for other samples of European origin or samples described as “Caucasian,” whereas non-European (mostly of Asian origin) sample allele frequencies were described as different from those of Europeans or Caucasians. Although reports of genetic association were only infrequently generalized beyond the specific named population sample, the generalizations we observed nearly always involved a population of European origin and generalization to the category “Caucasian” or to all humans. Therefore, in this study even population samples that were specifically characterized were sometimes, in their interpretation, treated as broader (typically racially described) units for purposes of comparison or generalization.

Attributing sample-specific genetic observations to broader race/ethnicity groups, either via an inappropriately broad description of a sampled population, or via generalization of findings to additional, non-sampled, populations, has been discouraged by bioethicists and social scientists concerned about the use of genetic observations to stereotype or stigmatize social groups (Sankar and Cho 2002; Race, Ethnicity, and Genetics Working Group 2005; Shields et al. 2005; Lee et al. 2008; Caulfield et al. 2009). Yet situating an individual study's findings in broader population context is an important component of epidemiological practice, and is especially relevant when investigation involves confirming or extending previously reported genetic associations (Little et al. 2009). Unless an identical population is under investigation, some degree of generalization is inevitable in the course of result interpretation. From this standpoint, calls to avoid unwarranted generalization may be misunderstood or, in the extreme, viewed by investigators as a suggestion to avoid generalization altogether—a response many would find scientifically counterproductive and undesirable. Instead, criteria which might help investigators better distinguish among justified and unjustified forms of generalization could

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

address ethical concerns while simultaneously promoting practices better suited to the robust confirmation of preliminary findings.

Similarly, greater ethical and analytical attention should be paid to the ways in which sample choice is justified in the confirmation of previously reported genetic associations. Recommendations suggest initial replication should be attempted in a "similar" population, followed by extension of the observation to other populations (Chanock et al. 2007; McCarthy et al. 2008; Igl et al. 2009). Once an association has been confirmed in a comparable sample (whether by analysis of separate cohorts of a single larger sample or by the independent recruitment of participants from an equivalent population) the question then becomes what kind of population samples are "too similar" or "sufficiently different" for follow-on investigation. Samples understood as too similar represent wasteful additional attempts at replication that should be avoided, whereas those deemed sufficiently different may be prioritized to extend the association to new contexts. However, it is not clear how often such considerations actually inform study design. Although the data presented here suggest that population similarity and difference were understood primarily in biogeographic and/or racial terms, the majority of the studies we identified replicated (arguably unnecessarily) the *PPARγ* association in a sample of European origin. Moreover, only one of those reports included a specific explanation for the choice of a European sample (Table 2).

These observations all serve to underline that while careful attention to population description is important in the dissemination of genetic research findings, it may ultimately be insufficient for addressing persistent ethical concerns surrounding the attribution of genetic observations to broad social groups (Smart et al. 2006). Not only might specificity of population description have relatively little to do with the way in which findings are generalized, but a focus on explanations for the choice of population labels may inadvertently cause us to overlook equally compelling questions about the choice of population samples and their broader research justification. Genetic scientists and their funders may be better served by addressing downstream research ethics concerns at earlier stages of study design and/or prioritization, and to do so in collaboration with representatives of the communities they hope to investigate.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Karen Edwards PhD, and members of the University of Washington Center for Genomics and Public Health, for providing access to the in-progress HuGE review of the *PPARγ* Pro12Ala association with type 2 diabetes and related phenotypes upon which our analysis is based. The authors also acknowledge Lizzeth Belskus-Amador, Aniel Solis, and Tessa Steel for research assistance. The article benefitted from the critical review of Melissa Austin, PhD, as well as several anonymous reviewers. This work was supported by the National Institutes of Health [P50 HG003374 to W.B.] and the Centers for Disease Control [U36 CCU300430-23 and U10 CCU025038 to K. Edwards].

References

Altshuler D, Hirschhorn JN, Klannemark M, Lindgren CM, Vohl MC, Nemesh J, Lane CR, Schaffner SF, Bolk S, Brewer C, Tuomi T, Gaudet D, Hudson TJ, Daly M, Groop L, Lander ES. The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. Nat Genet 2000;26(1):76–80. [PubMed: 10973253]

Census, race and science. Nat Genet 2000;24(2):97–98. [PubMed: 10655044]

The unexamined 'Caucasian'. Nat Genet 2004;36(6):541. [PubMed: 15167921]

Caulfield T, Fullerton SM, Ali-Khan SE, Arbour L, Burchard EG, Cooper RS, Hardy BJ, Harry S, Hyde-Lay R, Kahn J, Kittles R, Koenig BA, Lee SS, Malinowski M, Ravitsky V, Sankar P, Scherer SW,

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Seguin B, Shickle D, Suarez-Kurtz G, Daar AS. Race and ancestry in biomedical research: exploring the challenges. Genome Med 2009;1(1):8. [PubMed: 19348695]

Chanock SJ, Manolio T, Boehnke M, Boerwinkle E, Hunter DJ, Thomas G, Hirschhorn JN, Abecasis G, Altshuler D, Bailey-Wilson JE, Brooks LD, Cardon LR, Daly M, Donnelly P, Fraumeni JF Jr, Freimer NB, Gerhard DS, Gunter C, Guttmacher AE, Guyer MS, Harris EL, Hoh J, Hoover R, Kong CA, Merikangas KR, Morton CC, Palmer LJ, Phimister EG, Rice JP, Roberts J, Rotimi C, Tucker MA, Vogan KJ, Wacholder S, Wijsman EM, Winn DM, Collins FS. Replicating genotype–phenotype associations. Nature 2007;447(7145):655–660. [PubMed: 17554299]

Comstock RD, Castillo EM, Lindsay SP. Four-year review of the use of race and ethnicity in epidemiologic and public health research. Am J Epidemiol 2004;159(6):611–619. [PubMed: 15003966]

Deeb SS, Fajas L, Nemoto M, Pihlajamäki J, Mykkänen L, Kuusisto J, Laakso M, Fujimoto W, Auwerx J. A Pro12Ala substitution in PPARgamma2 associated with decreased receptor activity, lower body mass index and improved insulin sensitivity. Nat Genet 1998;20(3):284–287. [PubMed: 9806549]

Foster MW. Looking for race in all the wrong places: analyzing the lack of productivity in the ongoing debate about race and genetics. Hum Genet 2009;126(3):355–362. [PubMed: 19396464]

Fujimura JH, Duster T, Rajagopalan R. Race, genetics, and disease: questions of evidence, matters of consequence. Soc Stud Sci 2008;38(5):643–656. [PubMed: 19227816]

Hsieh HF, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005;15(9): 1277–1288. [PubMed: 16204405]

Hunt LM, Megyesi MS. The ambiguous meanings of the racial/ethnic categories routinely used in human genetics research. Soc Sci Med 2008a;66:349–361. [PubMed: 17959289]

Hunt LM, Megyesi MS. Genes, race and research ethics: who's minding the store? J Med Ethics 2008b; 34:495–500. [PubMed: 18511627]

Igl BW, Konig IR, Ziegler A. What do we mean by ‘replication’ and ‘validation’ in genome-wide association studies? Hum Hered 2009;67(1):66–68. [PubMed: 18931511]

Ioannidis JP, Boffetta P, Little J, O'Brien TR, Uitterlinden AG, Vineis P, Balding DJ, Chokkalingam A, Dolan SM, Flanders WD, Higgins JP, McCarthy MI, McDermott DH, Page GP, Rebbeck TR, Seminara D, Khoury MJ. Assessment of cumulative evidence on genetic associations: interim guidelines. Int J Epidemiol 2008;37(1):120–132. [PubMed: 17898028]

Kaplan JB, Bennett T. Use of race and ethnicity in biomedical publication. JAMA 2003;289(20):2709–2716. [PubMed: 12771118]

Lee C. “Race” and “ethnicity” in biomedical research: how do scientists construct and explain differences in health? Soc Sci Med 2009;68(6):1183–1190. [PubMed: 19185964]

Lee SS, Mountain J, Koenig B, Altman R, Brown M, Camarillo A, Cavalli-Sforza L, Cho M, Eberhardt J, Feldman M, Ford R, Greely H, King R, Markus H, Satz D, Snipp M, Steele C, Underhill P. The ethics of characterizing difference: guiding principles on using racial categories in human genetics. Genome Biol 2008;9(7):404. (Epub Jul 15). [PubMed: 18638359]

Little J, Higgins JP, Ioannidis JP, Moher D, Gagnon F, von Elm E, Khoury MJ, Cohen B, Davey-Smith G, Grimshaw J, Scheet P, Gwinn M, Williamson RE, Zou GY, Hutchings K, Johnson CY, Tait V, Wiens M, Golding J, van Duijn C, McLaughlin J, Paterson A, Wells G, Fortier I, Freedman M, Zecevic M, King R, Infante-Rivard C, Stewart A, Birkett N. Strengthening the reporting of genetic association studies (STREGA): an extension of the STROBE Statement. Hum Genet 2009;125(2): 131–151. [PubMed: 19184668]

McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JP, Hirschhorn JN. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat Rev Genet 2008;9(5):356–369. [PubMed: 18398418]

Outram SM, Ellison G. Anthropological insights into the use of race/ethnicity to explore genetic contributions to disparities in health. J Biosoc Sci 2006;32:83–102. [PubMed: 16266441]

Race, Ethnicity, and Genetics Working Group. The use of racial, ethnic, and ancestral categories in human genetics research. Am J Hum Genet 2005;77(4):519–532. [PubMed: 16175499]

Sankar P. Hasty generalisation and exaggerated certainties: reporting genetic findings in health disparities research. New Genet Soc 2006;25(3):249–264.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Sankar P, Cho MK. Genetics. Toward a new vocabulary of human genetic variation. Science 2002;298 (5597):1337–1338. [PubMed: 12434037]

Sankar P, Cho MK, Mountain J. Race and ethnicity in genetic research. Am J Med Genet 2007;143A(9): 961–970. [PubMed: 17345638]

Shanawani H, Dame L, Schwartz DA, Cook-Deegan R. Non-reporting and inconsistent reporting of race and ethnicity in articles that claim associations among genotype, outcome, and race or ethnicity. J Med Ethics 2006;32(12):724–728. [PubMed: 17145914]

Shields AE, Fortun M, Hammonds EM, King PA, Lerman C, Rapp R, Sullivan PF. The use of race variables in genetic studies of complex traits and the goal of reducing health disparities: a transdisciplinary perspective. Am Psychol 2005;60(1):77–103. [PubMed: 15641924]

Smart A, Tutton R, Ashcroft R, Martin PA, Ellison GTH. Can science alone improve the measurement and communication of race and ethnicity in genetic research? Exploring the strategies proposed by Nature Genetics. BioSocieties 2006;1(3):313–324.

Smart A, Tutton R, Martin P, Ellison GT, Ashcroft R. The standardization of race and ethnicity in biomedical science editorials and UK biobanks. Soc Stud Sci 2008;38(3):407–423. [PubMed: 19069078]

Weber, RP. Basic content analysis. Sage; Beverly Hills: 1990.

Yen C, Beamer BA, Negri C, Silver K, Brown KA, Yarnall DP, Burns DK, Roth J, Shuldiner AR. Molecular scanning of the human peroxisome proliferator activated receptor g (hPPARγ) gene in diabetic Caucasians: identification of a Pro12Ala PPARγ2 missense mutation. Biochem Biophys Res Commun 1997;241:270–274. [PubMed: 9425261]

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1

Population description of 91 samples used in 80 reports of *PPARγ* Pro12Ala genetic association

Origin of reported sample	Race/ethnicity (43 samples)	Nationality (30 samples)	Nationality-race/ethnicity (16 samples)	Geography (2 samples)
African (5 samples)	African-Americans (3), Black (2)	None	None	None
Asian (23 samples)	Chinese (2), Japanese-Americans (2), Chinese Ancestry (1), Full Japanese Ethnicity (1), Indian (1), Japanese Ancestry (1), Malay (1), Nauruans (1), Native Javanese (1), Native Japanese (1), Samoans (1), Taiwanese of Huanan descent (1)	Japanese (7), Korean (2)	None	None
European (56 samples)	Caucasian (11), White (3), French-Canadian (1), German Caucasian Ethnicity (1), Northern European Ancestry (1), Scandinavian (1)	Finnish (6), Czech (2), French (2), German (2), Spanish (2), Austrian (1), Canadian (1), Danish (1), Italian (1), Norwegian (1), Swedish (1)	Danish Caucasians (3), Italian Caucasians (3), French Caucasians (2), Spanish Caucasians (2), Brazilian Caucasians (1), Caucasians from Western Australia (1), German Caucasians (1), Scottish Caucasians (1), Swedish Caucasians (1), Uruguayan Caucasians (1)	Lazio Region of Italy (1), Scandinavian (1)
Native American (4 samples)	Oji-Cree (2), Parkataje Indians of Brazil (1), Pima Indians (1)	None	None	None
Other (3 samples)	Mexican-Americans (1), Arabian (1)	Israeli (1)	None	None

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 2

Explanations for samples used in nine of 80 published reports (no explanation otherwise provided)

Origin of reported sample	Special properties of the population	Expected population differences
African	n/a	African-Americans assumed to harbor ethnicity-related genetic background differences
Asian	Native Japanese and Japanese-Americans assumed to be genetically similar but with "different environmental factors"	Singapore study populations contrasted to previously studied Caucasian populations
	Japanese a "genetically homogeneous" population	Korean study population contrasted to previously studied Caucasian populations
European	Danish Caucasians a "homogenous" population	n/a
Native American	Canadian aboriginals "an isolated founder population" with low genetic heterogeneity	n/a
	Parkataje Indians were historically "largely isolated" but "recently underwent a rapid and intensive process of acculturation"	
Other	n/a	Prediction of greater insulin resistance in Arabians compared with Scandinavians

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 3

Allele frequency similarities and differences relative to sampled population

Origin of reported sample	Population with reported allele frequency	Comparison population with similar allele frequency
Asian (6 reports)	Japanese[a]	Other Japanese
	Japanese[b]	African-Americans; Nauruans; Chinese
	Japanese ethnicity	Other Japanese
	Native Japanese; Japanese-American	Other Japanese
	Native Javanese	Asian populations
	Korean[c]	East Asians; Chinese Taiwanese; Japanese
European (13 reports)	Caucasian	Other Caucasians
	Italian Caucasian	Other Caucasians
	Spanish Caucasian	Italian Caucasians
	Czech	Central Europeans
	Danish Caucasian	Caucasians
	Finnish	Caucasians
	German Caucasian	Caucasians
	Italian Caucasian	Caucasians; Japanese-Americans
	Lazio region of Italy	Caucasians
	Northern European ancestry	Unrelated Caucasians
	Not specified (conducted in Scotland)	Whites
	Spanish nationality	Caucasians
	White	Other studies of white samples
Other	Mexican-American	Caucasian population
Origin of reported sample	Population with reported allele frequency	Comparison population with different allele frequency
African	African-Americans	Caucasians
Asian (6 reports)	Chinese ancestry; Japanese ancestry	Other ethnic backgrounds
	Japanese[a]	Americans; Europeans
	Japanese[b]	Caucasians; Mexican-Americans
	Korean	Caucasians
	Korean[c]	Caucasians
	Chinese, Malay, Indian Singaporeans	Caucasians
European	Not specified (conducted in Finland)	Japanese
Other	Parkataje Indians of Brazil	Chinese; Caucasians

[a,b,c] Duplicates where both similarities and differences were reported

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Table 4

Generalization of sample-specific association findings

Level of generalization	Specific population	Broader grouping
To ethnicity or nation (2 reports)	German origin	German population
	Native Japanese and Japanese-Americans	Japanese
To race (5 reports)	French Caucasians	Caucasian
	French	Caucasian
	German Caucasian ethnicity	Caucasian
	German population	Caucasian
	Swedish	Caucasian
To all humans (12 reports)	Black and White biracial community (from US)	General population
	Finnish and Japanese-American	General population; humans
	Japanese	General population
	Spanish Caucasians	Diabetic women
	Italian Caucasians	General population
	Czech	Humans (implied)[a]
	Caucasian	Humans (implied)[a]
	Northern European ancestry	Patients
	Scandinavian, French-Canadian	Human population
	Mexican-American	Some populations
	Not specified (residents of Southern Poland)	Humans (implied)[a]
	Not specified (Nurses Health Study)	Humans

[a] Authors describe their conclusions in general terms without referencing a specific population or group, thus implying that their conclusions apply to all humans. For example, “the data presented here support an inverse association between 12Ala *PPARγ* allele and type 2 diabetes”