Original Paper

Public Health Genomics

Public Health Genomics 2010;13:368–377
DOI: 10.1159/000276767

Received: May 15, 2009
Accepted after revision: August 4, 2009
Published online: January 15, 2010

Public and Biobank Participant Attitudes toward Genetic Research Participation and Data Sharing

A.A. Lemke[a] W.A. Wolf[a] J. Hebert-Beirne[b] M.E. Smith[a]

[a]Northwestern University and [b]University of Illinois at Chicago, Chicago, Ill., USA

Key Words
Biobank · Biorepository · Data sharing · Ethics · Focus groups · Genetic research · Genome-wide association studies · Policy · Public attitudes · Qualitative research · Research participation

Abstract
Research assessing attitudes toward consent processes for high-throughput genomic-wide technologies and widespread sharing of data is limited. In order to develop a better understanding of stakeholder views toward these issues, this cross-sectional study assessed public and biorepository participant attitudes toward research participation and sharing of genetic research data. Forty-nine individuals participated in 6 focus groups; 28 in 3 public focus groups and 21 in 3 NUgene biorepository participant focus groups. In the public focus groups, 75% of participants were women, 75% had some college education or more, 46% were African-American and 29% were Hispanic. In the NUgene focus groups, 67% of participants were women, 95% had some college education or more, and the majority (76%) of participants was Caucasian. Five major themes were identified in the focus group data: (a) a wide spectrum of understanding of genetic research; (b) pros and cons of participation in genetic research; (c) influence of credibility and trust of the research institution; (d) concerns about sharing genetic research data and need for transparency in the Policy for Sharing of Data in National Institutes of Health-Supported or Conducted Genome-Wide Association Studies; (e) a need for more information and education about genetic research. In order to increase public understanding and address potential concerns about genetic research, future efforts should be aimed at involving the public in genetic research policy development and in identifying or developing appropriate educational strategies to meet the public's needs.

In parallel with the emergence of new technologies such as genome-wide association studies (GWAS), institutions worldwide are developing biorepositories that house large numbers of participants' genomic DNA and other health data [1–3]. Biorepositories are typically established to support future biomedical research and will increasingly be 'tapped' to support genome-wide studies. Studies of this magnitude will intensify the need to pool samples and data to identify statistically significant associations between gene variants and disease [4, 5]. To facilitate investigator access to and sharing of genotype-phenotype datasets from GWAS, the U.S. National Institutes of Health (NIH) has developed a centralized data repository – the database of Genotypes and Phenotypes (dbGaP) [6]. In addition, the NIH has a Policy for Sharing of Data in NIH-Supported or Conducted GWAS, which outlines data sharing procedures, data access principles, and research participant protections for submitting investigators [7].

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

Accessible online at:
www.karger.com/phg

Amy A. Lemke, MS, PhD
Center for Genetic Medicine, Northwestern University
676 N. St. Clair Street, Suite 1260
Chicago, IL 60611 (USA)
Tel. +1 608 441 6848, Fax +1 608 310 4340, E-Mail a-lemke@northwestern.edu

Although there are considerable scientific and potential public health benefits from collecting and sharing genetic data from GWAS and large scale studies, a number of concerns and potential participant risks have been raised. For example, concerns have been identified regarding confidentiality and privacy of participant research data, stigma and fear of genetic discrimination resulting from disease risk identification, and potential for non-research uses of data [8]. While recent research efforts have attempted to explore these issues, much work remains in understanding participant views toward data sharing. To address this gap, this paper will present findings of a focus group study and provide recommendations for policy development and future research.

Community engagement or consultation has been used to articulate community concerns and explore community-specific benefits regarding the complex ethical, legal and social issues involved in collecting and storing genetic research data [9, 10]. This process can enhance research quality, improve protection of participants, foster community trust, and address local public health needs [11]. A number of strategies have been utilized to explore community concerns and improve data repository/participant relationships while developing large-scale genetic databases in various countries and Canadian provinces including Iceland, Estonia, United Kingdom, British Columbia and Quebec [2, 12–14]. A multitude of strategies were utilized at these sites to engage or consult with communities including education sessions, interviews, surveys, focus groups, deliberative democracy and public discussion meetings. These types of approaches and future community consultation efforts will lead to better understanding of the public's needs and concerns regarding participation in GWAS and large scale genetic research studies.

Although a number of community consultation efforts and other studies have assessed public understanding of the collection and use of genetic research data [2, 13, 15–17], few have assessed issues specific to data sharing from the participant perspective. One such study by McGuire et al. (2008) assessed participant views on DNA data sharing and various levels of control over decision-making [18]. Focus group findings revealed that participants were interested in receiving information about data sharing and preferred multiple data sharing options. Study participants also indicated that genetic research data should not be released without explicit consent [18].

To address a lack of research about genetic research data sharing, a three-phase community consultation effort was designed to gather various stakeholder views of the consent process for high-throughput genomic-wide technologies and sharing of research data outside of a local institution. In phase 1 of this community engagement process, focus groups were used to obtain attitudes and perceptions of the general public and biorepository participants toward collecting and sharing genetic research data. This phase is presented here. Phase 2 involves a survey study of Institutional Review Board professionals and phase 3 will utilize consensus meetings of a variety of community and professional stakeholders. This study aims to elicit a better understanding of stakeholder views to inform best practices for conducting genetic research.

Subjects and Methods

Study Participants and Setting

Study participants included biorepository participants and members of the general public. Eligible biorepository study participants were NUgene participants, 18 years or older, not previously contacted for other research studies, and English-speaking. The NUgene project is a biospecimen repository with longitudinal medical information from participating patients at Northwestern University-affiliated hospitals and outpatient clinics [19]. Biorepository participants were randomly selected from existing NUgene participant lists. Six Chicago zip codes, representing the geographic areas in which the majority of NUgene participants reside, were used to identify and randomly select participants for this study. Biorepository participants were recruited by telephone to participate.

Eligible public participants were 18 years or older and English-speaking. Participants were recruited from 3 diverse Chicago neighborhoods representing the geographic areas from which the NUgene project recruits participants. Public recruitment strategies included engaging community partners at the focus group sites to advertise the research and posting flyers in the 3 targeted neighborhoods.

Participants received a confirmation letter, a participant information form, a consent form, and a one page fact sheet summarizing the NIH Data Sharing Policy prior to the focus groups. This study was approved by Northwestern University's Institutional Review Board. The fact sheet (Appendix A) described the reason for a data sharing policy and the types of information that will be included in the government database. Issues of identification and privacy protections were addressed. This general information was developed to provide a baseline understanding across focus group participants. Web addresses were provided on the fact sheet for further information on GWAS and the Policy for Sharing of Data in NIH-Supported or Conducted GWAS. The fact sheet was pretested with 3 individuals who met the focus group eligibility criteria and with the NUgene Community Advisory Committee.

Data Collection

Focus groups were chosen as a less structured method of eliciting data to allow for open discussion, varying viewpoints, discovery of unanticipated findings, and clarification of information. A brief participant information form, completed by each

Table 1. Focus Group Discussion Guide

1. What comes to mind when you think of genetic research data?
2. What might be some reasons that a person would participate in genetic research?
3. What might be some reasons that a person would not participate in genetic research?
4. How do *you* feel about participating in a genetic research study?
5. What kinds of information would you need to know before participating in a study to store and share your genetic research information?
6. If you agreed to participate in a study in which your genetic research information was stored in a database, what are your thoughts about sharing this data with other investigators?
7. How well do you think the privacy of individual genetic research data is protected?
8. What do you think the role of institutions and the government should be in protecting the privacy of those who participate in genetic research?

Table 2. Participant characteristics

	Public focus groups (n = 28)	NUgene focus groups (n = 21)
Gender		
Female	21 (75%)	14 (67%)
Male	7 (25%)	7 (33%)
Age		
Years, mean	42	48
Range	19–70	26–67
Education		
Less than high school	3 (11%)	0
High school graduate	4 (14%)	0
Some college	10 (36%)	5 (24%)
College graduate	6 (21%)	11 (52%)
Post college graduate	5 (18%)	4 (19%)
Employment		
Full-time	12 (42%)	13 (62%)
Part-time	5 (18%)	3 (14%)
Not employed	10 (36%)	4 (19%)
Race		
White	8 (29%)	16 (76%)
Black/African American	13 (46%)	4 (19%)
Other	6 (21%)	0
Ethnicity		
Hispanic or Latino	8 (29%)	0
Not Hispanic or Latino	19 (67%)	20 (95%)

One data value is missing for all categories except gender and age.

focus group participant, included standard demographic characteristic measures and 2 indicators about medical information and research.

Six audiotaped focus groups, 3 NUgene and 3 general public groups were conducted in May 2008. The focus groups consisted of 6–10 participants each and lasted between 1–2 hours. Prior to the start of the focus groups, informed consent was obtained from each participant. NUgene participant focus groups were conducted at Northwestern University's Chicago campus and the public focus groups were conducted at 3 Chicago neighborhood facilities. A USD 70 gift certificate to a general merchandise store and travel reimbursement were provided to focus group participants.

Experienced moderators used a focus group discussion guide consisting of 8 open-ended questions directly relating to the study's objectives (table 1). Specific probes were developed for each question to facilitate discussion. The focus group guide was pretested through 3 in-depth interviews with individuals who met the focus group eligibility criteria. Revisions to the focus group guide were made based on suggestions from the interviewees as well as from internal and external expert reviews.

Data Analysis

Focus group discussions were transcribed and independent checks, by 2 investigators, confirmed accurate and verbatim transcription. Transcripts were uploaded into Atlas.ti (version 5.2), a qualitative data management and analysis software program [20]. A codebook, with inclusion and exclusion criteria, was developed by the research investigators to identify key opinions and themes. Two investigators double coded a subset of transcripts and any coding differences were identified. Using this information, codes were further clarified in a revised codebook. An inter-coder reliability check was conducted using the revised codebook and agreement was reached in 89% or higher in 2 randomly selected transcripts. Data reduction and analysis were conducted through standard qualitative methods of coding, notation, and theme identification [21, 22]. Microsoft Excel was used to analyze the participant information form data.

Results

Participant Characteristics

All 49 individuals who participated in one of the 6 focus groups (21 in the 3 NUgene biorepository focus groups; 28 in 3 public focus groups) completed a participant information form. Demographic information from the public and NUgene participant focus groups is presented in table 2. In both group types, the majority of participants were women. Overall, the NUgene focus group participants were slightly older, more educated, and more likely to be employed than the public focus group participants. In the NUgene focus groups, the majority (76%) of participants was Caucasian; whereas in the public focus groups 46% were African-American and nearly one-third were Hispanic.

Two 5-point Likert scale questions were included in the participant information form to profile general participant views toward medical information and medical research. Overall, the participants' perspectives were similar between the public and NUgene participants. Approximately half of the public (46%) and NUgene (56%) participants indicated that they were somewhat or very concerned about confidentiality and privacy of medical information (table 3), and approximately half of all participants (public 46% and NUgene 48%) reported that they were somewhat or very trusting of medical research (table 4).

Table 3. Concern about confidentiality and privacy of medical information

	1	2	3	4	5
Public (n = 28)	2 (7%)	2 (7%)	9 (33%)	6 (21%)	7 (25%)
NUgene (n = 21)	1 (5%)	2 (10%)	5 (24%)	6 (28%)	6 (28%)

Range of one to two missing data values per group.
1 = Not very concerned; 5 = very concerned.

Table 4. Trust in medical research

	1	2	3	4	5
Public (n = 28)	1 (4%)	3 (11%)	10 (35%)	12 (42%)	1 (4%)
NUgene (n = 21)	1 (5%)	1 (5%)	8 (37%)	8 (38%)	2 (10%)

Range of one to two missing data values per group.
1 = Not very trusting; 5 = very trusting.

Table 5. Focus group major themes

1. A wide spectrum of understanding of genetic research
2. Weighing pros and cons of participation in genetic research
3. Influences on participation: credibility, trust and research integrity matter
4. Questions and concerns about sharing genetic research data and need for transparency in the NIH GWAS Data Sharing Policy
5. More information and education about genetic research needed

Major Themes Identified

Five major themes were identified in the focus groups (table 5).

A Wide Spectrum of Understanding of Genetic Research

The focus group participants' views and understanding of genetic research varied considerably within and between groups. The majority of participants associated genetic research with diseases and only a few participants in both focus group types mentioned specific genetic research data sources such as stem cells, DNA, tissues, 'bio-blood' and 'spit'. However, in one of the NUgene focus groups, there was considerable discussion about non-disease applications of genetic research, such as breeding Labradoodles, cloning cows, and conducting animal husbandry.

Expressions of both excitement and hope regarding the potential of genetic research were shared in all focus groups along with concern and fear. In one public focus group, a participant stated, 'So to me, what we are doing is great for the future … I think they are doing a great service.' In stating a specific concern about genetic research, one of the public group participants said, 'I live in a high crime area and I'm scared that what if they took my DNA and put it on somebody.' Other focus group discussions involved describing the opportunity for misuse of genetic research information such as 'playing god', 'genetic discrimination' and 'Big Brother'. Another concern raised was how genetic research findings are shared with the public. One public focus group participant stated, 'I'd have to say I'm a little skeptical of a lot of this data that they keep, you know … As far as some of the information, they're only going to feel like they only release what they want to release … they tell you what they want you to know.'

Genetic research was also discussed in terms of potentially altering human physical characteristics. One public focus group participant suggested this was a positive development while most other references to this type of genetic manipulation were negative. One NUgene participant stated, 'It's kind of scary. I think about back in the Hitler days, he was trying to make a perfect baby and now people are going around saying, "I want a blue-eyed baby, do this, do that I want a baby with brown hair." It's kind of spooky.'

In both the NUgene and public groups, there was lack of clear understanding about genetic research. In describing genetic research, one participant explained, 'Now, in

this day they have to process all of the blood to get all of these sicknesses out of it before they put it into your body … they take the sickness out before they give you the transfusion. That is the genetic research – that's what they do.'

Weighing Pros and Cons of Participation in Genetic Research

Focus group members discussed positive and negative aspects of participating in genetic research. There was more pro discussion for participating in genetic research in the NUgene focus groups than in the public groups. This may be because they had already decided to participate in genetic research. In the public focus groups, more discussion occurred identifying negative aspects of participating in genetic research. However, despite barriers and potential risks, most of the focus group participants indicated that they would consider participating in genetic research.

One of the positive reasons mentioned to participate in genetic research, in all focus groups, was the desire 'to help'. Language used to describe this included: 'for the greater good', 'it just speaks to your heart', 'to do my good deed'. Having a family member with a genetic condition was also mentioned as a motivating factor for participating in genetic research. For example, one public focus group participant said, 'Sometimes you don't feel like participating unless something happens within your own family or community.' Likewise, a NUgene focus group participant stated, 'I have a great-nephew who has sickle cell. I know his mother would be like, "Okay, let's do this DNA research."' Interest in genetic research and wanting to be a part of scientific achievement was also discussed by participants as reasons to participate. One public focus group participant said, '… out of curiosity for the science and stuff'. Another participant from the NUgene focus groups stated, 'The other thing is, and I think this touches on the very large numbers of people that are necessary in a lot of these studies. So even if any one of us, our genes are not that important, maybe we can do our little bit to share.'

Bettering society by curing disease was also mentioned by a number of focus group participants as a reason to participate in genetic research. One public focus group participant stated, 'That's your hope, you know, because at the end result you want results on diseases you know. You want some cures. That's the end result.' Similarly a NUgene participant said, 'I feel any disease if they can cure it, they can test me, yeah.' Other reasons mentioned by focus group participants as reasons to participate in genetic research included: potential cost savings to society and ease of participating.

Focus group participants mentioned several reasons for *not* participating in genetic research. Lack of information about and understanding of genetic research was reported as a potential barrier to participation. For example, one public focus group participant said, 'I don't really know a lot of what genetic researching does. Like, I haven't seen results of it … So I guess if I felt like it was doing something, you know, I'd be more likely to participate … But like there's so little information out there about it to feel like it's really accomplishing a whole lot.'

A significant barrier to participation mentioned in all focus groups was fear associated with genetic research. Fear of needles related to providing a sample was mentioned and fear of discovering a familial disposition toward a disease were discussed as barriers to participation in genetic research. One public focus group participant said, 'People, they're afraid of the unknown. So it's like if you don't know, then, … what you don't know can't hurt you.' Some participants mentioned that knowledge of genetic diseases in their family could influence their decision to conceive; therefore they did not want to have this information from genetic research.

Concern about genetic discrimination, due to the sharing of genetic information as a result of participation in genetic research, was also mentioned by participants as a barrier to participation. Discrimination was perceived possible by insurance companies, the government, the health care system, and employers. For example, one NUgene participant stated, 'If information is leaked, you could lose many of your personal privacy rights, … job discrimination, insurance discrimination, things like that.' Another NUgene participant said, 'I just wouldn't want to be denied care for it because of lack of insurance, because of information that was released about my family history.'

Influences to Participation: Credibility, Trust, and Research Integrity Matter

It was important for the focus group participants to know what organization/agency had oversight responsibilities for genetic research data as well as to clearly understand the research goal and motivation. Reputation and trust were key factors as one NUgene participant stated, 'I think part of it is probably the reputation of the organization performing the research. Clearly Northwestern and the processes, protocols that have been laid out here are reassuring and the fact that it's Northwestern. Now if some guy standing in front of a storefront medical clinic comes up with a piece of paper and a flipchart and says, "Okay, would you like to sign up here and

donate your DNA?" I'd keep walking on the street.' Insurance companies and pharmaceutical companies were noted as generally distrustful by participants and one public focus group participant suggested: 'If drug companies are the primary researchers and wanted to do genetic research, I would be very … skeptical. Yes, because I think they are already going into it expecting a certain outcome, you know.'

Across all 6 focus groups there was a consistent theme of distrust of the government as an oversight body for genetic research data. One public focus group participant said, 'The government does shady things, so you never know where it's going. So you can trust the government like only so far.' Regarding the government, one NUgene focus group participant stated, 'I don't even want them having oversight because to me oversight would become access. So that's why just no government involvement.' Some participants suggested the need for an independent oversight body but were unable to identify a particular entity. One NUgene participant suggested: 'a NGO (non-government organization) with a volunteer board of directors … who represent a wide slice of the population. That's the only way that you can get people who care.' Other participants wanted to know more about the underlying motivation for the oversight involvement. One public focus group participant stated, 'I think I would be more trusting if I could separate … whether that's research, medicine, what have you, from profit. If somehow those 2 are separated in society then I think I can be more trusting.'

Questions and Concerns about Sharing Genetic Research Data and Need for Transparency in the NIH GWAS Data Sharing Policy

In both focus group types, there were varying views on whether or not genetic research data should be shared with other investigators and participants discussed what they would require in order to feel comfortable having their data shared. Some requirements included: (a) similar study purpose, (b) adequate security and privacy checks, and (c) recontact to consent to participate in another researcher's study. Trust in the institution that stored data was again paramount. For example, one NUgene focus group participant said, 'It shouldn't be available to everyone but I would trust that there are people in charge of that who are looking in the right directions in terms of where that information should go and where it shouldn't.'

Some participants believed that genetic research data was really not going to be shared because the original institution would benefit more by keeping it for themselves. One public participant compared this to how drug companies deal with new research data on medications, 'And like I said, these drug companies which pay so much more for the research they don't want to share with anyone else.' Other participants mentioned that they felt a lack of control and that the data would be shared without their input anyway. One public participant stated, 'It's out of your hands, it's out of your control.'

Focus group participants were asked to review a one page fact sheet that summarized general information about the Policy for Sharing of Data in NIH-Supported or Conducted GWAS (Appendix A). Ninety percent of public and NUgene group members indicated that they had read the fact sheet prior to the focus group, however data was not collected on how many participants accessed the web link to the NIH Policy. Therefore, information discussed in the groups about the NIH Data Sharing Policy was likely primarily based on information provided in the fact sheet. In all groups, focus group participants discussed the need for more information and clarification about the policy. When discussing this policy, it became apparent that not all participants were aware of or familiar with the NIH. One public focus group participant stated, 'But I was wondering about this, the National Institutes of Health. What's that set up for? What is the make-up of the National Institutes of Health?' A NUgene participant asked, 'Who's the head of the NIH … who's making policy, who's involved in these decisions?' In addition, participants wanted to know more about how the data will be shared and with whom. In reference to the fact sheet summary of the NIH Data Sharing Policy one NUgene participant said, 'Name names, I thought it was vague.' Several participants wanted more information on consequences of a breach of protection. One public focus group participant said, 'I'll tell you what I missed – the consequences, like if this was somehow released in a way that was not covered by any of these policies. What's in place to correct that, to fix it – that's not addressed here.' Another public focus group participant wanted to know more about penalties: 'Penalties will be imposed for misuse of the database. Okay, so what is the misuse of the database? What specifically are you saying? And what are the penalties?'

More Information and Education about Genetic Research Needed

Throughout both focus group types, participants spontaneously discussed why more information is needed about genetic research and shared methods they felt would work best to increase public awareness. In the NUgene focus groups, some participants discussed that it is important to provide complete disclosure about

genetic research to reduce fears and build trust. One NUgene participant said, '… I think that education in general is a really good tool for the layperson to understand what's happening and I just think that that's probably the key, educating away the fears but also disclosing – what are we responsible for? What could happen? It's going to help people trust what you're doing a little more too.' Another NUgene participant wanted more accurate information on genetic research because of concern about how the media can distort scientific information: 'I can read this one day and then read something totally different the next day. The media could put something out like your DNA is going to do this, do that, … and you get scared. You don't know.' In the public focus groups, participants mentioned various reasons for wanting more education about genetic research including to learn more about research progress, to become more comfortable with the information, and that they had a right to know what is going on in this area. One public focus group participant stated, 'I believe there's a lot of things that's going on that maybe we just don't know about. You know, … groups like this (focus groups), … reminding me a little bit about what more is going on. If they let the public know, like with cancer research … getting that data out there, the information out there to the public, letting them know.' Another public focus group participant said, 'You're telling people information you know, to where it helps. It helps relax your mind to where you feel more comfortable with what you're receiving.'

Both group types discussed ways in which they thought public awareness about genetic research could be raised. In the NUgene focus groups, specific education strategies discussed included targeting various groups such as perceived disengaged groups (low-income, minority groups), young people, schools, neighborhoods, and disease support groups. Another suggestion was to target communities where a particular genetic condition occurs more frequently than in the general population. One NUgene participant said, 'If you wanted to target more culture to culture, I know that asthma incidence is really high in the Hispanic and black communities and there's some new genetic markers for asthma. So you would say, "Oh, you know, you can help out. We've found things for asthma", … because that is more likely to hit home, you know, and be relevant where it might not be for other populations.' Although the media had been generally discussed as a poor, yet common, source of accurate information about genetic research, the nightly news was reported by one public focus group participant as a good method of reaching people. Other suggested methods of education included the internet, conducting interviews and focus groups with specific disease groups and communities, and educating family members – a primary source for health-related information. More discussion occurred in one of the public focus groups about the need for separate education approaches for diverse communities, respecting beliefs, and that it might be 'taboo to talk about some of that stuff'. A number of participants indicated that the focus group experience itself was a helpful educational experience. An interest was expressed by some participants to meet again and one public focus group participant said, 'I think there should be more than this meeting. I found this very stimulating.'

Discussion

Members in NUgene and public focus groups demonstrated a wide range of understanding about genetic research. Diverse public perceptions of genetic research are similarly documented in a review of the opinion poll literature and a focus group study conducted by Bates et al. [23]. Although focus group participants discussed both positive and negative reasons for participating in genetic research, overall, most indicated they would consider participation to benefit individuals and society. Participants identified trust in the research organization as a key positive influence to participation in genetic research. This finding is consistent with other studies [18, 24–26], and trust in the research organization has been identified as particularly important in recruiting minority research participants [25]. Considerable discussion occurred in both group types concerning lack of trust in governmental oversight of genetic research data and the need for transparency in data sharing and monitoring policies. Concerns about data sharing in genetic research were more intensely debated, however, in the public participant groups. In contrast, the same types of concerns about data sharing have not been documented as key issues raised by participants in other biobank community consultation efforts [2, 12, 13]. This difference may be due to a greater involvement of governments in national biobanking efforts as well as existing and proposed legislation protecting biobank participants and data in other countries.

This study revealed several new insights into participant views toward sharing of genetic research data. Some of the public focus group participants had questions and concerns regarding the purpose and mechanism for sharing genetic research data, with a particular concern over

potentially undisclosed commercial interests. Participants outlined elements they felt were important to know and see documented in the NIH Data Sharing Policy: (a) identification of specific names and roles of oversight persons in the NIH; (b) great detail describing how data will be shared and under what specific circumstances; (c) clear explanations of circumstances constituting breach of confidentiality; (d) information on specific penalties imposed on researchers who misuse data sharing. Many participants expressed the need for additional information to feel secure in allowing their genetic research data to be shared with other investigators and institutions. This finding is consistent with McGuire and colleagues who found that research participants desired more explicit information in order to make decisions about sharing genetic research data [18]. Research participants' need for names of responsible oversight individuals, as well as clear penalties for violators of policy, may be important elements for obtaining trust and participation in research.

One topic that emerged throughout the focus group sessions was the need for more public education about genetic research. Other investigators have similarly found that education, particularly about the complex information involved in data sharing and genetic research, should be broadly communicated and aimed at the public and targeted research populations [18, 26]. Focus group participants offered education methods and strategies they thought would be most suitable for the public. Targeting specific groups where genetic research might be more relevant was suggested, such as disease advocacy groups, and educational interventions involving the whole family. Participants mentioned the importance of identifying appropriate educational methods and approaches for diverse groups. Participants also emphasized that genetic research information should be culturally sensitive and tailored to specific community constituents.

The issues and concerns expressed by the focus group participants illustrate some of the competing interests of various stakeholder groups involved in conducting genetic research [27]. The research community has an interest in sharing genetic research data. In fact, one of the primary objectives of the NIH Data Sharing Policy is to promote research efforts through the sharing of genetic research data [7]. Research participants have expressed different concerns. Many focus group participants in this study wanted mechanisms in place that would protect the privacy and confidentiality of their genetic information. Accordingly, policies should address both researchers' interests in collaboration and sharing of data with research participants' interests in confidentiality and privacy. Therefore, in addition to the public comment and town hall style meetings that were held to inform the data sharing policy [7], further mechanisms will need to be developed to capture a diverse and broad representation of the public's views on data sharing as well as to build partnerships between public and research communities. Potential mechanisms include utilizing existing community-academic partnerships and disease organizations to learn about most appropriate methods of reviewing, developing and disseminating genetic data sharing educational materials with their communities. Focusing on genetic research issues that are relevant and of interest to a community will be important in initiating dialogues and developing partnerships. In particular, traditionally underrepresented and underserved populations could be targeted so that a broader range of views is included. Additional mechanisms that may lead to greater transparency and public understanding could include: (a) an assessment of the accessibility of existing NIH data sharing policies to the public; (b) development of materials that are written at a general public literacy level; (c) dissemination of the educational materials to investigators enrolling participants into studies utilizing GWAS technologies.

While focus group studies can provide rich insights into specific topics and allow new issues to emerge, this methodology has inherent limitations. Such limitations include lack of generalizability and inability to discern the prevalence of the emergent themes across a general population. Given that the study participants resided in a metropolitan area, their views may not be representative of those from other geographic areas. In addition, study participants in both NUgene and the public focus groups were more highly educated, so their views may not represent the types of opinions held by individuals with less education. However, this study involved a good balance of racial and ethnic diversity of public focus group participants, allowing for a depth and range of divergent opinions. Another potential focus group limitation is that responses may represent the most vocal participants, despite moderator encouragement for input from all. Thus, differing views may not be expressed if participants do not feel comfortable voicing their opinions. To address these 2 issues, this study utilized experienced moderators and standardized focus group study procedures [28]. Another limitation is that an assessment was not undertaken to determine how many participants accessed the URL on the fact sheet to read the NIH Data Sharing Policy. Discussions were most likely based on the fact sheet information.

To our knowledge, this is one of the first studies conducted to assess public and biorepository participant views of key elements of the NIH Data Sharing Policy issued in 2007. Data from this study raise a number of policy considerations. These include: broader efforts to obtain input from the general public to address specific concerns about data sharing and control of genetic research data; the need for continuing public education about the interrelationship between genetic research efforts and data sharing; and an assessment of the impact of these issues on future participation in genetic research. To build community partnerships in the data collection and sharing process for GWAS, and other large scale genetic studies, genetic research oversight bodies should consider ongoing community involvement in oversight committees, policy development, and educational initiatives. Future studies will need to consider best methods to educate the general public and target populations about research goals, methods, and potential outcomes. To facilitate cultural sensitivity, target populations should be invited to participate in the development and review of these educational strategies.

Acknowledgements

We thank both the general public and NUgene participants who served as focus group participants, and we are grateful to the community partners at the public focus group sites for their support and assistance. We also thank the focus group moderators and assistant moderators for their contributions. Funding for this study is provided by the National Human Genome Research Institute, U01HG004609.

References

1 Ollier W, Sprosen T, Peakman T: UK Biobank: from concept to reality. Pharmacogenomics 2005;6:639–646.
2 Godard B, Marshall J, Laberge C: Community engagement in genetic research: results of the first public consultation for the Quebec CARTaGENE project. Community Genet 2007;10:147–158.
3 Kaiser Permanente News Center: AScribe Medicine News Service. Kaiser Permanente unveils groundbreaking genetic research program; researchers aim to reveal genetic and environmental causes behind deadly and disabling diseases. February 14, 2007. Available at http://www.accessmylibrary.com/coms2/summary_0286-29609891_ITM.
4 Wellcome Trust Case Control Consortium: Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 2007;447:661–678.
5 Knoppers BM, Fortier I, Legault D, Burton P: The public population project in genomics (P3G): a proof of concept. Eur J Hum Genet 2008;16:664–665.
6 Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R, Hao L, Kiang A, Paschall J, Phan L, Popova N, Pretel S, Ziyabari L, Lee M, Shao Y, Wang ZY, Sirotkin K, Ward M, Kholodov M, Zbicz K, Beck J, Kimelman M, Shevelev S, Preuss D, Yaschenko E, Graeff A, Ostell J, Sherry ST: The NCBI dbGaP database of genotypes and phenotypes. Nat Genet 2007;39:1181–1186.
7 National Institutes of Health: Policy for sharing of data obtained in NIH supported or conducted genome-wide association studies (GWAS). Federal Regist 2007;72:49290–49297 [cited 2009 January 16]. Available at http://grants.nih.gov/grants/guide/notice-files/NOT-OD-07–088.html.
8 Caulfield T, McGuire AL, Cho M, Buchanan JA, Burgess MM, Danilczyk U, Diaz CM, Fryer-Edwards K, Green SK, Hodosh MA, Juengst ET, Kaye J, Kedes L, Knoppers BM, Lemmens T, Meslin EM, Murphy J, Nussbaum RL, Otlowski M, Pullman D, Ray PN, Sugarman J, Timmons M: Research ethics recommendations for whole-genome research: consensus statement. PLoS Biol 2008;6:430–435.
9 National Institutes of Health, National Human Genome Research Institue: Design considerations for a potential US population-based cohort to determine the relationships among genes, environment and health: recommendations of an expert panel [cited 2009 February 6]. Available at http://www.genome.gov/Pages/About/OD/ReportsPublications/PotentialUSCohort.pdf.
10 National Institutes for Health (NIH), Office of Science Policy: SACGHS report chapter IV: Need for public engagement. Policy issues associated with undertaking a new large US population cohort study of genes, environment and diseases [cited 2009 February 6]. Available at http://oba.od.nih.gov/SACGHS/sacghs_focus_population.html.
11 National Institutes of Health (NIH), Office of Science Policy: Points to consider when planning a genetic study that involves members of named populations [cited 2009 February 6]. Available at http://bioethics.od.nih.gov/named_populations.html.
12 Haddow G, Cunningham-Burley S, Bruce A, Parry S: Generation Scotland: consulting publics and specialists at an early stage in a genetic database's development. Crit Public Health 2008;18:139–149.
13 Godard B, Marshall J, Laberge C, Knoppers BM: Strategies for consulting with the community: the cases of four large-scale genetic databases. Sci Eng Ethics 2004;10:457–477.
14 O'Doherty KC, Burgess MM: Engaging the public on biobanks: outcomes of the BC biobank deliberation. Public Health Genomics 2009;12:203–215.
15 Ormond KE, Cirino AL, Helenowski IB, Chisholm RL, Wolf WA: Assessing the understanding of biobank participants. Am J Med Genet A 2009;149A:188–198.
16 McCarty CA, Nair A, Austin DM, Giampietro PF: Informed consent and subject motivation to participate in a large, population-based genomics study: the Marshfield Clinic Personalized Medicine Project. Community Genet 2008;10:2–9.
17 U.S. Public Opinion on Uses of Genetic Information and Genetic Discrimination: genetics and public policy center, 2007. Available at http://www.dnapolicy.org/resources/GINAPublic_Opinion_Genetic_Information_Discrimination.pdf.
18 McGuire A, Hamilton JA, Lunstroth R, McCullough LB, Goldman A: DNA data sharing: research participants' perspectives. Genet Med 2008;10:46–53.
19 Wolf WA, Doyle MJ, Aufox SA, Frezzo TF, Smith ME, Kibbe KE, Chisholm RL: DNA banking study in an ethnically diverse urban university hospital. Am J Hum Genet 2003;73:423.
20 ATLAS.ti Scientific Software Development GmbH: c 2002–2009 [cited 2009 February 6]. Available at http://www.atlasti.com/.
21 Miles MB, Huberman AM: Qualitative Data Analysis. Thousand Oaks, Sage Publications, 1994.
22 Krueger RA, Casey MA: Focus Groups: A Practical Guide for Applied Research, 3rd edition. Thousand Oaks, Sage Publications, 2000.

23 Bates BR, Lynch JA, Bevan JL, Condit CM: Warranted concerns, warranted outlooks: a focus group study of public understandings of genetic research. Soc Sci Med 2005;60: 331–334.
24 Kass NE, Sugarman J, Faden R, Schoch-Spana M: Trust: the fragile foundation of contemporary biomedical research. Hastings Cent Rep 1996;26:25–29.
25 Corbie-Smith G, Thomas SB, Williams MV, Moody-Ayers S: Attitudes and beliefs of African Americans toward participation in medical research. J Gen Intern Med 1999;14: 537–546.
26 Henderson G, Garrett J, Bussey-Jones J, Maloney ME, Blumenthal C, Corbie-Smith G: Great expectations: views of genetic research participants regarding current and future genetic studies. Genet Med 2008;10:193–200.
27 Foster MW, Sharp RR: Share and share alike: deciding how to distribute the scientific and social benefits of genomic data. Nat Rev Genet 2008;8:633–639.
28 Morgan DL, Krueger RA: The Focus Group Kit, Volumes 1–6. Thousand Oaks, Sage Publications, 1998.

Appendix

National Guidelines for Sharing Genetic Research Data

The National Institutes of Health (NIH), the national agency responsible for overseeing government sponsored biomedical research, has worked on developing guidelines, or policies, for how research data is to be shared.

Why is a data sharing policy needed?
As scientists increasingly try to find the causes and cures for common conditions such as diabetes, heart disease, and arthritis, they must carry out their research on very large numbers of people in order to get useful results. It is difficult for any one investigator to have access to such large populations to study. By pooling research data into one database, many investigators can use the compiled data, thereby increasing the chance of finding answers to research questions. Additionally, this allows investigators to avoid repeating experiments that have already been performed, unnecessarily taking up resources and time.

What is the NIH Data Sharing Policy?
Certain studies that receive funding from the NIH are required to share their research findings with other investigators. Many of the studies contain genetic information from research participants of ***genome-wide association studies***.

The policy requires that research information be kept in a safe ***database*** maintained by the NIH. Limited use of the database will be allowed only for investigators who apply for access, pass a security check, and show evidence about why they need to use the information in the database. Investigators must agree to certain conditions in order to be granted access to the information. Once an investigator has been granted access, he/she can see all the information kept in the database about the research participants.

A **genome-wide association study** tries to identify genetic differences related to a particular disease by rapidly scanning markers across the complete sets of human DNA, or genomes, of many people.

A **database** is a collection of related information organized for convenient access, generally in a computer.

What type of information will be included in the database?
The database will contain information about research studies such as procedures, questionnaires, study manuals, and types of tests conducted. ***Coded*** information about research participants will also be included. The coded information may consist of medical test results, physical exam information, and medications taken. Demographic information, such as gender, age, or ethnicity may be included about the research participant but may not contain direct identifiers such as name, date of birth, address, or social security number. The database will also contain results of studies, including DNA information from genome-wide association studies.

Coded information is information in which a name or other direct identifier has been replaced with a random set of numbers and/or letters, such as a barcode.

Can a person be identified by their DNA alone?
In theory, current technology makes it possible to compare the results of a genome-wide association study to the DNA code from a known person, which could lead to the identification of a research participant. In this way, a third party might be able to figure out who the information contained in the database belongs to. The privacy protections listed below are meant to ensure only persons with a real research interest and who have appropriate reasons for using the data are able to access the information contained in the database.

How is the privacy of research participants in the database protected?
The privacy of research participants is protected in a number of different ways, including:

1) All information related to research participants is coded and no direct identifiers will be used.
2) The key to the code will be kept by the original investigator, not by the NIH.
3) Access to the database is restricted to investigators whose research has been approved.
4) Investigators using the database must agree to never attempt to identify any person in the database.
5) Penalties will be imposed for misuse of the database.
6) The NIH plans to obtain an exception to the Freedom of Information Act so that data will not be available to law enforcement or other government agencies.

Further information about Genome-wide association studies can be found at: www.genome.gov/20019523

Further information about the NIH Data Sharing Policy can be found at:
http://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html