Published in final edited form as:
Cancer Epidemiol Biomarkers Prev. 2008 August ; 17(8): 2019–2024. doi: 10.1158/1055-9965.EPI-07-2868.

Studying Genetic Research Participants: Lessons from the "Learning About Research in North Carolina" Study

G. Corbie-Smith[1,2,4], **C. Blumenthal**[4], **G. Henderson**[1], **J. Garrett**[4], **J Bussey-Jones**[7], **M. Moloney**[4], **R.S. Sandler**[2,5], **S.W. Lloyd**[4], **J. Dorrance**[4], and **J. Darter**[4]

[1]Department of Social Medicine University of North Carolina

[2]Department of Medicine University of North Carolina

[3]Department of Health Behavior and Health Education University of North Carolina

[4]Cecil G. Sheps Center for Health Services Research University of North Carolina

[5]Center for Gastrointestinal Biology and Disease University of North Carolina

[7]School of Medicine Emory University

Abstract

Given the prohibitive cost of recruiting large and diverse populations for genetic explorations in cancer research, there has been a call for genetic studies to engage *existing* cohorts of research participants. This strategy could lead to more efficient recruitment and potentially result in significant advances in the understanding of cancer etiology and treatment. The Learning About Research in North Carolina (LeARN) study responded to the National Human Genome Research Institute interest in research on how study participants from diverse populations who had participated in genetic research perceived the risks and benefits of participating in combined epidemiologic-genetic research, how well they understand the purpose of the research and the uses to which the research results may be put, and how involvement in such research affects perceptions of disease causality. In this paper we give an overview of the LeARN study, summarizing the methods we used, challenges we encountered, and lessons learned about recruiting participants who have previously participated in genetic research.

Keywords

cancer; genetic; epidemiology; research; methods

Introduction

Rapid advances in the Human Genome Project have scientists poised to realize the promises of genetic discoveries in the study of gene-environment interactions and genetic variation in cancer. The pace of these discoveries has led to an increasing need to engage populations in a variety of genetic research protocols. Involvement in such studies may provide no clinically relevant information to research participants, yet participants' willingness to become involved in such research is critical to continued advancement of the field. Given the prohibitive cost of recruiting large and diverse populations for genetic explorations, there has been a call for genetic studies that engage *existing* cohorts of research participants(1).

Corresponding Author: Giselle Corbie-Smith, Cecil G. Sheps Center for Health Services Research, 725 Martin Luther King Jr. Blvd., Chapel Hill, NC 27599, gcorbie@med.unc.edu.

This strategy could lead to more efficient recruitment and potentially result in significant advances in the understanding of cancer etiology and treatment.

Within the biomedical and research ethics communities, there is a call to carefully examine how groups of individuals and large populations would respond to such involvement, the most appropriate manner of engagement and an examination of the ethical, regulatory, scientific, and public health implications of such an endeavor (1–3). The Learning About Research in North Carolina (LeARN) study responded to the National Human Genome Research Institute call for research on how study participants from diverse populations perceive the risks and benefits of participating in combined epidemiologic-genetic cancer research, how well they understand the purpose of the research and the uses to which the research results may be put, and how involvement in such research affects perceptions of cancer causality.

The LeARN study takes advantage of a unique opportunity to examine participation in cancer genetic variation research by building on a series of ongoing and successful population-based studies of colorectal cancer in North Carolina, the North Carolina Colorectal Cancer Study (NCCCS). NCCCS sought to understand racial difference in cancer risk by collecting data on genetic, environmental and behavioral predictors of cancer. This research is an example of an emerging type of population-based genetic variation research that combines traditional epidemiological methods with contemporary laboratory analysis of genetic polymorphisms to understand individual and group differences in cancer risk.

As we considered the study design and data collection approaches for the LeARN study, we found the empirical literature lacking in guidance on the best research methods to accurately capture the views of potential or former participants in genetic research, particularly in groups that may have varied educational attainment and experience with genetics. This methodological gap in the literature has become a crucial omission as more investigators seek to engage diverse communities in genetic research. In this paper we give an overview of the LeARN study, summarizing the methods we used, challenges we encountered, and lessons learned about recruiting participants who have participated in genetic studies.

1. Methods

Study design

Learning About Research in North Carolina (LeARN) is an interview study that contains both a longitudinal component and a cross-sectional component. The first aim of the LeARN study was to describe participants' perceptions of causality of disease and how perceptions of disease causality might change after participation in a genetic-epidemiologic study. To accomplish this aim, we used a longitudinal design where NCCCS agreed to administer closed-ended questions on perceptions of disease causality at the beginning of the 2 hour NCCCS face-to-face interviews; the same questions were administered in the LeARN telephone interview. By comparing lay beliefs prior to participation in the NCCCS study and shortly thereafter, we were able to determine how perceptions about causation change before and after exposure to NCCCS.

The remaining aims of the LeARN study were to describe what motivates research participants to participate in a genetic-epidemiology study; to describe participants' perceptions of positives and negatives related to collecting genetic data in epidemiologic research, and to determine whether those perceptions vary by disease status and/or race/ethnicity; to describe how existing research participants understand the purpose of genetic research, and whether their understanding differed by disease status and/or race/ethnicity.

These aims were met through data collected in a cross-sectional study comparing cases and controls and race groups (whites and blacks).

Overview of the North Carolina Colorectal Cancer Study (NCCCS)

The NCCCS is a series of population-based studies that have examined the etiologic factors in colorectal cancer (4). The current case-control study includes 1,047 incident cases of invasive rectosigmoid cancer and 1,016 population-based controls from a 33-county area of central and eastern North Carolina. Cases were eligible for inclusion in NCCCS if they were between the ages of 40 and 80, resided in the 33-county area, and had a first diagnosis of invasive rectosigmoid cancer. Cases were identified from November 2001 through December 2006 using the rapid ascertainment system that has been implemented by the North Carolina Central Cancer Registry. NCCCS recruited all black participants and a probability sample of white patients with new rectosigmoid cancer. Age, race and sex matched controls were selected from two sources: Division of Motor Vehicles records for those under the age of 65 and Center for Medicare and Medicaid Services (CMS) tapes for those 65 years and older.

Once consent was obtained, trained nurse-interviewers collected data for NCCCS during in-person, in-home interviews for both cases and controls. The two-hour interview on determinants of colorectal cancer included data on dietary, lifestyle and environmental exposure, information about education, occupational status, poverty, and health care access and utilization. Blood and/or mouthwash samples were obtained from consenting participants at the conclusion of the interview. DNA and serum were stored for future analyses.

At the completion of the interview, all participants in the NCCCS were asked if they were interested in hearing about other research studies. (“From time to time other research studies become available. Should such a study become available, may we contact you?” [yes/no])

LeARN study sample

Participants for the LeARN telephone interviews were identified through the NCCCS database of participants interested in hearing about other studies. Participants from NCCCS were eligible to participate in Project LeARN if they met the following criteria: 1) self-reported race of black or non-Latino white, 2) completed the entire interview required of NCCCS, 3) agreed to be contacted about future studies, 4) lived in the state of North Carolina at the time of the LeARN study, and 5) had sufficient cognitive functioning, as assessed by the interviewer, to allow successful completion of the telephone interview.

Recruitment Procedures—Potential participants were mailed a letter of introduction by the NCCCS investigators that introduced the LeARN study, described the telephone interview, and alerted them to expect a follow-up telephone call. We contracted with the professional survey group, FGI, Inc., to conduct the telephone interviews using Computer Assisted Telephone Interviewing (CATI) methods. Potential participants were contacted by FGI on average 17 weeks after completing the NCCCS interview. LeARN interviewers were racially diverse, trained to address barriers to research participation and received extensive training on qualitative interviewing.

We obtained verbal consent for the telephone interview. After completion of the LeARN interview, each participant was mailed an incentive of $25 to partially compensate them for their time. We used a follow-up protocol to maximize response rates. Interviewers made at least 10 attempts to contact each person and perform one refusal conversion attempt per

refusal. All procedures were approved by the UNC and Emory University Institutional Review Boards.

Final LeARN Sample—During the recruitment period for the LeARN study, June 2003 to March 2007, 1292 NCCCS participants were eligible to be recruited to LeARN (Figure 1). Of these 1292, 97% (n=1257) had indicated during the NCCCS interview that they would be interested in hearing about other studies, and 95% (n=1196) of those agreed to be contacted after receiving the initial LeARN letter inviting them to participate. We closed recruitment to LeARN in March 2007, with a total of 810 participants, 63% of the original NCCCS sample. After excluding people with unusable calls (e.g. disconnected, deceased, or respondent unknown) or with no answer after 10 attempts, the final response rate was 73% (5).

Data Collection and Areas of Inquiry

In order to meet the specific aims of the LeARN study we asked participants questions on a range of topics having to do with research participation. We used a mixed methods approach of both closed and open ended questions in order to accurately capture participants' views on these topics. The final LeARN Participant Questionnaire contained 94 items- 16 open ended and 78 closed ended questions- and took on average 44 minutes to complete. Topics covered in our closed-ended questions included: demographic information, including self-identified race; history of colorectal cancer; perceived causes of colorectal cancer; how much they had heard about genetic research and genetic variation research and how positive or negative they felt about this type of research; whether or not they gave blood or saliva for the NCCCS study and their reasons for giving or refusing to give a sample; motivators for participation in genetic research; concerns with providing genetic samples for research; and trust of research, medicine and investigators. Our open ended questions were designed to facilitate in-depth understanding of participants' explanatory models of disease causality; their sources of information on cancer; and their views on positives and negatives for them, their family and society of genetic research and genetic variation research.

2. Lessons Learned

a. It is possible to recruit for “studies of studies”

While LeARN was an interview study that did not request use of previously donated biological samples, our overall response rate of 73% indicates the potential for successful recruitment of previous participants. LeARN participant characteristics were similar to those of people in the parent study (NCCCS) who expressed interest in participating in future studies (Table 1). However, when comparing the final response rate (n=810) to eligible participants (n=1108), we found that fewer blacks participated in the LeARN interview than whites (65% vs. 75%, p=0.001) and participants older than 60 were more likely to complete the interviews than those under 60 (p=0.022). 10% of blacks vs. 4% of whites in the NCCCS study had contact information that was classified as unusable (e.g. disconnected, deceased, or respondent unknown) and this may partially explain different participation rates. .

Competition from other studies was a surprising, additional challenge. During the 30 month recruitment period for LeARN, two other studies also recruited from the NCCCS study. One recruited all black cases into a genetic study of hereditary determinants of colorectal cancer.

The second recruited both black and white cases into a study of determinants of outcomes of colorectal cancer. Simultaneous recruitment may have been a factor in our differential recruitment of blacks into LeARN, especially considering that blacks were somewhat more likely to say that they were interested in hearing about other studies at the early stages of

recruitment. At the beginning of our collaboration with NCCCS we were able to institute procedures that minimized impact from other studies: Clear tracking procedures and monthly tracking reports, scheduled secure data transfers, regular and ad hoc in person and electronic communication..

b. Developing good questions about genetic research is challenging

To develop clear, understandable questions abut genetic research, we first used cognitive interviews, a qualitative research method, to assess existing measures and questions drawn from another ongoing study of participation in genetic research.[1] Cognitive interviewing is a well-accepted approach to questionnaire development for difficult topics (6–7). In a cognitive interview, a respondent is interviewed one-on-one by a trained interviewer who asks the respondent to "think out loud" as he develops his answers. In addition, the interviewer may also ask more directed probes after the respondent has answered the question to further assess comprehension of a particular term or concept. These techniques can help determine if questions are being interpreted and answered as intended. Twenty-one cognitive interviews were conducted by trained interviewers and were audio-taped and transcribed for analysis of open-ended questions. We used this process to modify questions as needed. For example, in our initial wording we asked participants "What do you think the researchers were trying to find out in the colorectal cancer study?" We found that participants needed more prompts to remind them about their participation in NCCCS and the context of that participation. This question was then modified to read "What do you think the researchers were trying to find out in the colorectal cancer study, the one where the nurses came to your house?"

Following the cognitive interviews, we conducted telephone interviews with subjects to pre-test and refine the questionnaire and test the feasibility of administering a structured or semi-structured questionnaire via telephone. However, despite in person cognitive interviews and telephone pre-testing, we found that our open ended questions were still problematic. Respondents did not understand terms like `genetic variation research', and had trouble providing their assessments.

In response to this, we conducted two focus groups, one with blacks and one with whites, to refine the wording of questions on the meaning of genetic and genetic variation research; the positive and negatives of participating in genetic research and genetic variation research. For example, the initial wording of one question was: "In general, what do you consider as the positives of taking part in genetic research studies?" Focus group participants recommended using the term "good things" rather than positives, separating questions into good things for self/family vs. good things for society, and giving a concise definition of genetic research to orient respondents. We modified this question to read, "We're going to be asking you about good things and bad things for you, your family and society about taking part in genetic research. This may seem repetitive, but we want to get all of your thoughts down. So, first, could you give me a list of the good things for yourself and your family about taking part in genetic research studies, the kind that look at whether genes put people at risk for disease or illness?" The subsequent question then reads, "And what are the good things for society about genetic research studies, the kind that look at whether genes put people at risk for disease or illness?"

Participants offered varied responses to these questions, ranging from very positive to concerns about possible misuse of findings. Despite our intensive efforts to create understandable ways to ask about perceptions of genetic research, for these 4 questions,

[1] S.C. Hull and B.S. Wilfond, personal communication

between 6.2 and 9.2% of the sample either said they did not understand or answered in a manner that made us question whether they understood these open-ended questions. We wondered if these were valid responses or were indications of remaining problems with the internal validity of these questions. For example, some answers seemed to imply that respondents did not understand why we would ask about the meaning of genetic research, a topic with little salience for them. As one participant noted after the interviewer attempted to reorient him to the question “I really had never given that a thought”. We don't know whether these respondents just needed more time to think through the questions or if there is a better way to frame questions to capture latent views. It is also possible that these “don't understand” responses simply reflect different levels of lay knowledge…In a separate analysis(8), we found participants with less education, lower income and no prior personal or family history of cancer were most likely to say that they didn't understand questions on the meaning of genetic research. As the field of genetics continues to become more established in lay parlance, we might find greater understanding of these types of questions.

Conclusions

LeARN provides a unique opportunity to understand the challenges of recruiting diverse populations that are part of existing genetic research, to understand their perceptions of genetic research and the motivations and concerns of subjects most likely to be asked to participate in large cohort studies. Our results are particularly relevant for new NIH funded collaborations in which biologic samples and data from multiple existing cohort studies are linked to provide a powerful resource for new types of genetic research. In the draft report the Secretary's Advisory Committee on Genetics, Health and Society calls for engagement of populations in all stages of the decision making, planning, and execution of genetic research. Included in these populations are those that might be involved as participants in genetic research (1).

We found that enlisting participation, even from those who have already agreed to participate in genetic research, has its own challenges. While the majority of individuals in NCCCS expressed interest in other studies, we were ultimately able to recruit a smaller subset of eligible and interested respondents. The accuracy of contact information had a differential impact on participation of blacks and whites. Our efforts were also complicated by the fact that we were not the only research team attempting to recruit individuals who have participated in genetic research. Second, we found that asking questions about genetic research is challenging. To ensure external and internal validity, we extensively pre-tested open ended questions on views of genetic research, and used a mixed methods approach to collect data that would provide more detailed and nuanced information on participants' views of genetic research. In fact, the data we collected were rich and varied, representing a broad spectrum of views. However, we were left pondering the responses of a small percentage of respondents who seemed not to understand the questions we posed.

The attractiveness of existing cohorts is clear. The cost of assembling new cohorts of individuals can be prohibitive; if sampled and assembled carefully, existing population-based studies provide an efficient method of recruitment for other research, particularly when a convenience sample rather than random probability sampling is appropriate. Despite clear efficiencies, our findings suggest that the informed consent process still needs to be deliberate and detailed. A substantial minority of LeARN participants said that they didn't understand questions on the positives and negatives of genetic research. Tailoring the consent process to educational attainment or other factors that affect participant understanding may increase understanding for this subgroup. Despite the extensive, multi-staged consent process that NCCCS used with all participants, it is possible that

participation in a prior study will only incrementally increase understanding of genetic research.

As evidenced by this research, raising genetic literacy and awareness among the lay public will also be an important step in ensuring full participation of communities in genetic research. One such attempt has been a sequence of regional conferences held annually since 2005. These Community Genetics Forums, sponsored by the Education and Community Involvement Branch at the National Human Genome Research Institute, are designed to create models of community engagement and to enhance existing models of public engagement and public participation (9). Over the course of a year they bring together the scientific and lay community to discuss the ethical, legal and social implications of genetic research for communities, using a variety of venues from lectures to performance art (10–11). Continued efforts to raise genetic literacy of potential participants will be critical to seeing the potential of genetic research realized in a manner that takes into account the perspectives of diverse groups that are to benefit from this research.

Acknowledgments

Jim Sorenson

Supported by grants from the National Institutes of Health (R01 HG002830, P20 HG03387, P30 DK34987, R01 CA 66635) and an NRSA training grant (T32 HS000032)

References

1. Secretary's Advisory Committee on Genetics, Health, and Society. Policy issues associated with undertaking a new large U.S. population cohort study of genes, environment, and disease. Mar 12007 [cited 2007 June 5]. Available from: http://www4.od.nih.gov/oba/sacghs/reports/SACGHS_LPS_report.pdf
2. National Institute of Health. Studies of the ethical, legal, and social implications (ELSI) of human genetic variations research for individuals in diverse racial and ethnic groups RFAHG-02-003. Nov 152001 [cited 2007 June 5]. Available from: http://grants.nih.gov/grants/guide/rfa-files/RFA-HG-02-003.html
3. National Human Genome Research Institute & National Institute of Health. Community Genetics Forum: A model community engagement program NHLBI-PB(HG)-2007-001/KLW. Oct 32006 [cited 2007 June 5]. Available from: http://www.genome.gov/Pages/About/OD/OPCE/ECIB/CommGenForum/2007-001CommunityGeneticsForumRFP.pdf
4. Satia-Abouta J, Galanko J, Potter J, Ammerman A, Martin C, Sandler R. Associations of total energy and macronutrients with colon cancer risk in African Americans and whites: Results from the North Carolina Colon Cancer Study. Am J Epidemiol 2003;158:951–62. [PubMed: 14607803]
5. American Association for Public Opinion Research. Standard definitions: Final dispositions of case codes and outcome rates for surveys. 3rd edition. AAPOR; Lenexa, Kansas: 2004.
6. Sudman, S.; Bradburn, N.; Schwarz, N. Thinking about answers: The application of cognitive process to survey methodology. Jossey-Bass Publishers; San Francisco: 1996.
7. Forsyth, B.; Lessler, J. Cognitive laboratory methods: A taxonomy. In: Biemer, P.; Groves, R.; Lyberg, L.; Mathiowetz, N.; Sudman, S., editors. Measurement errors in surveys. Wiley; New York: 1991. p. 393-438.
8. Bussey-Jones J, Henderson GE, Garrett JM, Maloney M, Corbie-Smith GM. Asking the right questions: views on genetic variation research among participants in a colorectal cancer genetic epidemiology study. J Gen Intern Med 2007;22(Suppl 1):44.
9. National Human Genome Research Institute. Community Genetics Forum. [cited 2007 June 5]. [modified 2007 Oct 27]; [about 3 screens]. Available from: http://www.genome.gov/19518473

10. National Human Genome Research Institute. A Community Forum on Genetics: DNA, health, and social justice. [cited 2007 June 5]. [modified 2007 Oct 2]; [about 3 screens]. Available from: http://www.genome.gov/14514655

11. National Human Genome Research Institute. A Community Genetics Forum 2006: Finding the genome. [cited 2007 June 5]. [modified 2007 Aug 20]; [about 3 screens]. Available from: http://www.genome.gov/19517845

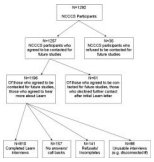

Figure 1.
Recruitment into LeARN Study

Table 1

Characteristics of potential and actual study participants recruited into the LeARN study

	Participants agreeing to be contacted N[1]	Eligible participants N[2]	Participants completing LeARN interviews N (%)[3]	P value
Overall	1196	1108	810 (73%)	
Race				
Black	277	240	155 (65%)	0.001
White	919	868	655 (75%)	
Disease Status				
Case	554	497	369 (74%)	n.s.
Control	642	611	441 (72%)	
Age				
40–49	143	131	91 (69%)	0.022
50–59	324	308	209 (68%)	
60–69	377	341	254 (74%)	
70–79	352	328	256 (78%)	
Education				
< High School	228	204	140 (69%)	n.s.
High School	323	303	225 (74%)	
Some College	292	273	197 (72%)	
College +	334	314	237 (75%)	

[1] Includes the NCCCS participants who agreed to further contact after the initial LeARN letter

[2] Includes eligible NCCCS participants with usable information

[3] Includes completed LeARN interviews; percent calculated from the number of participants completing interviews divided by the number of eligible NCCCS participants