Design of a Family Study Among High-Risk Caribbean Hispanics: The Northern Manhattan Family Study

Ralph L. Sacco, MD, MS; Edison A. Sabala, BA; Tanja Rundek, MD, PhD; Suh-Hang Hank Juo, MD, PhD; Jinaping Sam Huang, MS; Marco DiTullio, MD; Shunichi Homma, MD; Katihurka Almonte, MD; Carlos García Lithgow, MD; Bernadette Boden-Albala, MPH, DrPH

Background

Stroke continues to kill disproportionately more Blacks and Hispanics than Whites in the United States. Racial/ethnic variations in the incidence of stroke and prevalence of stroke risk factors are probably explained by both genetic and environmental influences. Family studies can help identify genetic predisposition to stroke and potential stroke precursors. Few studies have evaluated the heritability of these stroke risk factors among non-White populations, and none have focused on Caribbean Hispanic populations. The aim of the Northern Manhattan Family Study (NOMAFS) is to investigate the gene-environment interaction of stroke risk factors among Caribbean Hispanics. The unique recruitment and methodologic approaches used in this study are relevant to the design and conduct of research in non-White populations. The aim of the NOMAFS is to describe NOMAFS and report enrollment and characteristics of the participants. The NOMAFS will provide a data resource for the exploration of the genetic determinants of highly heritable stroke precursor phenotypes that are less complex than the stroke phenotype. Understanding the gene environment interaction is the critical next step toward the development of new and unique approaches to disease prevention and interventions. (Ethn Dis. 2007;17:351–357)

Key Words: Stroke, Hispanic, Genetics

From the Department of Neurology (RS, ES, TR, JH, BB), Department of Epidemiology (RS, SJ), Department of Sociomedical Science (BB), Department of Medicine, Division of Cardiology (MD, SH), Columbia University, New York, NY, USA; Graduate Institute of Medical Genetics, Kaohsiung Medical University (SJ), Department of Neurology, Chang Gung Memorial Hospital (SJ), Taipei, Taiwan; Department of Medical Research, Mackay Memorial Hospital, Taipei, Taiwan (SJ); and Clinicas Corazones Unidos, (CGL, KA), Santo Domingo, Dominican Republic.

Address correspondence and reprint requests to: Bernadette Boden-Albala, MPH, DrPH; Columbia University Neurological Institute; 710 West 168th Street; New York, NY 10032; 212-305-0501; 212-305-1658 (fax); bb87@columbia.edu

Methods

Enrollment of the Caribbean Family Cohort

The Northern Manhattan Family Study (NOMAFS) cohort was derived from the 1727 Caribbean Hispanic probands already enrolled in the multi-ethnic, community-based Northern Manhattan Study (NOMAS). The Northern Manhattan Family Study (NOMAS) was assembled between 1993 and 2001 from a population-based random sample by using random digit dialing methods. At baseline, the cohort was characterized as: 1) White, Black, or...
Hispanic residents of Northern Manhattan; 2) randomly derived from a household with a telephone; 3) age $\geq 40$ years; and 4) no baseline history of stroke.

**Eligibility**

High-risk Caribbean Hispanic probands were identified from NOMAS according to the following criteria: 1) reporting a sibling with a history of myocardial infarction or stroke; or 2) having two of three quantitative risk phenotypes (maximal carotid plaque thickness, left ventricle mass or homocysteine level) $\geq 75$th percentile in NOMAS. Families of the high-risk probands were considered eligible if the proband could provide a family history and had at least three primary relatives willing to participate. Children $< 18$ years old were not eligible because of the low prevalence of risk factors.

We defined inclusion in terms of the feasibility for enrollment as living within a geographically favorable range to enable complete, in-person evaluation. Our commitment to recruit large Caribbean Hispanic families included enrollment in the New York metropolitan area and the Dominican Republic. We chose the Dominican Republic as our second enrollment site because most of our Hispanic cohort self-identified as Dominican, and we were most likely to enroll first-degree relatives there. However, our study continues to recruit other Caribbean Hispanic subpopulations. For this manuscript, the characteristics of our population include families enrolled in the New York metropolitan area. Family members living outside the New York area or the Dominican Republic were excluded unless they committed to visit New York at least once during the study.

**Recruitment**

Selected probands were reached primarily by phone and were asked to make the first contact with their family members. Probands were asked to refer relatives who lived within the New York area or the Dominican Republic and were offered the option of calling their relatives from our facilities or using a calling card that we provided to defray any charges they might incur. After initial family contact to obtain approval, we followed up with relatives to solicit participation. Subjects who could not be contacted by phone were sent letters requesting that they contact our office. Bilingual interviews were conducted by using structured questionnaires.

In-person appointments were scheduled. A letter confirming the appointment was sent reiterating the benefits of the study. Reminder phone calls were made the day before each visit. Efforts to encourage participation included round-trip transportation, financial compensation, breakfast, flexible scheduling including evening, Saturday appointments, group appointments for families who wished to come together, and home visits. The study was approved by the Columbia University Medical Center Institutional Review Board and the National Bioethics Committee in the Dominican Republic.

Participants enrolled in the Dominican Republic were phoned before the field team arrived and contacted again by phone the day before their appointment at Corazones Unidos. All travel expenses were reimbursed. Unlike families enrolled in New York, most families in the Dominican Republic arrived as a group.

Subjects were enrolled during a one-day, four-hour visit. Written informed consent was obtained from all participants. Data were obtained directly from the subjects, and interviews were conducted to assess demographics, social resources, and medical history including medications, vascular risk factors, family history, dietary intake.

**Measurements**

Primary phenotypes for this study included homocysteine levels, carotid IMT, CD, and LVM. All measurements were taken at baseline only. Interviews were conducted in English or Spanish. Our baseline data collection is summarized in Table 1. Standardized questions were adapted from the Behavioral Risk Factor Surveillance System by the Centers for Disease Control and Prevention regarding the following conditions: hypertension, diabetes, hypercholesterolemia, peripheral vascular disease, transient ischemic attack, cigarette smoking, alcohol use, physical activity, and cardiac conditions such as myocardial infarction, coronary artery disease, angina, congestive heart failure, atrial fibrillation, other arrhythmias, and valvular heart disease. Questions regarding stroke and transient ischemic attack were adapted from the National Institute for Neurological Disorders and Stroke Data Bank and Asymptomatic Carotid Artery Stenosis Study surveillance instruments. Physical activity, alcohol use, family history, and other conventional stroke risk factors were measured as described previously.

Blood pressure was measured with a calibrated standard aneroid sphygmomanometer (Omron; Vernon Hills, Illinois); height and weight were measured with calibrated scales; and hip and waist measurements used standard protocols. Skinfold was measured in the right triceps and abdomen, and the mean of three measurements was calculated. Dietary intake was obtained from the Brief Food Frequency Survey developed at the National Cancer Institute; this questionnaire is both reliable and valid. It has been validated in Hispanic populations and covers foods, dietary habits, nutritional supplements, and specific traditional foods (eg, plantains, mango, rice).

Homocysteine, cystathionine, methylnalonic acid, and methylcitrate levels were analyzed in serum samples at the University of Colorado’s metabolite laboratory by using gas chromatography-mass spectrophotometry. Plasma levels of cholesterol, high-density
Table 1. Baseline data in NOMFAS

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Instrument</th>
<th>Data collected</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td></td>
<td>Gender, age, race-ethnicity</td>
</tr>
<tr>
<td>Social resources</td>
<td></td>
<td>Education, insurance, acculturation, social isolation</td>
</tr>
<tr>
<td>Medical</td>
<td>CDC Behavioral Risk Factor Surveillance</td>
<td>Vascular risk factors, cognition</td>
</tr>
<tr>
<td></td>
<td>NCI Block Diet Survey</td>
<td>Diet</td>
</tr>
<tr>
<td></td>
<td>Family pedigree</td>
<td>Family</td>
</tr>
<tr>
<td>Physical measures</td>
<td>Blood pressure</td>
<td>Blood pressure</td>
</tr>
<tr>
<td></td>
<td>Anthropometrics</td>
<td>Skinfold, height, weight</td>
</tr>
<tr>
<td>Bloods</td>
<td>Buffy coat collection</td>
<td>DNA</td>
</tr>
<tr>
<td></td>
<td>Homocysteine</td>
<td>Homocysteine and metabolites</td>
</tr>
<tr>
<td></td>
<td>Lipid panel</td>
<td>HDL, LDL</td>
</tr>
<tr>
<td></td>
<td>Chemistry</td>
<td>Blood glucose</td>
</tr>
<tr>
<td>Imaging</td>
<td>Carotid doppler</td>
<td>Carotid IMT</td>
</tr>
<tr>
<td></td>
<td>Echocardiogram</td>
<td>Carotid distensibility</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Left ventricular hypertrophy</td>
</tr>
</tbody>
</table>

NOMFAS=Northern Manhattan Family Study.

Lipoproteins, and triglycerides were determined using standardized enzymatic procedures (Boehringer, Mannheim, Germany). DNA was extracted from whole blood by using a DNA purification kit (Geantra Systems, Minneapolis, Minn). DNA samples were sent to the Center for Inherited Disease Research for genotyping. Genome-wide linkage analyses were performed at Columbia University to identify potential loci linked to vascular risk factors of interest.

**Carotid Ultrasound and Echocardiography Studies**

Carotid IMT is an excellent, reproducible, marker of subclinical atherosclerosis and was assessed by high-resolution B-mode carotid ultrasound according to a standardized scanning and reading protocol. Measurements of IMT were performed offline by using the automatic computerized edge tracking system. Carotid IMT was calculated as a composite measure of 12 sites.

Carotid distensibility (CD) has been introduced as a new measure of subclinical disease and a risk factor for cardiovascular disease and stroke. Carotid distensibility (CD) was performed in the 10-mm segment of the right common carotid artery (CCA) below the origin of the carotid bulb according to a standardized ultrasound protocol.

Brachial artery blood pressure measurements were taken with a semi-automated oscillometric blood pressure recorder (Dinamap Pro 100, Criticon, LLC, Tampa, Fl) twice, before and after each ultrasound exam, and averaged. The diameters for the right CCA were measured from 5 B/M-mode registrations and averaged.

Left ventricular mass (LVM) was assessed by two-dimensional transthoracic echocardiography. Standard echocardiography, including color-Doppler flow study was performed with patients in a left lateral decubitus position. Left ventricular mass (LVM) was calculated according to the modified American Society of Echocardiographers formula. The LVM was then divided by the body surface area to obtain an index used in the assessment of LVM phenotype.

The results of blood analysis, echocardiogram, and carotid Doppler ultrasound were mailed to each participant to continue interest in participation and alert them to any major findings. Significant abnormal findings were referred for medical followup and were monitored closely by the research team according to a predetermined safety protocol. A safety plan was in place to detect elevated blood pressure. Dangerously elevated blood pressure was triaged to the emergency room, and all other blood pressure elevations were discussed with the patients’ physicians.

**Statistical Methods**

Genetic analyses were performed to estimate the heritability of each phenotype of interest (homocysteine, carotid IMT, distensibility, and LVH), and genome-wide linkage as well as family-based association analyses were conducted to search for susceptibility loci. Our power estimate was based on our projected sample size of 1455 family members, and this estimate hinges on the heritability of the phenotypes of interest. For example, if the true heritability of the phenotype is 30%, our sample size will give us 92% power to estimate the heritability in the range of .22 to .38.

Simulation studies were performed to evaluate the power to detect linkage on the basis of the same family data (1502 subjects in 114 families). We considered the configurations in which the total heritability of the phenotype ranged from 30%–50%. We calculated the power to detect linkage in our sample assuming a single quantitative trait loci accounting for 20%, 40%, and 60% of the total heritability. For example, the simulation results indicate that for a phenotype with 50% heritability and a quantitative trait loci accounting for 40% of the heritability, the power to detect linkage at the significant level is almost 90%. These simulation studies suggest that the family data will provide excellent power to detect linkage.

**Results**

**Characteristics of our Caribbean Hispanic Population**

Approximately 38% (n=662) of the 1727 Caribbean Hispanics in NOMAS met high-risk criteria. We contacted 406 probands and found 361 eligible families; 27 families were too small, and 18 probands were cognitively impaired.

Ethnicity & Disease, Volume 17, Spring 2007
and unable to provide a reliable family history. Of eligible families, 200 were interested; another 62 families refused, 12 have not been contacted to date, and 87 were ineligible because of scattered geographic distribution.

To date, 87 families with a total of 670 subjects over the age of 18 have had a complete baseline assessment in New York. We have also enrolled 70 subjects from these families in the Dominican Republic (data not shown). No one has refused to participate because of issues pertaining to collection of genetic material. The mean age of the cohort is 48 years (±18 years). Among those enrolled, most (81%) were born in the Dominican Republic but have lived in New York >5 years (Table 2). Almost 51% did not complete high school. Our Caribbean Hispanic families are large, with a mean family size to date of 7.7 (±6.8) participants, and 59 families have at least five members.

The distribution of vascular risk factors is shown in Table 3. We report a high prevalence of hypertension (42%), diabetes (13%), elevated low-density lipoprotein, and high body mass index in our families. Distribution of vascular phenotypes is shown in Table 4.

DISCUSSION

This manuscript is one of the first to describe the methodologic issues involved with the recruitment, enrollment, and sampling of genetic material among Caribbean Hispanics living in and outside of the United States. Few epidemiologic studies have determined the heritability of stroke risk factors in Caribbean Hispanics. The advantages of the family study method include large extended pedigrees, known high vascular risk, sharing of environmental confounders, and ability to enroll US and Dominican Republic residents.

A number of family studies have examined the genetics of cardiovascular disease, primarily among White populations. Little data exist among Caribbean Hispanics. The National Heart, Lung, and Blood Institute (NHLBI) family study recruited from three epidemiologic studies: the Framingham Heart Study, the Atherosclerosis Risk in Communities Study, and the Utah Health Family Tree Study. The NHLBI population was primarily White with some data collection on African American families. The Detroit Project investigates the genetics of blood pressure among African Americans and Whites, while the Honolulu Heart Study examined cardiovascular disease and heritability among Japanese American families. The San Antonio Family Heart Study is a population-based study investigating the genetics of heart disease in Mexican Americans from San Antonio and is composed of 42 extended families (1431 subjects).²³ The San Antonio Family Heart Study is a population-based study investigating the genetics of heart disease in Mexican Americans from San Antonio and is composed of 42 extended families (1431 subjects).

Our success in the enrollment of an immigrant sample from urban and rural areas provides evidence that familial aggregation studies can be conducted in vulnerable minority populations. Studies have suggested that minority populations may be more reluctant to participate in studies because they mistrust the healthcare system. We suggest
that probands’ introducing the study to their family members may increase trust. Additionally, careful wording in the consent form may decrease confusion regarding the collection of genetic material. The use of Spanish-speaking interviewers who share cultural identification with the study population also improved our enrollment rate. Our ability to be flexible in terms of scheduling has also contributed to our successful recruitment.

Critical to the discernment of genetic and environmental factors for vascular disease is the rigor by which the phenotype is defined. Further, key to future genetic studies is the identification of phenotypes with high heritability. We have systematically measured phenotypes among the members of our high-risk Caribbean Hispanic families to determine heritability of putative stroke risk factors. Our standardized, validated protocols have helped reduce concern over measurement variability and precision.

Although the focus of this study is the genetic contribution to subclinical vascular risk factors, we collected extensive data on nativity, education, and other social determinants of health. We believe that controlling for proper and adequate environmental factors is critical to the determination of heritability. Factors including acculturation will be examined within the heritability model.

Since the probands of this family study were selected from the NOMAS study, we will have available data for association gene mapping to replicate and identify the susceptibility locus once we complete linkage analysis.

Our cohorts provide data obtained by the same protocols for both association and linkage mapping. Moreover, we have the ability to cross-validate genetic findings by using association study designs in the parent NOMAS studies. This combined approach has been considered an important method to detect susceptibility genes for a particular disease.

Limitations of the study include incomplete ascertainment of families due to geographical distribution. However, unlike other types of epidemiological studies, linkage analysis can effectively compensate for incomplete families, especially when family size is large. Other strengths include the efficiency of integrating a family study in the setting of NOMAS and the strengths of capitalizing on existing collaborations that will enhance the feasibility of examining family members.

The NOMAFS will provide a data resource for the exploration of the genetic determinants of highly heritable stroke precursor phenotypes that are less complex than the stroke phenotype. Understanding the gene and environment interactions is the critical next step toward the development of new and unique approaches to disease prevention and interventions.

**ACKNOWLEDGMENTS**

The authors wish to thank Kristen Coates, BS; Tania E. Corporan, MD; Roman Gomez, MD; Susanna Silverman, AB; and Meghan Spyes, BS for assistance in this manuscript. They also wish to thank Caroline Torres, MD; Gissette Reyes, MD; Marieta Herrara, MD; Norbelina Disla, BS; Lidia Infante, MD; Hsiu-Fen Lin, MD; Rui Liu, MD; Naeun Park, MS; Romel Ramas, MD; and Jonathan Saniel, MD, for their efforts. They gratefully acknowledge the collaboration of Luis Cuello Mainardi, MD from Clinicas Corazones Unidos, Dominican Republic. We acknowledge the National Institute of Neurological Disorders and Stroke, 1R01 NS 40807-04 for research funding.

---

**Table 3. Distribution of vascular risk factors in Northern Manhattan Family Study**

<table>
<thead>
<tr>
<th></th>
<th>Males (N =250)</th>
<th>Females (N =420)</th>
<th>Total (N =670)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
<td>n</td>
</tr>
<tr>
<td>Hypertension</td>
<td>94</td>
<td>38</td>
<td>189</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>38</td>
<td>15</td>
<td>50</td>
</tr>
<tr>
<td>Coronary artery disease</td>
<td>26</td>
<td>10</td>
<td>46</td>
</tr>
<tr>
<td>No physical activity</td>
<td>124</td>
<td>50</td>
<td>232</td>
</tr>
<tr>
<td>Moderate alcohol intake (&gt;0 to &lt;2 drinks per day)</td>
<td>142</td>
<td>57</td>
<td>140</td>
</tr>
<tr>
<td>Ever smoked</td>
<td>113</td>
<td>45</td>
<td>144</td>
</tr>
<tr>
<td>Current smoker</td>
<td>38</td>
<td>15</td>
<td>56</td>
</tr>
<tr>
<td>Body mass index</td>
<td>Mean ± SD</td>
<td>28.6 ± 5.4</td>
<td>29.6 ± 6.1</td>
</tr>
<tr>
<td>Waist:hip ratio</td>
<td>.93 ± .08</td>
<td>.92 ± .62</td>
<td>.95 ± .49</td>
</tr>
<tr>
<td>Skinfold thickness (mm)</td>
<td>22.3 ± 11.3</td>
<td>32.1 ± 10.5*</td>
<td>28.4 ± 11.8</td>
</tr>
<tr>
<td>Fasting glucose (mg/dL)</td>
<td>90.9 ± 43.9</td>
<td>85.9 ± 27.3</td>
<td>87.8 ± 34.4</td>
</tr>
<tr>
<td>Total cholesterol (mg/dL)</td>
<td>186.4 ± 43.7</td>
<td>186.9 ± 37.8</td>
<td>186.7 ± 40.0</td>
</tr>
<tr>
<td>HDL (mg/dL)</td>
<td>42.7 ± 11.6</td>
<td>51.8 ± 14.4*</td>
<td>48.4 ± 14.1</td>
</tr>
<tr>
<td>LDL (mg/dL)</td>
<td>115.2 ± 35.9</td>
<td>111.6 ± 35.6</td>
<td>112.9 ± 40.0</td>
</tr>
<tr>
<td>Triglycerides (mg/dL)</td>
<td>144.1 ± 127.7</td>
<td>116.8 ± 72.0*</td>
<td>126.8 ± 97.1</td>
</tr>
</tbody>
</table>

* P < .05 between males and females.

**Table 4. Distribution of phenotypes in Northern Manhattan Family Study**

<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>Mean ± SD</td>
<td>n</td>
</tr>
<tr>
<td>Intima-media thickness</td>
<td>228</td>
<td>.63 ± .07</td>
<td>377</td>
</tr>
<tr>
<td>Carotid distensibility</td>
<td>228</td>
<td>1.7 ± 1.04</td>
<td>375</td>
</tr>
<tr>
<td>Left ventricular mass</td>
<td>140</td>
<td>106.9 ± 29.82</td>
<td>272</td>
</tr>
<tr>
<td>Homocysteine(μmol/L)</td>
<td>106</td>
<td>9.56 ± 4.71</td>
<td>222</td>
</tr>
</tbody>
</table>
REFERENCES


AUTHOR CONTRIBUTIONS

Design concept of study: Sacco, Rundek, Hank Juo, Homma, Boden-Albala

Acquisition of data: Sacco, Sabala, Rundek, Huang, DiTullio, Homma, Lithgow, Almonte, Boden-Albala

Data analysis and interpretation: Sacco, Hank Juo, Huang, Homma, Boden-Albala

Manuscript draft: Sacco, Sabala, Rundek, Hank Juo, Huang, DiTullio, Lithgow, Almonte, Boden-Albala

Statistical expertise: Rundek, Hank Juo, Huang, Homma, Boden-Albala

Acquisition of funding: Sacco, DiTullio, Boden-Albala

Administrative, technical, or material assistance: Sacco, Sabala, Rundek, Hank Juo, DiTullio, Huang, Homma, Lithgow, Almonte

Supervision: Sacco, Sabala, Rundek