FAMILY STUDY

Cardiovascular Disease in American Indians
(Phase V)

Operations Manual - Volume One

GENERAL DESCRIPTION AND SURVEILLANCE PROCEDURES

THE NATIONAL HEART, LUNG AND BLOOD INSTITUTE
OF THE NATIONAL INSTITUTES OF HEALTH
THE STRONG HEART STUDY

Cardiovascular Disease in American Indians
(Phase V)

Operations Manual

Volume One

GENERAL DESCRIPTION

July 01, 2006

For copies, please contact

Strong Heart Study Coordinating Center
Center for American Indian Health Research
College of Public Health

P.O. Box 26901
Oklahoma City, OK  73190
ACKNOWLEDGEMENTS

The members of the Steering Committee of the Strong Heart Study would like to acknowledge that this manual and the extension of this study would not have been possible without the contributions and support of a large number of individuals and organizations. First, the Steering Committee wishes to express its appreciation to the thirteen Tribal Communities, whose approval and support have been so willingly offered and whose members are participants in the Strong Heart Study. For the preparation of the manual, we would like to acknowledge contributions and in some cases interview forms or instruction sheets from the following studies: Framingham, CARDIA, ARIC (Atherosclerosis Risk in Communities), CHS (Cardiovascular Health Study), The Longitudinal Diabetes Study of the Phoenix Epidemiology and Clinical Research Branch, National Institute of Diabetes, Digestive and Kidney Diseases, National Institutes of Health and the Diabetic Renal Disease Study. We wish to thank the Indian Health Service for providing us with access to medical records and reports, which have facilitated the planning and execution of the study. This SHS Phase V manual was compiled through the tireless efforts of Ms. JoAnne Whalen, Dr. Carl Schaefer, Ms. Lee Keesee, Dr. Jeunliang Yeh, Dr. Fawn Yeh, Ms. Karen Kimbley, Ms. Martha Stoddart, Ms. Debra Gates, Mr. Yiming Wang, and other staff at the Coordinating Center, and received careful oversight and many revisions from the SHS Steering Committee (especially Drs. Elisa Lee, Barbara Howard, Lyle Best, Richard Devereux, Jean MacCluer, Richard Fabsitz, Thomas Welty, Linda Cowan, and Jason Umans and the Study Coordinators (Dr. Tauerqer Ali, Ms. Marcia O’Leary, and Dr. Marie Russell). Special thanks are owed to Dr. Jan Beals, Dr. Thomas Welty and the Psychosocial Committee for developing a set of psychosocial instruments for use in Phase V and for developing the related Psychosocial Volume of the Manual (Volume 8). Additionally, we thank Lifescan, Inc. for supplying the glucose meters and supplies for the 3 Field Canters in Phase V. Finally, we wish to thank the staff of the Genetic Epidemiology Scientific Research Group, Epidemiology and Biometry Program, Division of Epidemiology and Clinical Applications Branch of the National Heart, Lung, and Blood Institute for making this study possible.
## VOLUME I

### GENERAL DESCRIPTION

Table of Contents

1. **GENERAL DESCRIPTION AND STUDY MANAGEMENT** ................................................................. 1

1.1 **BACKGROUND** ......................................................................................................................... 1

1.1.1 General ................................................................................................................................. 1

1.1.2 Scientific Background ............................................................................................................ 1

1.2 **RESEARCH OBJECTIVES** ........................................................................................................ 9

1.3 **STUDY DESIGN** ..................................................................................................................... 12

1.3.1 Surveillance .......................................................................................................................... 13

1.3.2 Clinical Examination ............................................................................................................ 13

1.4 **STUDY QUESTIONS** ............................................................................................................... 16

1.5 **STUDY MANAGEMENT** ......................................................................................................... 17

1.5.1 Introduction .......................................................................................................................... 17

1.5.2 Confidentiality of Data .......................................................................................................... 17

1.5.3 Communications ................................................................................................................... 17

1.6 **DATA MANAGEMENT AND STATISTICAL ANALYSIS** ......................................................... 19

1.6.1 Development and Production of Study Manual and Data Collection Forms ......................... 19

1.6.2 Procedures for data entry and verification for completeness .................................................. 22

1.6.3 Data Transmission ................................................................................................................ 22

1.6.4 Data Backup .......................................................................................................................... 22

1.6.5 Quality Assurance (QC) Program ......................................................................................... 22

1.6.6 Statistical Analysis and Power Estimates ................................................................................. 26
1.7 PUBLICATION POLICY .................................................................................................................................49
1.7.1 Submission of a Paper Proposal..................................................................................................................49

1.8 ANCILLARY STUDIES POLICY ......................................................................................................................59
1.8.1 General Policy...........................................................................................................................................59
1.8.2 Definition of an Ancillary Study ................................................................................................................59
1.8.3 Requirements for Approval of an Ancillary Study .......................................................................................59
1.8.4 Preparation of Request for Approval of an Ancillary Study ........................................................................60
1.8.5 Review of Ancillary Study Proposals..........................................................................................................61
1.8.6 Amendments of Ancillary Study Proposals................................................................................................61
1.8.7 Yearly Progress Report for Ancillary Study................................................................................................61
1.8.8 Analysis and Publication of Results of Ancillary Studies ..........................................................................61
1.8.9 Feedback of Results of Ancillary Studies to Participants ........................................................................62
1.8.10 Handling of SHS Data and Specimens .....................................................................................................62
1.8.11 Ancillary Studies Using DNA or Other Stored Samples..........................................................................63

Appendix
1 The Strong Heart Study V - - Principal and Co-Investigators .................................................................1-1
2 The Strong Heart Study V - - Organizational Chart ..................................................................................2-1
3 The Strong Heart Study V - - Steering Committee ....................................................................................3-1
4 The Strong Heart Study V - - Subcommittees ..........................................................................................4-1
5 The Strong Heart Study V - - Other Key Personnel .................................................................................5-1
6 The Strong Heart Study V - - Consultants ...............................................................................................6-1
7 The Strong Heart Study V - - Confidentiality Pledge .................................................................................7-1
8 The Strong Heart Study V - - Publications and Presentations Committee (P&P) Forms ....................8-1
CHAPTER ONE

GENERAL DESCRIPTION AND STUDY MANAGEMENT

1.1 BACKGROUND

1.1.1 General

A review of existing data by the Subcommittee on Cardiovascular and Cerebrovascular Disease of the Secretary of Health and Human Service's Task Force on Black and Minority Health concluded that information on cardiovascular disease (CVD) in American Indians (AI) is inadequate and strongly recommended epidemiologic studies of this problem. The Strong Heart Study (SHS) was designed to respond to this recommendation.

1.1.2 Scientific Background

A. Rationale for studying heart disease in American Indians

CVD is the leading cause of death of American Indians. Approximately 30% of Indian deaths for all ages are associated with diseases of the heart. The number of deaths associated with heart disease and stroke among Indians aged 45 years and older exceeds the next three leading causes of death (cancer, diabetes and unintentional injuries) combined. The decline in age-adjusted CVD death rates experienced by the general population in recent decades is not being observed in the Indian population. Among most Indian groups, CVD morbidity and mortality are increasing. SHS offers by far the best, and perhaps the only, prospect of understanding why this increase in CVD is occurring and, more importantly, what can be done to reverse the trend.

An extremely practical reason exists for continued study of CVD among American Indians: - little systematic information about management of heart disease, stroke and hypertension among Indians is available to guide health care workers in identifying effective treatment and intervention programs. The SHS provides by far the greatest body of information to help guide therapeutic and preventive measures for those providing CVD care to AI. As management and prevention of CVD become more complex, the need for epidemiologic, pathophysiologic and genetic information becomes greater. Without information that can only be obtained from the SHS, the extension of such advances to AI will be problematic, subject to potential error, and require years of study. It also must be emphasized that the SHS population serves as the model for examining diabetes-related CVD; our results have been applied to other populations in the US and throughout the world.

B. Description of Strong Heart Study, Phases I - IV
The SHS includes cohort and family/genetic studies of CVD among AI men and women. SHS has been supported by the National Heart, Lung, and Blood Institute (NHLBI) from October 1, 1988 (Phases I-IV), and funding has now been continued for Phase V of SHS, which is a 2nd exam of all of the family members enrolled in the family pilot study of Phase III and/or the full-blown family study in Phase IV. SHS is the largest longitudinal study and the largest study of extended families ever undertaken among American Indians. The study population includes members of 13 communities in three geographical areas.

The SHS has two major components. The cohort study is a comprehensive investigation of CVD morbidity and mortality and associated risk factors. It employs standardized methodology for CVD epidemiology, and is designed to estimate CVD mortality and morbidity and prevalence of known and suspected CVD risk factors and target organ damage among American Indians and to assess the significance of these risk factors in a longitudinal analysis. It contains the largest cohort of individuals with diabetes under continuous CVD surveillance in the U.S.

During the Phase I baseline examination, conducted between 1989 and 1991, 4,549 tribal members (62% of the total population aged 45-74 years) were examined. A second examination (Phase II), involving 89% of surviving original cohort members, was conducted between 1993 and 1995. A third and final exam in 1998-1999 (Phase III) involved 88% of the surviving cohort (3,197 participants).

Continuous surveillance of the cohort has been in effect since the conclusion of the first examination. Information on each member is obtained yearly, and all deaths and all nonfatal CVD events are classified by standardized criteria.

The second major component is a genetic study (the Strong Heart Family Study, SHFS) using linkage analyses to localize genes influencing CVD and its risk factors. This study is noted for the large size of its families and extensive evaluation of cardiovascular risk factors and cardiovascular phenotypes by carotid and cardiac ultrasound measurements. The SHFS was initiated in Phase III as a (feasibility study), expanded to a fully-powered genetic study in Phase IV, and continues as a second exam of all of the surviving family members in Phase V. In Phase III, between 9 and 12 extended families (more than 300 members at least 18 years of age) were recruited and examined in each of the three field centers beginning in 1997. The exam included all elements of the Phases I – III exams except the echocardiogram, gall bladder sonogram, and pulmonary function tests. A 10-centimorgan (cM) map has been constructed and linkage analysis is being performed to assess inheritance of CVD risk factors. In Phase IV, an additional 18 to 25 extended families (a total of about 900 members at least 15 years of age) were recruited from each of the field centers from 2001 – 2003. This effort provided a total of 3,797 individuals from 94 families, of whom 825 are Phase III participants re-examined in Phase IV. In Phase IV, both cardiac and carotid ultrasound exams were done. Major goals for Phase IV were to estimate heritabilities, covariate and household effects, and genetic and environmental correlations for a large set of CVD risk factors and measures of preclinical disease; to generate a 10 cM map of nearly 400 short tandem repeats (STRs) for the 2,700+ Phase IV participants; to screen the phenotypes for linkage; and to begin finer scale mapping to localize quantitative trait loci (QTLs). In Phase V promising signals will be pursued.
High-throughput microsatellite genotyping for a genome scan has been completed in 3,797 SHFS participants (1,240 or more family members from each of the three centers). Work continues to screen the phenotypes for linkage using a variance component approach in full pedigrees; and to begin finer scale mapping with additional STRs to more precisely localize QTLs within targeted chromosomal regions, using SNPs in positional candidate genes for linkage/association analysis to identify genes that are responsible for linkages detected by the initial genome scan.

C. Rationale for Phase V of the Strong Heart Study

One of the most promising routes for expanding knowledge of CVD is exploration of genetic linkage of CVD phenotypes. It is particularly important among Indian populations, because public health policies designed for majority populations are not likely to be not applicable or possibly even to be risk enhancing when applied to AI.

The high prevalence of diabetes in most Indian communities suggests that investigation of diabetes and CVD genetics in SHS may be particularly fruitful. There are several reasons why this might be true. First, possibly unique genetic etiologies may be operative. Second, the more homogeneous genetic background of SHS communities tends to reduce the variation seen in more heterogeneous populations, increasing power to detect genetic influences. Third, the high prevalence and incidence of these conditions enhance the statistical power. Lastly, SHS communities contain large extended families with limited migration.

The goal of the SHFS is detection and mapping of genes influencing variation in risk factors for CVD and related disorders. We successfully recruited and examined more than 1,200 family members in each of the three centers. A 10-centimorgan map is now complete for all of the family members, and we have localized several regions of interest. Our heritability assessments indicate that several CVD risk factors and measures of atherosclerosis and cardiac function have a strong genetic basis, and we have obtained promising preliminary indications of chromosomal regions that may contain genes for CVD risk factors and cardiac function. The sources of genetic variation responsible for these preliminary linkage results require further study, and additional disease-predisposing genes remain to be discovered and mapped. We have power to detect a range of previously unidentified risk factor genes that may be important determinants of cardiovascular health and disease among AI, and the availability of echocardiographic and carotid ultrasound measures provides an innovative approach to understanding genetic mechanisms involved in CVD. While our communities will not allow the transformation and establishment of permanent cell lines from the WBC samples, we have collected adequate amounts of DNA from the cohort at all 3 exams, and from family members at their baseline and repeat exams. Availability of DNA for SHS participants (including both cohort and family study participants) will enable us to analyze positional candidate genes identified by linkage analysis, as well as other candidate genes that may influence CVD risk factors (see below). We will also examine selected candidate genes using a conservative approach due to the large number of conflicting results from association studies, particularly where non-functional
polymorphisms are investigated. Our highly selective candidate approach is complementary to the linkage strategy. Our analyses will examine relationships with major vascular end points particularly where intermediate phenotypes for pathways affected are less well represented in SHS (e.g., the innate immune system). Decreases in cost of genotyping mean that this can be achieved relatively inexpensively.

We will assess a limited number of compelling candidate genes and their relationship to major vascular endpoints without guidance from the linkage study.

1) Mannose binding lectin (MBL). This serum protein opsonizes pathogenic microorganisms by binding mannose moieties on their surface and activating complement via the lectin pathway prior to antibody formation. Decreases in opsonization detected in 5-7% of Caucasians and commonly among other populations result from markedly decreased levels of MBL related to variations of both structural and promoter portions of this gene. Previous reports suggest an association between MBL genotypes and CVD, and our preliminary data showed a high prevalence of variant MBL alleles and their relation to coronary artery disease (CAD) with potentially important public health implications.

2) Interleukin 6 (IL6) -174 C to G. Interleukin 6 stimulates release of acute phase proteins, including fibrinogen and C-reactive protein. A -174 polymorphism in the IL-6 promoter has been described, and in vitro expression supports a functional role. In non-disease states, the C allele is associated with lower concentrations of IL6 relative insulin sensitivity and higher endothelium dependent vasodilatation. The C allele has been associated with reduced carotid intimal medial thickness and coronary heart disease in some but not all studies.

3) Thrombospondin 4 (A387P variant). Thrombospondins are a family of extracellular matrix glycoproteins involved in cell adhesion. A large-scale screen of functional polymorphisms in the GeneQuest study highlighted the A387P variant as having the strongest association with vascular disease (Odds Ratio of myocardial infarction for P allele of 1.89).

4) Lymphotoxin-α (G252A, A804C variants). Variants in the lymphotoxin-α gene were significantly associated with myocardial infarction (OR 1.78) in a large Japanese case control study using genome wide SNP analysis screening.

5) Toll-like receptor-4 (TLR-4). Toll-like receptors, such as TLR-4, respond to microbial lipopolysaccharide (LPS) by activating the NK-kB signaling pathway and induce a wide variety of cytokines and other inflammatory mediators. Many in vitro investigations provide evidence of the influence of TLR-4 receptors on processes related to atherosclerosis. Genotypic variants of TLR-4 are associated with CRP and WBC responses to pulmonary LPS challenge in humans. Various lines of clinical evidence suggest a role for TLR-4 in the pathogenesis of CVD.

6) Peroxisome Proliferator-Activated Receptor (PPARγ) Pro12Ala. PPARγ is an important candidate gene for insulin sensitivity and has been related to vascular disease. In addition to very rare loss of function mutations (PPARγ mutations, digenic mutations of PPARγ and PPP1R3), it is now clear that a common mutation (Pro12Ala) of PPARγ has functional consequences in vitro and relates reliably to development of type 2 diabetes in large populations. Recently the Alanine 12 allele has been associated with protection against incident myocardial
infarction (OR 0.71). As the costs of SNP analyses continue to fall, other promising candidate genes may also be analyzed.

**Rationales for Major Components of Phase V of the Strong Heart Study**

1. **Rationale for Mortality and Morbidity Surveillance of the Original Cohort plus the SHFS Members**

   Cohort surveillance is necessary to identify fatal and non-fatal CVD outcomes. Because we will not examine the original SHS cohort as part of Phase V, surveillance is the only tool for identification of incident CVD events. In 2006, the SHS cohort will range in age from about 62 to 91 years with a mean of 75. These participants constitute a cohort of elders who will have been under repeated observation for about 20 years.

   An advantage of continued follow-up is the ability to study the development and progression of heart disease over time in a population with especially high prevalence rates of diabetes. Results will be applicable to diabetic populations throughout the world. Inclusion of the over 3000 Phase IV SHFS non-cohort participants will provide information on mortality within families at younger ages. This information can be analyzed to study clustering of risk factors and preclinical disease.

   Continued surveillance will provide: (1) a sufficient period of observation following collection of risk factor data to ensure biologically plausible latent periods for vascular disease, (2) improved accuracy in estimating age-specific CVD mortality and morbidity rates and age-specific all-cause mortality rates and increased power to identify risk factors for development of CVD, (3) the opportunity to examine factors related to CVD incidence and mortality in mid-life (early or premature events) vs. those occurring among the elderly (late or non-premature events), and to explore factors associated with survival, (4) a dataset permitting study of the relationship between diabetes and CVD, and (5) information on family mortality patterns including younger individuals.

2. **Rationale for Re-exam of Family Members**

   The re-exam will be conducted approximately 5 to 6 years after the original exam and will permit evaluation of genetic factors that contribute to changes in CVD risk factors. This re-exam will include both carotid and cardiac ultrasound measures, so that the preliminary data on progression can be examined in more detail with a larger number of participants, and among participants as young 15 yrs. Because of the high rates of insulin resistance, obesity, and diabetes among the young people in this population, the re-examination will permit detailed examination of the effects of these disorders on progression of preclinical CVD. Popliteal artery (artery in the leg) ultrasound has been added to provide better measures of peripheral arterial disease (PAD) because of the high rates of PAD observed using other indicators during our cohort exams. Thus, we will be able to compare popliteal and carotid atherosclerosis and their risk factors with emphasis on smoking- and diabetes-related phenotypes. Blood measures will include measures of risk factors that could be expected to change in this period of time – i.e.,
lipoprotein, hemostatic and inflammatory factors plus indices of glycemia – but not ones that would not be expected to change in this timeframe (e.g., Lp(a), apo E genotype).

The exam will include additional biomarkers. We will measure C-Reactive Protein (CRP), which has been shown to predict both CVD and type 2 diabetes mellitus in many populations, and genetic factors are known to strongly influence CRP expression. We will measure leptin because of its role in obesity and because numerous studies, including ours, have implicated a locus on chromosome 2p in determination of leptin levels and/or other obesity-related phenotypes. Leptin represents one of the first adipocyte-derived proteins that has offered clues into the potential endocrinological role of adipose tissue. We will also measure free fatty acids (FFA) because elevated FFAs decrease cellular glucose metabolism and impede insulin-mediated glucose disposal. Fatty acids suppress insulin secretion and are postulated to be mediators of beta cell failure in type 2 diabetes. Elevations of FFAs influence vascular function by causing endothelial damage leading to vasoconstriction, release of inflammatory cytokines and enhanced thrombosis.

In addition to the value for our genetic analyses, the re-exam will allow us to ask a number of questions about changes in risk factors and preclinical disease. Although the family cohort is not a population-based sample, families were recruited from all areas of our communities; and a comparison of the original cohort data with the family study data indicate that the cohort and family data do not differ in any meaningful way. Our preliminary analyses of risk factors for younger family members show alarming rates of obesity and diabetes and high prevalences of key risk factors. At the end of the Phase IV exam (2003), pilot family members were reexamined. Among those family members, 479 cohort members had echocardiograms in SHS Phases II and IV, and 825 cohort members and pilot family study participants had carotid ultrasound studies in Phases III and IV. These data provided repeat cardiac and arterial assessments in these individuals. Preliminary results revealed interesting and disturbing changes in carotid measures of atherosclerosis and in echocardiographic measures of left ventricular hypertrophy (LVH) and valvular disease over an average of only 4 to 8 yrs of follow-up.

Thus, despite the limitations of the relatedness of the sample and its non-random selection, we believe the data obtained in the re-exam will be both informative and trail breaking, pointing the direction for investigators wanting to plan further studies in younger individuals.

3. Rationale for New Biomarkers to be Measured in Stored Specimens

As an established study, SHS maintains a unique reservoir of stored biological samples. During the 15 years the cohort has been followed, we have obtained reliable information about participants who developed or who were free from atherosclerosis. Samples from these individuals can be used to assay new risk factors in relation to clinical or subclinical CVD. In Phase V, we will measure apo CIII, a small protein on the surface of apo-B containing lipoproteins and HDL that plays a major role in metabolism of VLDL via inhibition of lipoprotein lipase. Apo CIII also retards the clearance of VLDL and influences HDL metabolism. Concentrations of apo CIII are associated with CVD independent of triglyceride (TG) concentrations. Apo CIII is elevated in individuals with diabetes, and it potentiates
atherosclerosis. We will perform this test on a subset of the entire cohort using a case-cohort design.

4. Rationale for Benefits to Indian Communities and Clinical/Public Health Practice

Four prime considerations make imperative study of disease processes among American Indians and Alaska Natives (AI/AN): First, support for research is part of the federal government’s obligation to raise the health status of American Indians to the highest possible level. Second, the SHS is essential for estimating as accurately as possible Indian morbidity and mortality, since the IHS mortality surveillance system depends upon state death certificates, which have been shown to be unreliable as an accurate index of Indian death rates. Third, Congress has directed the National Institutes of Health (NIH) and other federal agencies to address health disparities that exist among certain minority populations, including American Indians. Fourth, continuing (and in the case of CVD, growing) disparities in health status of Indians compared to the general population require investigation because information gained from studies among Indian populations will increase understanding of disease mechanisms important to all populations.

The SHS has demonstrated that a multi-center, longitudinal, epidemiologic and pathophysiologic study, including extended family studies, can be carried out in Indian communities and be of high quality, with high participation rates. Numerous logistical problems have been successfully overcome with the cooperation and support of the Indian communities. Cultural differences, which often lead to reluctance to attend exams have been successfully dealt with, and concerns about genetic analyses have been answered. Thus, the SHS serves as a model for epidemiologic studies in minority communities. Perhaps most significantly, the participating communities are eager to continue the study, and both the communities and staff are committed to continued success.

The SHS findings will continue being disseminated at local, regional, state, and national scientific meetings. The SHS has provided many vital services to AI communities by disseminating findings on the major known CVD risk factors and their contributions to CVD. Tangible improvements in community educational and medical infrastructure have also accrued. SHS places considerable emphasis on soliciting input and sharing results with participating Indian communities. Community representatives and physicians are members of the Steering Committee, and the advice of many community members has been implemented throughout all phases of the study.

Participants have received a thorough medical examination emphasizing the cardiovascular system, and extensive laboratory testing. Examinations have included an echocardiogram, carotid artery studies, measurements of ventilatory function, and testing for sensitivity to a variety of skin allergens. These evaluations would be prohibitively costly in the private sector, even if they were readily available to this population. Clinically useful information is regularly shared with participants and their health care providers (with the consent of the participants). Participants receive educational materials and advice on how to reduce their own cardiovascular risk factors. SHS newsletters are distributed twice per year to participants.
When significant medical conditions are detected, participants are referred for medical care according to established SHS referral criteria. Important, unrecognized conditions detected during these exams have included many cases of incident diabetes, hypertension, and occasional cases of congenital heart disease. Such measures are especially important in a rural epidemiologic study whose participants have limited access to health care.

SHS investigators, staff, and coordinating center worked with the NHLBI to compile and publish a data book summarizing SHS findings that are useful to health care providers and community health officials. These have been widely distributed to other Indian communities and programs, and have even been utilized in one rural community of Native Hawaiians. Clearly, SHS data have had a major public health impact. These data serve as the reference for the IHS and other public health agencies in planning programs for health care delivery, education and prevention strategies; SHS data are presented and discussed at all major meetings involving health care providers to AI. Our data have been used for several national reports from the Surgeon General’s office, the American Heart Association and the American Diabetes Association. They formed the basis for the current IHS recommendations on management of dyslipidemias and the ongoing IHS strategic plan for CVD prevention and therapy. SHS data have been used by the Native Elder Research Center at the University of Colorado for the training of AI researchers, and they formed the basis for the PATHWAYS and SANDS Trials now funded by NHLBI. SHS methodology was used as the basis for a new NIH-funded study of CVD in rural Native Hawaiians. The knowledge gained through SHS has allowed tribal leaders to promote healthier lifestyles, especially for the younger generations. The value of utilizing locally derived, community-specific data, rather than regionally derived data for health planning, is readily apparent. SHS has also motivated investigators to design and implement studies to promote healthy dietary and exercise habits among AI children and adults in the SHS communities, for example, the Oklahoma Native American EXPORT Center, funded by the National Center on Minority Health and Health Disparities (NCMHD) of NIH.

Whenever possible, community members have been hired to conduct the study. During the first four phases of SHS, more than 130 community members were employed as part of the SHS clinic staffs. Particularly noteworthy is that more than 100 health profession students have worked in Phases I – IV and that more than half of these were Indian. In several instances, student participation stimulated pursuit of health careers, and some were motivated toward serving Indian communities upon completion of training. Some of these students were research assistants with support from the NIH Minority Supplement program. Others have been mentored by field personnel, who have served as role models for young community members.

The continuation of the SHS in Phase V promises to produce additional benefits in all of these areas. Ultimately, the data will enable the IHS and the Tribes to better allocate limited health care resources and to implement community-specific preventive interventions.
1.2 RESEARCH OBJECTIVES

Specific Aim #1: Conduct genetic studies emphasizing the genome scan approach but also including investigation of carefully selected candidate genes:

a. Linkage analysis of genome scan data to localize genes that contribute to overt and preclinical CVD and CVD risk factors:

We have completed the genotyping of the 3,797 members of extended families in each of the three SHS field centers, of which 2,100 have been used in initial linkage analyses with the goal of localizing genes that influence risk factors for clinical and subclinical CVD, diabetes, and obesity, genes that influence measures of cardiac and vascular function, and genes that influence the progression of these traits over time. After identifying promising chromosomal regions that contain quantitative trait loci (QTLs) using data from a 10-centimorgan map, we will narrow regions of interest by fine mapping. This will involve prioritization of candidate genes within the region of each QTL for extensive resequencing to identify single nucleotide polymorphisms (SNPs), and then SNP typing of all SNPs that we identify in candidate genes. Our analyses will enable us to test whether specific SNPs account for our linkage signals. We also will test whether the SNPs account for population-level association in the SHS cohort.

In a screening analysis of more than 900 members of extended Phase III families and 1,140 of the members of Phase IV families using the 10 centimorgan map, we have identified numerous chromosomal regions containing promising linkage signals. More detailed analyses of some of these signals have shown:

1) Left ventricular mass normalized for height (LOD=5.3) on chromosome 12p.
2) Weight (LOD=5.17) and body mass index (BMI; LOD=5.08) on chromosome 4q.
3) Plasma insulin level (LOD=3.5) and lean body mass (LOD=2.6) on chromosome 2p.
4) Ejection fraction (LOD=3.5) on chromosome 1q.
5) Clusters of insulin resistance syndrome variables identified by factor analysis: a glucose/insulin/obesity factor on chromosome 4 (LOD = 2.3), a dyslipidemia factor on chromosome 12 (LOD = 2.7), and a blood pressure factor on chromosome 1 (LOD = 1.5).
6) LDL-C (LOD=3.7) on chromosome 10p.
7) PAI-1 (LOD = 3.03) on chromosome 11p.

As costs of genetic analyses fall, additional polymorphisms will be considered. Additional signals will be identified in ongoing screening analysis of nearly 3,800 family members. We have budgeted funds to localize and identify the QTLs responsible for three signals. If we are unable to identify the polymorphism responsible for a linkage signal, we will move on to another of the signals on our list. We also will examine whether the linkage signals increase with analysis of additional covariates or additional families or with bivariate analysis, and whether the signals are present primarily in specific subsets of the data (specific centers, specific families).

b. Examination of candidate genes not guided by linkage analysis

The main emphasis of genetic investigation in the SHFS is linkage analysis and subsequent identification of positional candidate genes. In addition, we plan to assess the relation of major
vascular endpoints to a limited number of compelling polymorphisms in candidate genes not guided by the linkage study. This will be designed to test hypotheses regarding intermediate phenotypes and pathways that are less well represented in the SHS (and therefore candidate genes that are difficult to exclude on the basis of the linkage studies), but are of potential importance to vascular disease. Most important in this respect are ‘novel’ vascular risk factors including the innate immune system, cell adhesion, endothelial dysfunction, inflammation, and insulin sensitivity. Investigation will be limited to polymorphisms showing some evidence of functionality. Given reductions in the cost in genotyping, where convincing functional variants exist, this approach allows testing of novel hypotheses and areas of biology more cheaply than additional phenotyping. It also allows examination of key polymorphisms (and contribution to meta-analysis) of data from American Indians - a group generally underrepresented in other large studies. Finally, it allows assessment of associations of these genes with markers of preclinical disease obtained from coronary and carotid ultrasound measurements.

As examples of this approach, we will investigate:

1) **Mannose binding lectin (MBL)** - We have already shown that functional variants are associated with incident CHD (adjusted OR 3.2) and will extend investigation to other variants in the gene.

2) **Interleukin 6 (IL6)** -174 C to G. IL6 is a key stimulator of release of acute phase proteins. The 174 C to G influences IL6 expression in vitro and has been associated with both metabolic and vascular phenotypes.

3) **Thrombospondin 4 (A387P variant)** - Influences cell adhesion and has previously been associated with vascular disease.

4) **Lymphotoxin-α (G252A variant)** - Influences transcription of a key cytokine and, in turn, expression of a range of adhesion molecules and cytokines and also shows significant association with vascular disease.

5) **Toll-like receptor-4 (TLR-4)** - TLR-4 variants influence innate immunity.

6) **Peroxisome proliferator-activated receptor γ (PPARγ) Pro12Ala** - PPARγ is an important candidate gene for insulin sensitivity and has previously been related to vascular disease.

Specific Aim #2: To continue mortality and morbidity surveillance:

Annual CVD mortality and morbidity surveillance of the original SHS cohort members (4,549 original, approximately 3,000 survivors, ages 60 to 89 years) will be continued, and annual mortality surveillance and limited morbidity follow-up of the over 3000 non-cohort SHFS participants will be initiated.

Questions to be addressed are:

a. What risk factors are related to the incidence of CVD across different age strata? What are the age- and gender-adjusted risk factor profiles for premature CVD deaths vs. non-premature CVD deaths?

b. How are incidence rates for various manifestations of CVD (e.g., coronary, cerebral, peripheral) influenced by age, gender, and diabetes status?

c. What are the relations between quantitative measures of systemic atherosclerosis, cardiac hypertrophy, and cardiovascular dysfunction (e.g., LV mass, carotid plaque, carotid wall...
thickness) and CVD incidence and mortality? Are these potential predictors related to other established CVD risk factors, such as, diabetes?

d. What are the incidence rates and major risk factors for specific types of stroke (atherothrombotic, cardioembolic, hemorrhagic, etc) and do they differ by gender or diabetic status?

e. What factors are significantly related to long-term survival? Are there differences in the factors that predict longevity between individuals with and without diabetes at baseline?

f. What are the age-specific CVD incidence and mortality rates and all-cause mortality rates including rates of premature CVD death in American Indians in the three SHS geographic areas? Do these rates differ significantly among the three areas and if they do, what are the explanations for the differences?

g. Have age-specific mortality rates and proportional mortality ratios for CVD and other causes of mortality changed over the 20 years of the SHS follow-up (1989-2009)? Do changes differ between individuals with diabetes and without diabetes? If they do, what are the explanations for the differences?

h. What are the health-adjusted life expectancies (HALE) of American Indians in the three geographic areas? What is the impact of chronic health conditions such as CVD, diabetes, obesity, and renal disease on the HALE?

Specific Aim #3: To re-examine family members:

All SHFS participants will be re-examined so that changes in risk factors can be analyzed and genetic effects on changes can be estimated. The high prevalences of diabetes, obesity, and other risk factors in younger family members warrant follow-up. In addition, we have seen striking changes in cardiac and arterial findings in the relatively few cohort members who were examined in Phase IV, which merit repeat cardiovascular evaluation.

This aim offers several potential scientific opportunities:

a. to assess changes in key CVD risk factors that are the focus of the linkage analysis.

b. to examine changes in intermediate vascular phenotypes, to try to detect genes that are related to these changes, and to assess interactions of other risk factors (adiposity, insulin resistance, hyperglycemia) with these changes. In the first exam we found many young adults with diabetes or metabolic risk factors such as obesity and impaired fasting glucose. Identification of factors promoting atherosclerotic progression in this young at-risk age group is a high priority. Additionally, we will capitalize on members of the original cohort who also are in the Family Study (over 500) to examine long-term changes (for example, left ventricular hypertrophy and carotid intima medial thickness by standardized methods used since the 2nd SHS exam).

c. to examine subclinical atherosclerosis in peripheral arteries using popliteal ultrasound, assess its heritability and relations to risk factors (with focus on diabetes and smoking), and compare these to both carotid intima-media wall thickness (IMT) and ankle-brachial index (ABI).
d. to add measures of additional phenotypes – free fatty acids (FFA), C-reactive protein (CRP) and leptin. All three of these vasoactive substances are important indicators of inflammation and/or are related to obesity, insulin resistance and diabetes, and they will be valuable additions to the linkage analysis.

1.3 STUDY DESIGN

Timeline

<table>
<thead>
<tr>
<th>Activity</th>
<th>Timeline</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SIHS Phase V</strong> (5 yrs, 09/30/05 - 09/29/10)</td>
<td>09/30/05 09/30/06 09/30/07 09/30/08 09/30/09 09/29/10</td>
</tr>
<tr>
<td><strong>Phase I-IV &amp; Surveillance data analyses</strong></td>
<td>X------------------------X</td>
</tr>
<tr>
<td>(5 yrs, 09/30/05 - 09/29/10)</td>
<td>X------------------------X</td>
</tr>
<tr>
<td><strong>Surveillance of Cohort</strong> (4 yrs, 09/30/05 - 09/29/09)</td>
<td>X------------------------X</td>
</tr>
<tr>
<td><strong>Mortality surveillance of members of Family Study</strong> (4 yrs, 08/30/05 - 09/29/09)</td>
<td>X------------------------X</td>
</tr>
<tr>
<td><strong>Train ultrasound staff</strong> (9 mos, 10/01/05 - 06/30/06)</td>
<td>X-----X</td>
</tr>
<tr>
<td><strong>Develop protocol, manual, forms</strong> (9 mos, 10/01/05 - 06/30/06)</td>
<td>X-----X</td>
</tr>
<tr>
<td><strong>Purchase supplies</strong> (9 mos, 10/01/05 - 06/30/06)</td>
<td>X-----X</td>
</tr>
<tr>
<td><strong>Train field staff</strong> (06/06)</td>
<td>X</td>
</tr>
<tr>
<td><strong>Re-exam of Family Members</strong> (3yrs, 07/01/06 - 06/30/09)</td>
<td>X------------------------X</td>
</tr>
<tr>
<td><strong>Fine Mapping</strong> (5 yrs, 09/30/05 - 09/29/10)</td>
<td>X------------------------X</td>
</tr>
<tr>
<td><strong>Candidate Genes</strong> (5 yrs, 09/30/05 - 09/29/10)</td>
<td>X------------------------X</td>
</tr>
<tr>
<td><strong>Analyses and papers on genetic analyses</strong> (5 yrs, 09/30/05 - 09/29/10)</td>
<td>X------------------------X</td>
</tr>
</tbody>
</table>

**Timeline**

09/30/05 09/30/06 09/30/07 09/30/08 09/30/09 09/29/10
1.3.1 Surveillance

Surveillance of the SHS cohort for CVD morbidity and mortality has been ongoing since 1989. Surveillance methods for Phase V of the SHS are the same as those used successfully in Phases II-IV. Mortality surveillance includes annual ascertainment of deaths in survivors of the original cohort and in participants in the SHFS of all ages (i.e., age 15 years and older at the Phase IV examination). Inclusion in the mortality surveillance of SHFS members will add over 3000 new individuals from the SHFS to the annual mortality surveillance cohort and permit continued examination of CVD risk factors in relation to “early” events and comparisons of these factors to those associated with CVD at older ages. Focusing on age-specific risks will allow us to indirectly take into account the “relatedness” of participants from the SHFS, who will have been added to the surveillance cohort. Very few members of our cohort have been lost to follow-up (N=11).

Morbidity surveillance will be done in the original SHS cohort using the same methodology as in Phase IV. For participants in the SHFS, non-fatal CVD events that have occurred since their Phase IV examination will be identified at the time of the Phase V re-examination by means of a physical examination, ECG, and history and through review of the participants’ medical records covering the time period between the Phase IV and Phase V examinations.

Individuals are designated at each center, who are specifically responsible for mortality and morbidity surveillance activities. Surveillance contacts are accomplished using a variety of approaches specific to the SHS populations. These approaches include home visits, monitoring of IHS facility records, telephone calls and mail contacts. All reports of primary endpoints and selected secondary events of interest obtained through surveillance procedures will be validated from medical records. (See Volume Two – Morbidity and Mortality Surveillance)

1.3.2 Clinical Examination

Components of the Clinical Examination. The clinical examination includes a personal interview and a physical examination. Most of the procedures will be the same as those applied to the Phase IV family exam. Procedures are described in brief below, with details presented in Volume III of the manual.

1. Personal Interview

The following questionnaires will be administered:

i. Demographic information: income, residence, marital status, number of household members, and education will be determined.

ii. Health habits: Smoking, alcohol intake.

iii. Medical and reproductive history, including the Rose questionnaire for angina pectoris and intermittent claudication, medication history.
iv. Dietary survey: The Block 98 FFQ was used in Phase IV, modified slightly to add certain foods commonly eaten in our communities. In Phase V, the Block 2005 Food Frequency Questionnaire (FFQ) is being used (again modified slightly to include the additional foods commonly eaten in the communities). Block Dietary Data Systems (BDDS) will enter and analyze the data. Participants will receive individualized nutrient estimates and associated Recommended Dietary Intakes (RDI), providing additional health benefits.

2. Physical Examination

The physical examination includes the following procedures that were used previously. (Anthropometric measurements will be made with participants in loose clothing with shoes off and heavy objects removed from pockets.)

i. Weight: The scale will be balanced on a level and firm surface prior to weighing a participant.

ii. Height: The participant will be measured with special attention to posture, using a standard stadiometer.

iii. Waist circumference: Anthropometric tape will be applied at the level of the navel in the supine position as in previous phases.

iv. Body fat measurement: Resistance and reactances are recorded using an RJL bioelectric impedance meter. Percent body fat will be estimated by the RJL formula based on total body water.

v. Arm circumference: After proper positioning, this will be measured at the midpoint between the acromion and olecranon. The measure will be used to select the proper size blood pressure cuff.

vi. Amputated extremities will be recorded.

vii. Pedal pulses: The presence of posterior tibial and dorsal pedal pulses will be determined in both legs. The sensitivity and specificity of these evaluations will be compared to "gold standards" such as the ABI and ultrasound examination of the popliteal artery.

viii. Ankle edema: Pitting edema will be evaluated anteriorly from the mid-tibia to the ankle. The degree of edema (absent, mild, marked) will be recorded.

ix. Blood pressure measurements: Using the measured arm circumference to select the proper cuff size, the sitting arm blood pressure will be measured three times using a mercury sphygmomanometer. Three blood pressure measurements will be obtained with a 1-minute waiting period between successive measurements. The average of the last two measurements will be used for analysis.

x. Ankle/Brachial Index (ABI): Using a Doppler detector, brachial and ankle systolic pressures will be measured twice.

xi. Resting ECG: A Marquette MAC 1200 will obtain a standard 12 lead ECG. ECGs will be electronically transmitted to Cornell University. Tracings will be Minnesota-coded by computer.
xii. Fasting blood samples will be obtained for measurements of total triglyceride (TG) and cholesterol, LDL and HDL cholesterol, fibrinogen, and glucose, creatinine, insulin, HbA1c, CBC, chemistry profile, FFA, CRP, and leptin.

xiii. Urine will be collected at the beginning of the physical examination for measurement of albumin and creatinine.

xiv. Peripheral sensation will be measured in one foot by 5.07 monofilament, allowing estimation of peripheral sensory neuropathy prevalence and its relationship to diabetes and CVD.

xv. Echocardiography and ultrasound exams of the carotid artery and popliteal arteries.

xvi. Medical records from the Indian Health Service and/or other medical providers will be abstracted to ascertain hospitalizations, outpatient evaluations, or other manifestations of CVD that are SHS endpoints. For family participants, information since the first SHFS examination will be reviewed.

xvii. Physical activity: The Accusplit pedometer will be used to record the number of walking steps. Information will be recorded in the participant’s activity record twice daily. The pedometers were used successfully in Phase IV, and the participants will again be allowed to keep them as a gift.

xviii. Psychosocial measures: The following questionnaires are administered to all Phase V SHS participants: Quality of Life – SF-12; CES-D depression scale; Social Support, Posttraumatic Stress Screening Scale, Generalized Anxiety Screening Scale, Spirituality, and Fatalism. All of these factors have been associated with CVD in other studies.

The clinical examination will last approximately three hours. If possible, all components will be completed in one visit. If for some reason the examination is not completed, every effort will be made to complete the remaining components of the examination within 1 month. The personal interview, consent, and FFQ may be completed up to two weeks prior to the physical examination. For pregnant women, the examination will be conducted no earlier than six weeks after delivery. Lactating women will be included in the study when six weeks or more postpartum.

The participant will arrive at the clinic fasting in the morning. After registration, a study staff member will explain the study and procedures to the participant, answer questions, if any, and administer the consent form. The participant will then be instructed to go to the laboratory for blood drawing and to obtain the urine specimen. The participant is then offered a light snack. The nurse clinician and other staff will then conduct the personal interview, obtain anthropometric measurements, blood pressure, impedance measurement for body fat composition, and ECG measurements. Project staff who have been trained and certified will perform echocardiography and ultrasound exams of the carotid and popliteal arteries. After all the procedures are completed, the participant will receive payment or sign a payment form, will be provided appropriate health educational material to reduce his/her cardiovascular risk, and will be thanked for his/her participation.
1.4 STUDY QUESTIONS

1. What risk factors are related to the incidence of CVD across different age strata? What are the age- and gender-adjusted risk factor profiles for premature CVD deaths vs. non-premature CVD deaths?

2. How are incidence rates for various manifestations of CVD (e.g., coronary, cerebral, peripheral) influenced by age, gender, and diabetes status?

3. What are the relations between quantitative measures of systemic atherosclerosis, cardiac hypertrophy and cardiovascular dysfunction (e.g., LV mass, carotid plaque, carotid wall thickness) and CVD incidence and mortality? Are these potential predictors related to other established risk factors such as diabetes?

4. What are the incidence rates and major risk factors for specific types of stroke (atherothrombotic, cardioembolic, hemorrhagic, etc) and do they differ by gender or diabetes status?

5. What factors are significantly related to long-term survival? Are there differences in the factors that predict longevity between individuals with and without diabetes at baseline?

6. What are the age-specific CVD incidence and mortality rates and all-cause mortality rates including rates of premature CVD death in American Indians in the three SHS geographic areas? Do these rates differ significantly among the three areas and if they do, what are the explanations for the differences?

7. Have age-specific mortality rates and proportional mortality ratios for CVD and other causes of mortality changed over the 20 years of the SHS follow-up (1989-2009)? Do changes differ between individuals with diabetes and without diabetes? If they do, what are the explanations for the differences?

8. What are the health-adjusted life expectancies (HALE) of the American Indians in the three geographic areas? What is the impact of chronic health conditions such as CVD, diabetes, obesity, and renal disease on the HALE?
1.5 STUDY MANAGEMENT

1.5.1 Introduction

The Strong Heart Study Phase V is funded by the National Heart, Lung, and Blood Institute and directed by the Genetic Epidemiology Scientific Research Group, Epidemiology and Biometry Program, Division of Epidemiology and Clinical Applications Branch. The SHS Observational Study Monitoring Board (OSMB) was established for SHS by NHLBI in 1997 and has provided extremely valuable oversight of the project since that time. The Principal and Co-investigators are listed in Appendix 1 below. An organizational chart of the Strong Heart Study Phase V is given in Appendix 2. The operations of the study are directed by the SHS Steering Committee (SC), which includes members from each center and the NHLBI Project Manager (see Appendix 3 for the members of Steering Committee). The SHS OSMB provides guidance and ideas during the annual OSMB meetings when SHS progress and plans are presented; they also review ancillary studies as they are considered by the SC, and make suggestions on potential collaborators. The Oklahoma Center, in addition to being a field center, assumes the responsibility of the Coordinating Center, and the Arizona Center acts as the Core Laboratory. The Cornell University Reading Center under the direction of Dr. Richard Devereux serves as both the ECG Reading Center and the Ultrasound Reading Center. Analysis of the Family Study genetic component is directed by Dr. Jean MacCluer at the Southwest Foundation for Biomedical Research. SHSV Sub-Committee members are listed in Appendix 4. Other key personnel at each center and consultants of the Study are listed in Appendix 5 and Appendix 6, respectively.

1.5.2 Confidentiality of Data

All personnel with access to data collected for the study at each center are required to sign a confidentiality pledge, which states that they understand the sensitive and confidential nature of the data and that divulgence of any information will result in disciplinary action. The pledge will be co-signed by the principal investigator. A sample of the confidentiality pledge is given in Appendix 7.

Completed data forms will be placed in locked file cabinets in offices assigned to the study at each study center. Only authorized staff members have the key to the office and access to the data forms.

Data on computers at the Coordinating Center will be safeguarded by a password, which is known only to authorized personnel.

1.5.3 Communications

1. Newsletter:

The Coordinating Center periodically prepares and distributes a newsletter to facilitate communication among Study staff and with the SHS participants. In general, each edition
includes: (1) reports from the Program Office, the Steering Committee, the Coordinating Center, the Core Laboratory, the Cornell Reading Center (ECG, Carotid Artery and Popliteal Artery Ultrasound, and Echocardiogram), and the Southwest Foundation for Biomedical Research (Genetic Study Center), (2) a description of the facilities and staff of a field center or central agency, (3) general information on data management, (4) information about new ancillary studies, and (5) upcoming events. The newsletter also provides reports on issues such as recruitment and participant follow-up rates, the development and use of new equipment, and preliminary study results and abstracts.

2. Electronic Mail:

E-mail through Internet and FAX continue to be the major electronic mail facilities used by all field centers, the Coordinating Center, Core Laboratory, Cornell Reading Center, Genetic Study Center, and the Program Office. This electronic mail network allows rapid and efficient communication among centers for messages such as announcements, meeting agendas, abstracts for clearance, and acknowledgments of receipt of data.

3. Web Site http://strongheart.ouhsc.edu:

The list of SHS scientific papers, both published and in press, is available and linked to abstracts posted on the National Library of Medicine PubMed website. The Manual of Operations for each of the five phases of SHS is also available, along with a wealth of other information including annotated data collection forms, virtually all of the SHS newsletters in Adobe Acrobat format, and downloadable slide presentations on various aspects of SHS.

4. Field Center Visits:

The Program Office and staff from the Coordinating Center, Cornell Reading Center, Core Laboratory, and Genetic Study Center conduct periodic monitoring visits to field centers as needed to: (1) maintain channels of communication with field center investigators and staff, (2) monitor participant recruitment and surveillance procedures, (3) monitor adherence to the protocol, and (4) provide technical support for activities such as data management and quality control.
1.6 DATA MANAGEMENT AND STATISTICAL ANALYSIS

1.6.1 Development and Production of Study Manual and Data Collection Forms

The Coordinating Center worked closely with the Steering Committee in the development and production of the study manual and data collection forms. A Forms Committee reviewed all forms and made recommendations for revisions, deletions, and additions of forms. The Psychosocial Committee held frequent conference calls and devised a set of psychosocial forms comprised of forms used previously in SHS and elsewhere. The Manual was revised by Steering Committee members, Field Coordinators and CC personnel. Revisions were circulated by email attachments, and further input and improvements were provided during the training sessions held in Oklahoma City (March 14 – 16, 2006). After initiation of the Phase V exams in May 2006, the entire manual was reviewed page by page and modifications were incorporated.

a. **Sources of data**

Data forms for the SHS are generated from a variety of sources.

i. From the three field centers: Clinical examination forms (personal interview, medical history, physical examination, quality of life and other psychosocial forms, machine reading of ECG, and CBC by local clinic labs), Death Certificate Form, and Morbidity Survey Medical Chart Review Form.

ii. From the Penn Medical Lab (Core Lab) at Medstar Research Institute: total triglyceride and cholesterol, LDL and HDL cholesterol, fibrinogen, and glucose, creatinine, insulin, HbA1c, CBC, chemistry profile, FFA, CRP, leptin, and urinary albumin and creatinine.

iii. From the Cornell University Reading Center, cardiologist's ECG reports, computerized Minnesota ECG codes, echocardiography, carotid artery and popliteal artery ultrasound data.


v. From Dr. Maurice Sievers: Mortality study final decision package (Mortality Study Chart Review Form, Final Decision Form, and Informant Interview Form).

vi. From Mortality and Morbidity Review Committees (Mortality or Morbidity Study Chart Review Form, Mortality or Morbidity Final Decision Form, and Mortality Informant Interview Form).

vii. From Dr. Jean MacCluer: genotyping data on all family study members.

viii. From Block Dietary Data Systems: Analyses of the Food Frequency Questionnaire.

b. **Database development**

In SHSV, the Coordinating Center continues to use a distributed data entry system. In Phase V the Coordinating Center used Microsoft (MS) ACCESS 2003 to develop the data entry programs (similar to previous phases) and MS Windows Terminal Services to support real-time data entry (as opposed to batch transmission as used in Phase III) via high-speed Internet connections with state of the art field center computers. Separate files have been created for
each data form; these files and the data files are stored solely on the server(s) at the Coordinating Center. Maintenance of the data programs and files occurs on the server(s) at the Coordinating Center; the field centers transmit the exam data to the Coordinating Center for data cleanup and permanent storage.

The laboratory data and data from special studies are transmitted to the Coordinating Center electronically over the Internet or by sending data-containing media such as diskettes, CDs, and DVDs. The Coordinating Center stores the raw data sent from the specific study centers and converts them into SAS data files for analysis.

c. Procedures for data entry and verification of completeness

Each field center reviews every data form for completeness and accuracy before entering it into the field computer. Details of the data entry process and data management can be found in Volume 7 of this SHS V Operations Manual. The completeness of data entry for each form is checked again by the Coordinating Center. Any incomplete items (missing, questionable, unclear) are recorded, and the corresponding field center is contacted to find out the reason. When these items are completed by the individual center and received by the Coordinating Center, the information is updated in the Coordinating Center’s database. To ensure the data quality, the field centers are required to double-enter 10% of the forms each month (or at least one double entry per transmission). The Coordinating Center checks the disagreement rate for double entry on a monthly basis. If the disagreement is greater than 0.5% in any transmission, that center is asked to re-enter (as second entry) the data of all the forms in that transmission.

The data received from the Core Laboratory via the Internet as ASCII files are directly converted into SAS datasets. Before these data are merged into the permanent data files, various quality and consistency checks are performed.

Uniform data entry forms for all information to be collected have been designed by the Coordinating Center for use by each Study Center. Each study subject has a unique identification number (ID number). Please see the Strong Heart Study Phase I Manual page 12a for the detailed procedure to assign the study ID number. For those original cohort members who participate in Phase V, the original ID number assigned in the Strong Heart Study Phase I will still be used. The ID number will be stamped on every page of all forms at each center. For laboratory specimens, printed labels supplied by the Core Lab are used. For each of the family members enrolled in the family study, Family IDs were assigned during the Phase III Pilot Family Study or during the Phase IV Family Study as follows:

<table>
<thead>
<tr>
<th>Center</th>
<th>Family ID</th>
<th>SHS ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arizona</td>
<td>AZxxyyyy</td>
<td>360001 - 36zzzz</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>Okxxyyyy</td>
<td>260001 - 26zzzz</td>
</tr>
<tr>
<td>South and North Dakota</td>
<td>Dkxxyyyy</td>
<td>160001 - 16zzzz</td>
</tr>
</tbody>
</table>

Where  xx : family number.
yyy : 001 - 999 for each family member.
zzzz : a unique number for each family member who participated in the examination and interview.

Standard IHS community codes are used to identify the community where the participant resides. A list of community codes for the three centers is given in Appendix A-1 of Volume 2. Hospitals where an SHS participant died or was treated for CVD are also coded. Standard IHS facility codes are used to identify IHS hospitals and clinics. Codes for other non-IHS hospitals are assigned by each center. The hospital/clinic codes are given in Appendix A-2 and A-3 of Volume 2, respectively. In addition, every member of the Study is assigned a Personnel Code, which will be used to identify the person who filled out a specific data form. The Personnel Codes for the three centers are listed in Appendix A-4 of Volume 2. Additional codes are added sequentially as new employees begin to work on the project.

All data forms must be filled out legibly and completely. Each and every form must be reviewed and checked for completeness and legibility before it is entered into the computer.

1. All forms should be filled out in black pen. Print all information in block capital letters, with one letter only in each box, so that data entry errors can be minimized. For example, one should differentiate: 7 from 1, U from V, 4 from 6, P from D, M from N, C from O, and T from J.

2. For names and addresses, start from the leftmost box and leave the unused boxes blank. Include periods for initials.

3. For numerical values, fill in the boxes in a right justified manner and leave the unused boxes blank.

4. For dates, two digits are allowed for the month and day, and four digits for the year. If the number has only one digit, use zero in front of the number.

5. When recording dates, use the following rule for missing dates:

   If date is unknown/missing: 01/01/1001
   If only year is known: 06/30/year (assign mid-year as the date)
   If only year and month are known: month/15/year (assign mid-month as the date)

6. To correct an error, draw a single line through the mistake and write the correct value above.

7. Fractions should be rounded up to the nearest whole number if the fraction is 0.5 or more, otherwise, drop the fraction, e.g. 2.25 = 2; 2.75 = 3; 3.5 = 4.
8. If an interval is given, record the midpoint of the interval if it is a whole number. If the midpoint includes the fraction 0.5, use the rounding rules previously given.

9. Unless otherwise instructed, no item on any of the forms should ever be left blank. Codes to be used in the event of missing or incomplete data are given under the heading of each specific item. If there is not a code for the "unknown" category, draw two parallel lines horizontally through the box or boxes to indicate that the interviewer or abstractor did not ignore the question. For example, if the time of death is unknown, draw two lines across the boxes.

1.6.2 Procedures for data entry and verification of completeness -- See SHSV Operations Manual Volume VII - Data Entry

1.6.3 Data Transmission

The lab data, ECG data, and ultrasound data will be transmitted to the Coordinating Center through a secure protocol. If the data transmission is via an email attachment, an encrypted, password-protected file will be required. The data will be converted to SAS datasets. However, before these data are merged into the permanent data files, they will undergo various error checking procedures to ensure the SAS conversion has been done accurately.

1.6.4 Data Backup

Several backup procedures are used to ensure the safety of the SHS data files in both field centers and the Coordinating Center.

a. Daily backup: Two sets of cartridges are rotated to backup the data every day from Monday through Thursday (one for Monday and Wednesday and the other for Tuesday and Thursday).

b. Weekly Backup: Similar to daily backup, two sets of cartridges are rotated, each for every other week. Backup of the week's data set is done every Friday.

c. Optical disk backup: Additional permanent files are stored on optical disks (CDs and DVDs) for long term storage.

d. Storage of backup data: Cartridges and optical disks are stored in locked file cabinets in different offices, and one set of them is stored in a different building.

1.6.5 Quality Assurance (QC) Program

The quality control (QC) program includes close monitoring of the quality of all measurements and interview data. A Quality Control Subcommittee oversees the QC program of
the Study. The members of this Subcommittee include the NHLBI Project Manager, a representative from the Coordinating Center, one principal investigator, and the three Field Center Coordinators. The Quality Control Committee meets periodically via conference calls during the examination period to assess the results of quality control activities. The QC Committee reviews the QC data and summary statistics provided by the Coordinating Center and reports to the Steering Committee with recommendations. Recommendations are made to the appropriate centers when problems are identified. Follow-up procedures are established and monitored for all the QC activities. After each site visit, reports are reviewed. If indicated, field staff are retrained, re-certified, and re-monitored by the QC personnel. For lab data, aberrant pairs are investigated and corrective actions are taken both in the core lab and in the field sites. The quality control program includes: a) data collection, b) site visits, c) routine maintenance and monitoring of instrument performance, d) duplicate measures for physical examinations, laboratory tests, observations of personal interviews, QC for cardiology tests, and QC for surveillance (each of these components is described below). Each clinical center has a quality control officer who is responsible for all aspects of quality control at that center. The Coordinating Center closely monitors the recruitment and progress of the Study. According to the target numbers to be recruited by each center during the whole study examination period, the CC develops a timetable to indicate the projected goal for each month. The field centers report the number of participants actually examined to the CC, and the CC then compiles these numbers on a monthly basis. Cumulative achievement for each field center is then calculated by comparing the actual number examined to the projected number for the corresponding month. In addition to recruitment, the CC also monitors whether the field centers have completed their quotas for double entry of data, QC physical examinations, and QC blinded blood samples. The CC submits progress reports to the SHS Steering Committee as a tool to monitor the progress of the study. If the percentage of projected recruitment in a certain field center falls below 80%, the PI and the field coordinator are informed, so that the efforts can be focused on recruitment in the following months. Field center coordinators are responsible for reviewing all QC data as they become available and following up on any problems that are detected. The QC committee monitors the efficacy of retraining and problem solving.

1. **Data Collection**

   Every data form will be checked for completeness at the field center. Ambiguous or erroneous items will be clarified and corrected. The data entry programs generated by the Coordinating Center provide additional quality control checks by built-in range and logic checks. The program refuses to accept suspect data until the errors are corrected. Throughout the study, 10% of the examinations are selected for double entry. The Coordinating Center tracks the data entry error rates. If the data entry error rate of any field center is greater than 0.5% for any data transmission, that center has to double-enter all of the examination data for that transmission. Computer printouts of inconsistent data items are sent back to each field center for clarification or correction. Summary statistics such as mean, median, range, maximum and minimum for continuous variables and frequency distributions for categorical variables are calculated monthly for each center, and data not meeting consistency checks are flagged. Summary statistics will be
generated quarterly to identify any peculiar or unreasonable values. Further verifications will be made and errors corrected.

2. **Quality Control Site Visits**

   Quality control site visits will be made periodically to each of the three centers during the examination period. The site visit teams will include representatives from the program office at NHLBI and investigators and staff members from each of the centers. Procedures used in the clinical examination will be carefully observed for adherence to protocol. Equipment will be inspected and problems noted. The site visitors then will meet with all of the clinic staff to inform them of any observed discrepancies. In addition, a written evaluation, including corrections or improvements needed, will be sent to each center.

3. **Quality Control -- Equipment**

   Other quality control measures will include maintenance of the scale, measuring tapes, impedance, glucometer, sphygmomanometer, and ECG machine. The scale will be zeroed daily and calibrated with a known weight (50 lbs) every month or whenever the scale is moved. The standard sphygmomanometer will be inspected once a month. These inspections will include checking of the zero level, mercury or air leakage, manometer column for dirt or mercury oxide deposit, and the condition of all tubing and fittings. Other quality control measures for the blood pressure measurements will include simultaneous Y-tube observation of each technician and frequent staff meetings to provide feedback.

4. **Quality Control -- Examination**

   1) **Anthropometry and blood pressure**

   Duplicate measures of brachial artery blood pressure (systolic and diastolic) simultaneously using a double head stethoscope with two observers will be taken quarterly. Duplicate measures of anthropometry (height, weight, waist, hip, and electrical impedance measurements) will be performed with a second observer on a quarterly basis. These data will be sent to the Coordinating Center for analysis. In addition, distributions of measurements and digit preference for each staff member will be compiled and repeated quarterly. Results of the analyses will be provided to the field centers and the Steering Committee. Differences between duplicate measures exceeding the following values will be considered unacceptable:

   i.) Systolic Blood Pressure: 4 mmHg, using Y-shaped stethoscope for two simultaneous observations.
   ii.) Diastolic Blood Pressure: 4 mmHg, using Y-shaped stethoscope for two simultaneous observations.
   iii.) Height: 1 cm
   iv.) Weight: 1 Kg
   v.) Resistance: 15 ohms
vi.) Waist circumference: 2 cm
vii.) Hip circumference: 2 cm
viii.) Arm circumference: 1 cm

2) Laboratory tests

Duplicate blood and urine specimens are collected on approximately 5% of the participants in Phase V. These duplicates are sent to the Core Laboratory in a blind fashion. Results obtained for each test will be analyzed quarterly by the Coordinating Center for accuracy and consistency. The percent of pairs with differences within 5% and 10% will be computed. Correlation coefficients and technical error rates will be calculated. Poor correlations or unreasonably high technical errors will be reported to the Laboratory and the Steering Committee.

3) Personal interview

Personal interviews by new staff will be observed monthly by the study coordinator until the staff member meets the standards of the study. Then new staff will be observed on a quarterly basis along with experienced interviewers. Problems and errors are identified using a checklist and corrected immediately.

4) Food Frequency Questionnaire (FFQ)

The Block FFQ is self-administered; participants receive guidance from SHS staff in how to fill out the questionnaire. The developer, Block Dietary Data Systems (BDDS), has provided documentation (see Volume 9 of this manual) that describes each question. During the March 2006 training sessions in OK, Jean Norris, MS, RD, DrPH (BDDS) provided training for the field staffs in how to instruct participants and how to check the FFQs for completeness, for proper pencil entries on the FFQ bubble forms, and for correction of the bubble forms if improperly filled in (e.g., pen instead of pencil). Trained staff members will assist any participants having difficulty with the FFQ.

5) Quality control for surveillance data

Surveillance activities at each center are monitored on a monthly basis by the Coordinating Center. Contact rates, numbers of potential events, rates of medical record abstraction and forwarding of packets for review are evaluated each month according to pre-set, expected completion rates. Final decisions on possible CVD deaths and morbid events are made by members of the Mortality Review and Morbidity Review Committees. These surveillance committees also evaluate the quality of chart reviews and advise clinic staff when changes are needed. The Mortality Review Committee is composed of a primary physician reviewer (Dr. Maurice Sievers) who reviews all deaths and a group of six physicians who serve as secondary reviewers for all potential CVD deaths. Each physician independently determines the classification of the cause of death, and the Coordinating Center then compares the results from
both physicians. All fatal events judged to be strokes by Dr. Sievers are directly forwarded to Dr. Jorge Kizer at Cornell Medical Center, Division of Cardiology, New York-Presbyterian Hospital but not to the next member of the Mortality Review Panel. The entire Mortality Committee adjudicates potential CVD cases when there is a disagreement between the primary and secondary reviewers. A detailed description of the steps in the process of identifying deaths and confirming the underlying cause is given in Volume 2. Monthly reports are reviewed by the Steering Committee in order to monitor the progress of surveillance and event reviews. An example of a monthly surveillance report is included in Volume 2.

6) Certification of technicians

Each center recruits the most qualified personnel. Clinical staff were centrally trained and certified before the examination began, and newly hired personnel will be trained at each clinic. The study coordinators will monitor the technicians quarterly to ensure accurate and consistent performance.

7) Confidentiality and security of data

All personnel with access to the collected data are required to sign a confidentiality pledge (see Appendix 7 below). Completed data forms are placed in locked file cabinets at every center and are accessible by authorized staff members only. At the Coordinating Center, the data are stored on computers that are used exclusively by the Strong Heart Study and are safeguarded by passwords that are known only to authorized personnel. The data are stored on hard disk and four copies of optical diskettes. Two of the Zip disks/optical diskettes are stored in two different locations other than the Coordinating Center office.

8) Monitoring of study progress

The Coordinating Center works closely with the field centers to monitor recruitment and progress of the examinations. At the beginning of the study, a projected monthly number of participants to be recruited was generated, and the Coordinating Center monitors the progress of each field center according to these projected numbers and provides quarterly progress reports to the Steering Committee. If the percentage of projected recruitment in a certain field center falls below 80%, the PI and the field coordinator will be informed, so that the efforts can be focused on recruitment. This program proved to be an efficient tool for monitoring the progress of SHS in previous phases and will be continued in Phase V of SHS. The Coordinating Center will also monitor the number of double entries, QC physical exams, and QC blinded blood samples and report to the Steering Committee quarterly.

1.6.6 Statistical Analysis and Power Estimates

There are two major types of statistical analyses for Phase V of SHS. The first type is characterized by genetic or linkage analyses (for Specific Aim #1 and part of Specific Aim #3).
The following presents detailed information regarding genetic analyses and their respective statistical power calculations.

**Statistical Genetic Methods**

**Genotypic Data Cleaning**

By the beginning of Phase V, all family members (more than 1,240 from each Field Center) will have been genotyped for approximately 400 markers spaced at intervals that average 10 cM. Before these data can be used in statistical genetic analyses, they must be “cleaned”, i.e., any apparent pedigree discrepancies must be resolved and Mendelian and double recombinant errors must be eliminated. We also must estimate allele frequencies, construct genetic maps, and calculate multipoint identity-by-descent (IBD) matrices. The cleaning of all of the data in the 10 cM map will be finalized at the beginning of Phase V.

To verify the correctness of pedigrees we use PREST (McPeek and Sun, 2000; Sun et al., 2002) to sequentially answer the following two questions: 1) Which relationships are not consistent with the genotype data? 2) Among the rejected relationships, what is the most supported relationship for each one rejected? We focus on the full-sib and half-sib relationships rejected by the first stage tests. We combine a maximum likelihood approach at the relationship level with a maximum parsimony requirement at the level of pedigree configuration. The final pedigree structure is the one most consistent with the data and most parsimonious (i.e., requiring the fewest pedigree changes).

Clearly, we must resolve genotyping errors before linkage analyses are carried out. Genotyping errors influence our estimations of both map distance between markers and IBD sharing among relatives. We use SimWalk2 (Sobel et al., 2002) as the basis of our PEDSYS program PRESWALK to estimate error probabilities for each individual for each marker genotype. PRESWALK uses mistyping probabilities generated by SimWalk2 to blank genotypes using an iterative procedure. SimWalk2 can also detect genotyping errors due to spurious double-recombinants, which are difficult to detect because they may be consistent with Mendelian segregation.

Maximum likelihood techniques that account for pedigree structure, implemented in SOLAR, are used to estimate allelic frequencies. Since we type additional microsatellite markers in any (positional) candidate region, a population-specific genetic marker map of each region is constructed utilizing the known marker map positions and known sequence data. We use the marker orders specified in the Marshfield map, and the program CRI-MAP (Lander and Green, 1987) to estimate distances between markers. We use the program Loki (Heath, 1997; Heath et al., 1997) to compute multipoint IBD matrices. Loki employs MCMC methodology to compute the expectations of IBD sharing at points throughout the genome conditional on the information available at other neighboring points.

Cleaning of the SNP data and of the data for additional microsatellite markers in regions where we have linkage signals will be done in Phase V. The SNPs will be used only in measured genotype analysis and Bayesian QTN analysis, which do not rely on map information and do not require IBD computation. Given that these intra-genic SNPs are extremely close, we clearly couldn’t hope to accurately estimate their genetic map distances.
A. **Quantitative Genetic Analysis**

Initial heritability estimates have been done for risk factors measured in Phase IV. Scripts enable us to automate the analyses rather than performing them separately for each phenotype. In all of these initial analyses, age, sex, and their higher order terms and interactions are included as covariates. More refined heritability estimates, including other covariates and allowing for center, household, and other effects, will be done in the remainder of Phase IV and in Phase V, as will bivariate analyses to estimate genetic correlations among risk factors. When phenotypic data generated in Phase V become available, these too will be subjected to quantitative genetic analysis.

Heritabilities and genetic correlations are estimated using maximum likelihood variance decomposition methods (Hopper and Mathews, 1982; Lange and Boehnke, 1983) that have been implemented in SOLAR (Almasy and Blangero, 1998). If the phenotype vector is assumed to be multivariate normal, the likelihood of the pedigree (i.e., the likelihood of the observed array of phenotypes in a family under a specific genetic model) is easily calculated. Optimization methods are used to estimate parameters, and subsequent hypothesis testing is performed using likelihood ratio tests. Multivariate normality does not always hold (in fact it is clearly violated when there are major gene effects). Theoretical and simulation studies indicate that variance decomposition methods are relatively robust to deviations from normality (Beaty et al., 1985; Amos, 1994; Allison et al., 1999), although they are sensitive to kurtosis. However, kurtosis is easily identified, and simulation studies show that use of the multivariate t distribution, implemented in SOLAR, recovers the correct test statistic distribution even in cases of strong kurtosis.

The variance terms that will be included in our analyses are the additive genetic variance, the variance due to shared household effects, and the random environmental variance. The simple model can be extended to include additional (or alternative) components such as shared spouse or sibling environments, dominance genetic effects, and mitochondrial effects. Our analyses will provide estimates of the relative importance of genetic, shared environmental, and random environmental effects on disease-related phenotypes. Phenotypes that are significantly positively skewed will be ln transformed prior to analysis. To reduce Type I error in screening for covariate effects (Blangero et al., 1992), the effects of potential covariates will be simultaneously estimated in all analyses using variance component methods, which directly account for the non-independence of relatives. Any covariates whose effects are significant at the $p \leq 0.10$ level will be retained in subsequent analyses.

Quantitative genetic analyses will be done separately by center. To test for heterogeneity among centers, analyses also will be done on the combined data set, with center as a covariate. A significant center effect for a disease-related trait will indicate that the centers differ in the relative contribution of genetic and/or shared environmental factors to variation in that trait. Whenever we find a center effect, we will do additional analyses using SOLAR to determine whether the differences between centers are due to differences in heritabilities between centers or to differential expression of genes in different environments.
Bivariate quantitative genetic analyses: The quantitative genetic methods described above are readily extended to multivariate traits (Lange and Boehnke, 1983). We will perform bivariate analyses to examine the genetic correlations between pairs of phenotypes. For example, we can test whether significant genetic correlations exist between left ventricular mass and blood pressure or between fibrinogen concentration and PAI-1. Large genetic correlations between traits imply that the same genes influence both traits.

We have previously used this approach in the San Antonio Family Heart Study (SAFHS) to reveal significant genetic correlations between hormone measures and measures of body fat (Comuzzie et al., 1996), indicating that these constellations of traits share underlying genetic determinants. We also have found that HDL and triglycerides are not significantly genetically correlated with measures of body fat distribution in the SAFHS (Mahaney et al., 1995). Similar analyses, using likelihood ratio tests, will enable us to estimate the magnitude of pleiotropic effects of underlying genes in American Indians.

Genotype x age interaction and longitudinal genetic models

G×E interaction is likely to be an important influence on continuous physiological variation. Our group has pursued a number of approaches for the examination of G×E interaction in quantitative genetic analysis (Blangero et al., 1990; Blangero, 1993; Jaquish et al., 1997a,b) and in variance component-based linkage analysis (Towne et al., 1997, 1999). Our earlier work concentrated largely on G×E interactions involving dichotomous environments, such as sex (Towne et al., 1997, 1999) or smoking (Martin et al., 2000). We have since expanded our work on G×E interaction methods for continuous environmental variables such as age (Almasy et al., 2001). A simple way to model G×age interaction in a quantitative genetic analysis is to use a matrix of differences in age between individuals (or between measures on the same individual) to structure the additive genetic component of variance. One approach we have used is a simple exponential decay model. Another approach is to make the additive genetic variance a function of age and test for changing genetic variance with age. We also include in the model a correlation in genetic effects at different ages. A correlation significantly different from 1 suggests that different genes influence the trait at different ages. Positive evidence for G×age interaction is interpreted as evidence for a heritable basis to phenotypic change with aging. The strategy of modeling variance as a function of age is easily extended to G×age interaction at a QTL by making the QTL variance a function of age.

Blangero (1993) has shown the conditions necessary for this approach to work when only cross-sectional data are available, so that each individual is measured in only one environment (e.g., one age) and data across the range of environments (ages) must be obtained from different individuals. Cross-sectional data are sufficient, although less powerful than longitudinal data, to provide valid inferences concerning G×E interaction. Fortunately, we will have longitudinal data for most of our phenotypes: measures at three points in time for each individual who participated in Phases III and IV and also in Phase V and at two time points for family members newly recruited in Phase IV and re-measured in Phase V. Our multivariate analysis routines incorporated in SOLAR allow for this type of unbalanced mixed longitudinal design and efficiently use all available data.
Longitudinal data will allow us to detect more precisely and powerfully the changing effects of genes with age. There are few QTL mapping methods suitable for the analysis of longitudinal data. Multiple measurements of the same trait over time can be viewed as measurements of a single multivariate trait. However this naïve perspective is dangerous unless we properly model the correlations between measurements on the same person. Similarly, we must allow the different variance components to change with age and allow the correlations between the expressions of the trait at different ages to be a function of the age difference. In this regard, members of our group are developing a model for longitudinal QTL analysis that is very similar to our GxE model. The central difference is that we allow multiple phenotypes per individual and employ parametric (co)variance functions that depend upon the ages of the individuals involved. Additionally, because we will have multiple measures on individuals, we also can estimate environmental correlations that may change with time in addition to QTL and residual additive genetic correlations (Towne et al., 2000).

**Linkage Analysis**

Genomic screening involves a complete search of all chromosomes for genes influencing quantitative or qualitative disease risk factors. Linkage analyses in a genome screen involve genotypic data for hundreds of anonymous markers that are distributed across all of the chromosomes.

Methods of linkage analysis that exploit identity by descent (IBD) allele sharing between pairs of relatives are widely used in the genetic analysis of complex traits as these methods generally require few assumptions about the genetic model underlying expression of the trait. Basically, these types of linkage analyses do not require that one previously specify or simultaneously estimate genetic model parameters such as allele frequencies, genotype mean effects, and dominance relationships among alleles for the putative disease locus. Instead, it is sufficient to estimate only chromosomal location and a summary measure of the relative importance of the disease locus such as the locus-specific heritability.

The best known of these methods is the sib-pair approach of Haseman and Elston (1972). Recently, variance component linkage analysis methods have been developed which are more powerful than relative-pair based approaches (Amos, 1994; Goldgar, 1990; Schork, 1993; Blangero and Almasy, 1997). These variance component methods have been extended to accommodate general pedigrees of arbitrary size and complexity and to allow analyses that include genotype by environment interaction, epistasis, threshold models for discrete traits, and pleiotropy, as well as multivariate and oligogenic analyses. The variance component approach fully exploits all of the genetic linkage information in extended pedigrees, considering all possible biological relationships simultaneously. Since the variance component method requires estimation of fewer parameters than in the fully parametric penetrance-based linkage methods, it is also more efficient. Formal mathematical development of variance component linkage analysis is given in Almasy and Blangero (1998).

In linkage analyses, we test the null hypothesis that \( \sigma_m^2 \), the additive genetic variance due to the trait locus, equals zero (i.e., no linkage) by comparing the likelihood of this restricted model with that of a model in which \( \sigma_m^2 \) is estimated. From these analyses, we identify
chromosomal regions that are most likely to contain a QTL for a CVD risk factor, focusing on those regions with the highest LOD scores. These analyses utilize SOLAR (Almasy and Blangero, 1998), developed at the Southwest Foundation for Biomedical Research.

Each marker is screened for linkage to each of the quantitative traits that show significant heritabilities, using data generated for the extended families recruited into the Family Study. For each phenotype, we will perform two-point and multipoint variance component linkage analyses using SOLAR. When necessary, the linkage analyses will incorporate the effects of shared household, spouse and sib environmental effects, or genotype by environment interaction. We will test the null hypothesis that the additive genetic variance due to the trait locus equals zero (i.e., no linkage) by comparing the likelihood of this restricted model with that of a model in which the additive genetic variance is estimated. All of our analyses will include the simultaneous estimation of covariate effects.

After we have identified chromosomal regions that are most likely to contain a QTL for a disease risk factor, we will focus on those regions with the highest LOD scores. We will do saturation mapping with additional STRs in each region and will repeat our linkage analyses with the goal of strengthening evidence for linkage and narrowing the region containing the QTL.

Initial linkage analyses for Phases III and IV already have been done using an automated procedure that we have developed. Output is entered into Excel spreadsheets that indicate LOD scores by chromosomal position (every 10 cM) for each phenotype. Also indicated in the spreadsheet for each phenotype are the number of individuals and number of pedigrees included in each analysis, heritability, proportion of variance attributable to covariates, covariate effects and their p values, the maximum LOD score, and its chromosomal location. This automated process allows us to concentrate our efforts on those chromosomal regions that appear most promising. More refined analyses have been done for a few of the linkage signals detected in the initial analyses, but analyses for additional signals remain to be done. These will be pursued in Phase V. As the phenotypic data generated in Phase V become available, we also will subject them to linkage analysis and will begin to examine longitudinal changes.

We will pursue selected linkage signals detected in the Phase III data as well as linkage signals revealed in our ongoing analyses. Further analyses will involve incorporation of additional covariates for some traits as well as bivariate analyses and analyses to assess shared household effects and genotype by environment interaction. Since we have found a significant linkage signal for insulin, we plan to analyze derived insulin sensitivity and secretion traits (e.g., Melchionda et al., 2002). Along these lines, we also plan to use recently developed data imputation methods (e.g., Soler and Blangero, 2003) in our analyses of the insulin phenotypes to account for the well-known effects of diabetes medication on insulin and glucose phenotypes.

After identifying promising chromosomal regions that contain QTLs using data from the 10 cM map, we will narrow the regions of interest by fine mapping and evaluate positional candidate genes. We will genotype SNPs in known genes in the region, and use measured genotype analysis along with information garnered from literature and genetic database searches to identify candidate genes in the region. We will sequence one or two of the most promising candidate genes within the region of each QTL to identify single nucleotide polymorphisms.
(SNPs), and then genotype all SNPs that we identify in candidate genes. We will use our new statistical functional genomic approach. This includes a method for quantitative trait nucleotide analysis that will allow us to quantify the posterior probability that a given genetic variant is functional. Thus, our analyses will enable us to test whether specific SNP polymorphism(s) account for our linkage signals. If we find evidence of functional SNPs, we can then test whether the SNPs account for population-level association in the SHS cohort.

**Heterogeneity among centers:** The three Strong Heart Study centers differ in many aspects of their history, culture, and environment and may differ in their genetic risk for CVD as well. We hope to be able to map risk factor genes that are common across centers as well as those that are unique to one center. As discussed below, the families recruited in Phases III and IV provide excellent power to map genes that contribute to risk of CVD in separate analyses of each of the three centers. To detect heterogeneity among centers, we will evaluate models in which linkage parameters are separately estimated for each of the three centers, comparing them with models in which they are constrained to be the same across centers. (Analyses that combine the data across centers will include a covariate for the center effect.) If a gene in a specific chromosomal region influences a disease risk factor in only one of the three centers, then we would expect little evidence of linkage in the other two centers. The mapping of genes that influence disease risk factors in just one center will allow us to focus on specific families and perhaps to identify individuals who are at increased genetic risk of disease. If there is evidence of linkage in all three centers, then the relevant genes are more likely to be important in the general population.

**Measured Genotype Analysis**

As described below, in regions where strong linkage signals have been found, measured genotype analysis (Boerwinkle et al., 1986) will be used to examine the extent of association between phenotypes and SNPs. Using the likelihood ratio test, we will compare the likelihood of a model in which variation in the phenotype is influenced by polygenic factors, covariates, and random environmental effects (reduced model) with a model in which the effects of the SNP also are included (saturated model), using SOLAR. If a significant effect is found, we will estimate the proportion of the variation that is attributable to the SNP. The measured genotype models assume that the distribution of the phenotype is normal given the candidate locus (SNP). If necessary, we will perform measured genotype analyses that incorporate a simultaneous transformation of the data. The goodness of fit statistics suggested by Hopper and Mathews (1982) will be used to evaluate the validity of the fitted models. In these analyses, measured genotypes are assumed to have the same (fixed) effects on means in all pedigrees. Alternatively, we may perform association analyses using the Bayesian QTN approach. The benefit of this approach is that it automatically handles the multiple testing issue and can be used with either SNP genotypes or haplotypes.

**Statistical Functional Genomics**

After positional candidate loci are revealed by linkage analysis, we must attempt to determine the actual functional variants that are responsible for the observed linkage signal. This final activity takes us from the QTL to the responsible nucleotide differences (the QTNs [Long et
al., 1998; Phillips, 1999)) influencing the phenotype. Molecular sequencing and functional genetic analyses are traditionally relied upon to pinpoint the actual genetic variants involved. We propose to utilize a novel statistical functional genomic analysis that will bring rigorous statistical procedures to the final stage of identifying the specific variants involved in determining variation in disease risk.

We will apply a multi-step strategy to prioritize our analyses to identify potentially functional polymorphisms in three or more chromosomal regions where we have the strongest evidence for linkage to quantitative traits. This strategy is described in detail below. Briefly, we will:

1) Refine our linkage signals by genotyping additional microsatellites in the region of linkage.
2) Use genome databases to identify all known expressed genes in each region of linkage.
3) Genotype several previously-identified SNPs in each of the known genes in the region of linkage in those individuals contributing alleles to our study population ("founders").
4) Use measured genotype and/or QTN analyses to determine the extent of association of any of these SNPs with our linked traits.
5) Use information on function and expression obtained from the literature and from genetic databases, along with the results of the measured genotype analyses, to select the most promising candidate gene or genes from the linkage region.
6) Re-sequence the candidate gene(s) for SNP discovery in the set of individuals contributing alleles to our study population (i.e., founders).
7) Genotype the SNPs identified in step 6 in the portion of the data set in which the signal was initially detected.
8) Perform QTN analysis to identify the potentially functional polymorphisms.
9) Perform linkage analysis conditional on the potentially functional SNPs, to determine whether or not they are responsible for our linkage signal.
10) For those SNPs that explain the initial linkage result based on analysis of our partial data set, genotype them in all SHFS participants if linkage is confirmed in the entire dataset, and repeat the linkage analysis conditional on the SNPs.
11) Genotype the polymorphisms that are responsible for our linkage signals in the original SHS cohort and use measured genotype and/or QTN analysis to determine whether they are associated with quantitative traits.

This tiered approach will allow us to explore our linkage regions using multiple sources of information to prioritize positional candidate genes, and then focus our re-sequencing and QTN analyses on the most promising gene or genes. This general approach may be modified for a specific gene region. For instance, if there is a very strong positional candidate gene in the region, we might proceed directly to re-sequencing of that gene for SNP discovery. Throughout our study, at every step, we will, if necessary, modify our strategy based on the results of our molecular and quantitative analyses to give us the best chance of identifying functional polymorphisms responsible for our QTL.

Our tiered strategy for choosing multiple QTLs to follow up for prioritizing candidate genes to pursue for each QTL and for choosing the appropriate analysis based on the results from a previous step gives us flexibility to investigate and focus resources on the most encouraging
results. As we continue our phase III/IV linkage analyses, if our analysis involving one linkage region gives no direction, we will refocus our efforts and resources and move on to another QTL from our linkage results. We wish to emphasize that we will not abandon our proposed linkages without strong justification. We anticipate that our strategy will optimize our chances of finding functional polymorphisms that represent genetic risk factors for CVD-related phenotypes.

**Refining the linkage region**: In chromosomal regions where we have identified QTLs using variance component analyses with our 10-cM map, we will increase the density of markers to refine our linkage signal and reduce the genomic region containing candidate genes. We will identify additional microsatellite (STR) markers from the Applied Biosystems Linkage Mapping sets HD5, as well as several human genome databases. We will purchase primer pairs, with a fluorescently labeled reverse primer, and amplify these STRs using standard conditions (Cole and Hixson, 1998). Reaction products for each family member will be pooled according to size and fluorescent dye and analyzed by capillary electrophoresis on an ABI 3100 Automated DNA Analyzer (Applied Biosystems, Foster City, CA) using Genescan and Genotyper software. These additional microsatellites will be subjected to data cleaning, and linkage analysis will be repeated, as described above.

**Prioritization of candidate genes in regions of linkage**: To reduce the amount of resources devoted to re-sequencing in the SHFS participants, we will use a combination of in silico genetic database searches, literature searches, and measured genotype analyses to prioritize candidate genes for re-sequencing based on several factors including genetic location, function, expression, and potential association with the linked phenotype. This is the overall general strategy of prioritization, but it will be tailored to address the specific circumstances and state of knowledge regarding each QTL.

We will use results from genetic database (in silico) searches to identify known and expressed genes in our regions of linkage. We will genotype known SNPs in these genes, available as validated SNP genotyping assays from Applied Biosystems (Foster City, CA). We will genotype all individuals contributing alleles that are segregating in families responsible for our initial linkage findings. For instance, in our pilot (Phase III) families, approximately 400 individuals out of 900 meet the criteria of 1) having been genotyped, 2) having been phenotyped, 3) having offspring in the study, and 4) not having parents who were genotyped. Thus they contribute the alleles segregating in the study population and are defined as founders. We estimate that we will genotype approximately 250 SNPs per gene region in 400 SHFS participants. For details of the molecular genetic techniques, see below.

**DNA sequencing to identify polymorphisms in candidate genes**: We will use the results from our measured genotype analyses, combined with information obtained from our in silico genetic database searches and literature searches, to prioritize candidate genes for high-throughput re-sequencing analysis for SNP discovery in our SHFS subjects. Our approach, based on our experience with the chromosome 2 QTL in the San Antonio Family Heart Study, is to thoroughly examine these positional candidate genes and identify, genotype and analyze all polymorphisms in these genes.
We will obtain our gene sequence information through several sources, including published articles as well as human genome databases. For our sequencing strategy, we will first re-sequence the coding regions and 1 kb of the proximal promoter, and then continue with the noncoding regions by extending through the 5' flanking region with potential regulatory sequences, and then introns. We estimate we will sequence approximately 10 kb of each candidate gene, and will sequence at least two promising candidate genes in each of three QTLs. For large genes, our efforts on noncoding regions will focus on those that show evolutionary conservation revealed by inter-species genomic sequence comparisons (reviewed in Pennacchio and Rubin, 2001) using global sequence alignment software tools such as VISTA (Mayor et al., 2000), Exonerate Mouse (Ensembl), and Exofish (Genoscope). These conserved regions have a reasonable likelihood of having gene regulatory properties. This strategy will become more important if we are unable to identify functional polymorphisms in the coding regions, and we must move away from the structural gene to find more distant regulatory regions, or on to other candidates in the linkage region. For details on the molecular genetic techniques, see below.

**Genotyping of novel SNPs in candidate genes**: We will genotype all the new SNPs that we identify from our re-sequencing efforts in the SHFS families that are responsible for the initial linkage. We estimate we will identify approximately 25 SNPs per candidate, and type them in at least 900 additional SHFS participants (400 will already have been genotyped during SNP discovery, above). We assume that the majority of the polymorphisms we detect will be single nucleotide polymorphisms (SNPs). Our method of choice for genotyping SNPs will be the allelic discrimination assay (Holland et al., 1991) on an ABI Prism 7900HT Sequence Detector (Applied Biosystems, Foster City, CA) (see Molecular Genetic Techniques, below). Primers and probes for each SNP specific assay will be developed using the Primer Express (Applied Biosystems) software, or purchased from Applied Biosystems using their custom design service, and PCR reactions will use TaqMan Universal PCR Master Mix PCR reagents. For those few polymorphisms that we may detect that are not SNPs, we will use more appropriate methods for detection. For instance, if they are STRs, we will use the methods described above for microsatellite markers. In some instances, we might genotype individuals using direct sequencing.

**Estimation of linkage disequilibrium among SNPs**: For our examination of intragenic positional candidate gene SNPs, we will estimate all of the pairwise linkage disequilibria parameters between all pairs of intragenic SNPs using a standard pedigree-based maximum likelihood method that can handle any pattern of missing data using the program MENDEL (Lange et al., 1988).

**Quantitative Trait Nucleotide Analysis**

Given complete sequence data for a gene harboring a functional site, we can identify statistically which polymorphism(s) is/are most likely to be affecting our phenotype. Although determination of the mechanism by which a genetic variant leads to phenotypic variation will still require molecular investigation, it is possible to formulate a first-line statistical genetic approach to limit the number of genetic variants to be examined in the molecular laboratory and to prioritize them in terms of their likely importance in the population. This approach requires enumeration of all polymorphisms within the positional candidate loci and will thus require re-
sequencing of a substantial number of individuals to establish the polymorphic sites in the population. Once the polymorphisms are found, they must be typed in a large number of individuals for whom phenotypic information is available (e.g., the extended pedigree sample in which we conducted our linkage analyses). Although large volumes of re-sequencing and SNP typing are labor intensive, recent advances in technology have rendered them practical, and new technologies may make this step even more efficient in the near future.

**The QTN model:** The QTN model that we have employed represents a simple extension of the classical variance component model. If a candidate locus has numerous polymorphic nucleotide sites, one of which is functional, then the variance associated with a marker in disequilibrium with the functional site will generally be less than that due to the functional polymorphism unless the genotypes at the two loci are completely correlated. We model the phenotype as a linear combination of fixed effects and random variables. Estimation of the various fixed effects and variance components associated with the random effects can be performed using standard maximum likelihood methods.

**Model selection using the Bayesian Information Criterion:** Once the extensive polymorphism within a positional candidate gene is assayed, Bayesian model averaging/model selection will be employed to determine the functional polymorphisms. We first applied this powerful methodological framework to the study of multiple QTLs in linkage analyses (Blangero et al., 1999; Martin et al., 2001). Because there may be a large number of SNPs to evaluate in a candidate gene, there can be many possible models of QTN action. If we consider only additive QTN effects, there are $2^m$ possible models, where $m$ is the number of QTNs considered. Our approach is to evaluate all such models and utilize Bayesian methods to estimate the probability that each SNP is functional. The Bayesian QTN (BQTN) method is designed to separate potentially functional variants from neutral polymorphisms in linkage disequilibrium (LD) with them. In this framework, functional variants are those that are responsible for a displacement in the observed phenotype values. This method is predicated on the assumption that we have the complete collection of variants in the positional candidate gene. Hence the extensive re-sequencing of the candidate gene and surrounding conserved and known regulatory regions described above. The BQTN model incorporates each variant one by one, evaluating the likelihood of a model in which the trait mean varies by genotype at that variant. Then it evaluates models with all possible combinations of two variants, all possible combinations of three variants, and so on.

The phenotypic variation explained by a marker in linkage disequilibrium with a functional site is a function of the variation due to the actual functional site and the strength of LD between the marker and the functional site. Thus, the explanatory power, or effect size, of a genuine functional polymorphism is always greater than or equal to that of variants that are merely in LD with it. The effect sizes of the genuine functional variant and the neutral marker are equal only if they are in complete LD, which also requires that they have identical allele frequencies. Thus, barring complete LD, functional polymorphisms can be distinguished from markers in LD with them by their greater explanatory power in the multivariate BQTN models. In the rare case of complete LD, a SNP set can be treated as a single unit in the analyses. This will allow us to test whether one or more of the variants in complete LD is functional, but the
identification of which one will require either laboratory testing or replication in another population with different LD structures between the markers. The Bayesian Information Criterion (BIC) will be used to assess whether the QTN model explains sufficient variation in the phenotype to justify the number of parameters used. BIC differences greater than 2 are indicative of positive evidence of support for one model over another with posterior probabilities of greater than 75% (Raftery, 1995; Blangero et al., 1999). Similarly, BIC differences of 6 units represent strong support favoring a model with 95% posterior probabilities, and BIC differences greater than 10 units indicate posterior probabilities of greater than 99% and thus represent very strong support.

**Bayesian model averaging in QTN analysis:** The BIC also can be used to formulate a simple model averaging approach to estimation that explicitly allows for model uncertainty (Raftery, 1995). The main utility of this approach is that it provides an estimate of our faith that a given SNP is itself functional. We have incorporated the Bayesian model averaging/model selection procedures for QTN analysis into our program SOLAR.

**Linkage analysis conditional upon marker polymorphisms:** We will combine the QTN analysis with our IBD-based variance component linkage analysis. This will allow us to assess whether the putative functional polymorphisms found by the QTN method can account for a given linkage signal. We employ marker polymorphisms as fixed (for markers with 3 or fewer alleles) or random (for markers with 4 or more alleles) effects and then calculate the conditional LOD score after removing the effects of the marker. We have employed this method to test whether a putative functional polymorphism can adequately account for a prior linkage signal (Soria et al., 2000). If the conditional LOD score is zero, then there is no residual linkage signal, which is evidence that the marker may be the primary functional polymorphism.

**Repeating linkage analysis conditional upon marker polymorphisms in the entire SHFS data set:** Linkage analysis will be repeated upon completion of the Phase IV genotyping and data cleaning. If the linkage persists in the larger sample, we will test whether the putative functional polymorphisms we have detected using the QTN analysis explain the linkage. This will be done by genotyping those SNPs (we estimate 5 per gene) in the Phase IV SHFS subjects (~2,700), and repeating the linkage analysis conditional upon marker polymorphisms. If the putative functional polymorphisms detected in the Phase III families do not explain the linkage in the larger data set, we will continue with SNP discovery and additional QTN analysis.

**Measured genotype analysis in the SHS cohort:** One benefit of the SHS design is the opportunity to assess the effects of putative functional polymorphisms in the large SHS cohort. We will genotype these polymorphisms in the SHS cohort (~4000 individuals) and perform measured genotype analysis to determine the potential association of the polymorphism with the traits of interest. The results of these additional analyses in the cohort will not impact on the results of the QTN and conditional linkage analysis in the SHFS, since rare functional polymorphisms might not have a population-level effect. But for those polymorphisms that might not be rare, or might have a large effect, utilizing the cohort will confirm their effect and usefulness as markers of CVD risk in the SHS population as a whole.
Power Considerations for Genetic Analysis

We have performed an extensive series of computer simulations to evaluate our power to detect quantitative trait loci by genomic scanning in the Family Study. We assumed a sample size of 1,200 family members for each center, with structures like those of the families recruited in the pilot study. For a 10 cM map, a QTL will be at most 5 cM from a polymorphic marker. Table 1 indicates the QTL heritability for which we will have 80% power to detect the QTL for a range of LOD scores. A LOD of 1.175 corresponds to a p value of 0.01. It can be seen that even without typing additional markers in regions of interest, we will be able to obtain suggestive evidence for linkage (LOD > 2) for QTLs that account for as little as 18% to 21% of the residual variance in a CVD risk factor (after covariate effects have been taken into account) in separate analyses of each center, and for as little as 9% of the variance in situations in which it is appropriate to analyze the combined data across centers. The power to detect linkage is nearly as great for the variance component method, which does not require estimates of penetrance parameters, as for penetrance-based linkage analysis in which the model is known.

Table 1. Estimated QTL Heritability Detectable at Recombination Fraction \( \theta = 0.05 \) with Power = 80%.

<table>
<thead>
<tr>
<th>LOD score</th>
<th>&gt;3</th>
<th>&gt;2</th>
<th>&gt;1.175</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arizona</td>
<td>0.24</td>
<td>0.21</td>
<td>0.16</td>
</tr>
<tr>
<td>Dakotas</td>
<td>0.24</td>
<td>0.18</td>
<td>0.14</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>0.23</td>
<td>0.18</td>
<td>0.15</td>
</tr>
<tr>
<td>Combined</td>
<td>0.11</td>
<td>0.09</td>
<td>0.05</td>
</tr>
</tbody>
</table>

Table 2. Power for Projected Sample Size of 1,200 at Each Center.

<table>
<thead>
<tr>
<th></th>
<th>Power to obtain a LOD score</th>
<th>Power to obtain a LOD score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( h^2 )</td>
<td>( ELOD )</td>
</tr>
<tr>
<td>Arizona</td>
<td>0.05</td>
<td>0.7595</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>1.7544</td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>2.6484</td>
</tr>
<tr>
<td></td>
<td>0.20</td>
<td>4.6704</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>6.3888</td>
</tr>
<tr>
<td></td>
<td>0.30</td>
<td>9.1376</td>
</tr>
<tr>
<td>Dakotas</td>
<td>0.05</td>
<td>0.8950</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>1.9475</td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>3.3470</td>
</tr>
<tr>
<td></td>
<td>0.20</td>
<td>4.7829</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>8.0876</td>
</tr>
<tr>
<td></td>
<td>0.30</td>
<td>10.4270</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>0.05</td>
<td>0.7965</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>1.7891</td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>3.3872</td>
</tr>
<tr>
<td></td>
<td>0.20</td>
<td>5.1546</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>7.2513</td>
</tr>
<tr>
<td></td>
<td>0.30</td>
<td>10.0687</td>
</tr>
<tr>
<td>Combined</td>
<td>0.05</td>
<td>2.4510</td>
</tr>
<tr>
<td></td>
<td>0.10</td>
<td>5.4910</td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>9.3826</td>
</tr>
<tr>
<td></td>
<td>0.20</td>
<td>14.6079</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>21.7276</td>
</tr>
<tr>
<td></td>
<td>0.30</td>
<td>29.6334</td>
</tr>
</tbody>
</table>

For any linkages detected in the genome screen, the next step is to saturate the region with additional markers, thus reducing the maximum distance of markers to the trait locus, and increasing the power to localize the gene. After a genomic screening using the 10 cM map, additional markers will be genotyped in the region. Results in Table 2 indicate power to detect a
QTL linked at recombination fraction $\theta = 0$ for a range of LOD scores, and with heritability ($h^2_q$) ranging from 0.05 to 0.30. Also shown are expected LOD scores (ELODs) for specific QTL heritabilities in each center and for the combined data set. For each of the three centers, power is greater than 80% to detect a QTL that accounts for as little as 20 percent of the variance, with a LOD greater than 3.0. In analyses across all three centers (which would be appropriate if there is linkage homogeneity across centers), a QTL with heritability of 0.10 will be detectable with a LOD greater than 3.0, with a power of 90%.

Note that these power calculations are for detection of a particular locus. Since we expect these traits to be influenced by multiple loci, our power to detect at least one of them is greater than our power to detect a particular locus. For example, while our power to obtain a LOD of 3 for a QTL that accounts for 10% of the variance is only 19% in the Arizona sample, if there are 5 such QTLs influencing the trait, our power to detect one or more of them would be $(1 - (1 - 0.19)^5)$ or 65%.

Power to detect a functional effect: We performed a large number of computer simulations to assess the power to detect a functional effect of a given relative size in the population. Table 3 shows the results of this simulation. The allele frequency of a QTL was varied (with three possible values, 0.5, 0.3, and 0.1) and the displacement between homozygous means was altered to obtain a total QTL-specific heritability of 0.01, 0.02, 0.03, and 0.04. We analyzed the simulated data using our QTN fixed-effect model. Our results indicate that we have outstanding power (>88%) to identify functional effects that account for as little as 2% of the total phenotypic variation in a trait.

<table>
<thead>
<tr>
<th>QTL-specific heritability</th>
<th>$p_q$</th>
<th>0.01</th>
<th>0.02</th>
<th>0.03</th>
<th>0.04</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td></td>
<td>52.4</td>
<td>90.5</td>
<td>98.4</td>
<td>99.8</td>
</tr>
<tr>
<td>0.3</td>
<td></td>
<td>48.4</td>
<td>92.7</td>
<td>97.8</td>
<td>99.8</td>
</tr>
<tr>
<td>0.1</td>
<td></td>
<td>44.9</td>
<td>88.1</td>
<td>96.4</td>
<td>99.8</td>
</tr>
</tbody>
</table>

Molecular Genetic Techniques

Our strategy applied to the use of the following techniques is described in detail in the Statistical Functional Genomics section, above. Below are descriptions of each technique.

High-throughput SNP genotyping: We will use SNP genotypes in our measured genotype analyses and QTN analyses described above. Our method of choice for genotyping SNPs will be the allelic discrimination assay (Holland et al., 1991), which allows direct detection of the PCR product by the release of a fluorescent reporter as a result of PCR using the 5'-nuclease. This technique is more robust, allows higher throughput, and requires less up-front assay development than other SNP genotyping assays (Holloway et al., 1999). For known SNPs, we will genotype our study subjects using the Applied Biosystems validated SNP genotyping assays and the TaqMan Universal PCR Master Mix PCR reagents on an ABI Prism 7900HT Sequence Detector. For novel SNPs that we identify, we will develop primers and probes for each SNP-specific assay using the Primer Express (Applied Biosystems) software, or we will use Applied Biosystems custom design service. Two probes that hybridize to the target sequence containing the SNP are used in the assay. Each probe consists of an oligonucleotide with a 5'-
reporter dye and a 3'-quencher dye. When the probe is intact, the proximity of the reporter and quencher dyes results in suppression of the reporter fluorescence. As the Taq polymerase cleaves the probe with its 5' to 3' nuclease activity, the reporter dye is separated from the quencher dye, resulting in increased fluorescence. This fluorescence is read and quantified on the ABI Prism 7900HT Sequence Detector.

DNA sequencing to identify polymorphisms in candidate genes: Our general approach and strategy applied to the selection of genes for re-sequencing is described above. For sequencing reactions, we will design PCR primers to amplify approximately 1 kb overlapping gene fragments, as well as internal sequencing primers for each fragment. PCR reactions will use standard conditions. The PCR products will be treated with enzymes to inactivate the unincorporated primers and deoxynucleotide triphosphates in the samples (Nickerson et al., 1998). The amplified fragments will be used in cycle sequencing reactions using the ABI PRISM BigDye Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems, Foster City, CA), according to the manufacturer's instructions. The sequencing reactions will be precipitated with ethanol, re-suspended, and loaded into an ABI 3730XL DNA Analyzer (Applied Biosystems). The ABI sequence software will be used for lane tracking and first pass base calling, and the sequence data will be analyzed to identify polymorphisms using Sequencher Version 4.1 (GeneCodes, Ann Arbor, MI), or transferred to a UNIX workstation (Sun Microsystems, Inc.) for analysis using the programs Phred, Phrap, PolyPhred, and Consed (Nickerson et al., 1997; 1998). Both strands will be sequenced to help resolve polymorphisms in heterozygous individuals.

Epidemiologic Analysis Methods

The second type of statistical analyses is characterized by epidemiologic analyses for the cohort and family study data (for Specific Aim #2 and the first part of Specific Aim #3). These analyses will be performed at the SHS Coordinating Center or by investigators at MedStar and Cornell, and are described below.

a. Power estimation and epidemiologic analysis for Specific Aim #2 (Morbidity and Mortality Surveillance)

In the following, we present our analysis plan and describe the adequacy of our sample size by providing power estimates. We address these issues for each question listed in the specific aim.

Question a. What risk factors are related to the incidence of CVD across different age strata? What are the age- and gender-adjusted risk factor profiles for premature CVD deaths vs. non-premature CVD deaths?

For the first question, we will use the data from those participants who were CVD-free at the baseline exam to assess the association of CVD incidence with potential risk factors measured at the baseline exam. The Cox proportional hazards model with stepwise selection procedure will be used to identify risk factors that are significantly related to the time to a CVD
event (first occurrence of a CVD event or censored at last follow-up or May 2009) within each age stratum and over all age strata. Interactions between risk factors will also be assessed.

The average CVD mortality and morbidity (M&M) surveillance follow-up time of the original SHS cohort from the baseline exam to May 2009 will be approximately 18 years. There was a total of 4372 CVD-free participants at the baseline exam (1467 in AZ, 1452 in OK, 1453 in SD/ND), or an average of 1457 per center. Our power analyses for detecting associations of the development of CVD/CHD with several major risk factors are based on the average 18 years of follow-up and the average numbers of participants per center. We selected the following risk factors for the power analysis: hypertension (HTN), LDL-C, diabetes (DM) and macro/micro albuminuria, based on analyses of the currently available surveillance data. Furthermore, from the mortality and morbidity surveillance data currently available, the overall CVD (CHD) incidence was 2.18 (1.79) per 100 person-years. Gender-specific incidence rates of CVD (CHD) were 1.85 (1.48) and 2.73 (2.32) per 100 person-years for women and men, respectively. Age-specific incidence rates were 1.51 (1.28), 2.61 (2.11) and 3.61 (2.91) per 100 person-years, respectively, for age groups 45-54, 54-64 and 65-74 years. There was an annual reduction of participants under M&M surveillance of 1.69%, due predominantly to non-CVD deaths of participants never having a CVD event, with a small number of participants lost to morbidity surveillance (only 0.2% cumulative loss to mortality follow-up). All of these figures were used in calculating statistical power.

Table 4. The smallest hazard ratio (SHR) of a risk factor for a disease that can be detected with 80% power at the 0.05 level of significance based on the average no. of participants per center, and the hazard ratio (HR) observed in preliminary analyses of the currently available data.

<table>
<thead>
<tr>
<th>Disease</th>
<th>Risk factor (RF)</th>
<th>% with RF</th>
<th>N=1457</th>
<th>N=874</th>
<th>N=583</th>
<th>N=730</th>
<th>N=479</th>
<th>N=248</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVD</td>
<td>Hypertension</td>
<td>38</td>
<td>1.90</td>
<td>1.30</td>
<td>1.79</td>
<td>2.08</td>
<td>1.44</td>
<td>1.80</td>
</tr>
<tr>
<td>DM</td>
<td>45</td>
<td></td>
<td>2.41</td>
<td>1.42</td>
<td>1.67</td>
<td>1.43</td>
<td></td>
<td>2.39</td>
</tr>
<tr>
<td>LDL-C&gt;130 mg/dl</td>
<td>34</td>
<td>1.47</td>
<td>1.30</td>
<td>1.22</td>
<td>1.78</td>
<td>1.46</td>
<td>1.75</td>
<td>1.54</td>
</tr>
<tr>
<td>Micro/Macro Albuminuria</td>
<td>29</td>
<td>1.75</td>
<td>1.32</td>
<td>1.80</td>
<td>1.46</td>
<td>1.72</td>
<td>1.48</td>
<td>1.73</td>
</tr>
<tr>
<td>CHD</td>
<td>Hypertension</td>
<td>38</td>
<td>1.88</td>
<td>1.33</td>
<td>1.79</td>
<td>2.04</td>
<td>1.48</td>
<td>1.71</td>
</tr>
<tr>
<td>DM</td>
<td>45</td>
<td></td>
<td>2.35</td>
<td>1.47</td>
<td>1.69</td>
<td>1.47</td>
<td></td>
<td>2.21</td>
</tr>
<tr>
<td>LDL-C&gt;130 mg/dl</td>
<td>34</td>
<td>1.65</td>
<td>1.33</td>
<td>1.43</td>
<td>1.49</td>
<td>1.89</td>
<td>1.49</td>
<td>2.04</td>
</tr>
<tr>
<td>Micro/Macro Albuminuria</td>
<td>29</td>
<td>1.69</td>
<td>1.35</td>
<td>1.78</td>
<td>1.52</td>
<td>1.61</td>
<td>1.52</td>
<td>1.67</td>
</tr>
<tr>
<td>X</td>
<td>15</td>
<td></td>
<td>1.46</td>
<td>1.68</td>
<td>1.69</td>
<td></td>
<td>1.83</td>
<td>1.81</td>
</tr>
</tbody>
</table>

Table 4 gives the smallest hazard ratios (SHR) that can be detected with 80% power at the 0.05 level of significance. Table 4 above also gives the observed hazard ratios (HR) for the same risk factors according to the currently available data. The SHR with the estimated sample sizes are in most cases smaller than the HR based on existing data. In addition, we calculated the SHRs for risk factors with prevalence proportions of 10%, 15% or 20%. Even with a 10%
prevalence, the SHR obtained were mostly less than 2.0 with only one slightly above 2.0 for CVD in the 65-74 age group. Included in Table 4 as an example is a hypothetical risk factor (X), which has a prevalence proportion of 15%. The detectable SHRs are all less than 2.0 except one slightly above 2.0, in the 65-74 age group for CHD. Thus, our sample sizes are adequate to study the center-specific associations of CVD/CHD with risk factors with a reasonable range of prevalence proportions, as well as the gender- and age-specific associations within a center.

For the second question, we define premature CVD deaths as deaths before age 55 for men and age 65 for women. These age cut-points were selected based on the SHS age-specific CVD incidence rates provided above, the SHS community mortality data and age- and sex-specific CHD prevalence and mortality data from NHLBI. All of these indicate that CVD mortality and morbidity increase sharply after these ages. To address this question, we will include participants who were < 65 years (women, 2205) or <55 years (men, 961) at the baseline exam (total 3166) in the analysis of risk factors for premature CVD death. The follow-up time will be from the baseline exam to May 31, 2009, or age 65 for women and 55 for men, whichever comes first. Those participants who died of CVD during the follow-up period will be uncensored observations, and those who died of a non-CVD cause or those who are still alive at age 65 (women) or 55 (men) or alive at the end of May 2009 will be censored observations. We will use the Cox proportional hazards model to identify significant risk factors, adjusting for age and gender. For non-premature death risk factors, included in the risk factor analysis will be participants who were ≥ 65 (women) or ≥ 55 (men) at the baseline exam plus those individuals who were < 65 years (women) or <55 years (men) at the baseline exam and turned 65 (women) or 55 (men) at the second or third exam. Including the latter two groups will increase our sample size. The follow-up period will be from the baseline exam (first group), second exam (second group) or third exam (third group) to the end of May 2009. Those participants who died of CVD during the follow-up period will be uncensored observations and those who died of a non-CVD cause or are still alive will be censored observations. Similarly, we will use the Cox proportional hazards model, adjusting for age and gender. The stepwise selection method will be used in both analyses to obtain the most significant risk factors. Risk factors identified in these two models (both are age and gender adjusted) can then be compared.

The data from the 3166 participants at the SHS baseline exam will be used in analyses for premature CVD death. Based on currently available data, the annual reduction rate was 1.69%, and an estimate of the premature CVD death rate was 0.43 per 100 person-years and the prevalence proportions of diabetes and hypertension were 45% and 35%, respectively. With these figures, we can detect an HR for premature CVD death among diabetics as low as 1.58 compared to non-diabetics, and for hypertensives as low as 1.61 compared to normotensives, with 80% power at the 0.05 level of significance. The observed HRs using currently available surveillance data from the 3166 participants were 2.52 for diabetes and 1.8 for hypertension, which were higher than those we will be able to detect in Phase V. For risk factor analysis of non-premature deaths, we have 1383 participants who were ≥65 (women) or ≥ 55 (men) at baseline (possible follow-up time is 18 years), 501 participants who turned 65 or 55 at the second exam (possible follow-up time is 14 years), and an additional 544 participants who turned 65 or 55 at the third exam (possible follow-up time is 10 years). Thus, we have a total of 2428
participants with an average follow-up time of 15.4 years. Based on currently available data, an estimate of the non-premature CVD death rate was 1.55 per 100 person-years, and the prevalence proportions were 47% for diabetes and 50% for hypertension. With the same power (80%) and significance level (0.05), for non-premature CVD death, the smallest HRs we can detect are 1.27 for both diabetes and hypertension while the observed HRs were 2.34 for diabetes and 2.09 for hypertension based on currently available surveillance data.

**Question b.** How are incidence rates for various manifestations of CVD (e.g., coronary, cerebral, peripheral) influenced by age, gender and diabetes status?

To study the relationships between incidence rates for these various manifestations of CVD, we will apply the Cox proportional hazards model to the time from the baseline exam to the first occurrence of each individual type of CVD and potential risk factors such as age, gender and diabetes status. The significance level obtained for the coefficient of each of these risk factors and the HRs indicate if the specific risk factor has significant influence on the incidence of the specific manifestation. In addition, the age-, gender- or diabetes status-specific incidence rates for different types of CVD can be estimated separately by the respective fraction in which the sum of the disease-free time in the specific age (or gender, or diabetes status) group is the denominator and the number of disease cases in that specific group is the numerator. If we use the CVD-free time (the time from the baseline exam to the first occurrence of any one of the manifestations of CVD), the different manifestations of CVD can be treated as competing risks during the follow-up period for the participants who were CVD-free at the baseline exam. The competing risks model and stepwise selection method can be applied to the time to the first such event computed from the baseline exam with risk factors such as age, gender and diabetes. The coefficients and the HRs obtained for these risk factors will demonstrate their influence on the incidence of these various manifestations of CVD. Other potential risk factors can be used to adjust for possible confounding effects.

Based on the total 4372 CVD-free men and women at the baseline exam, incidence rates of 1.76, 0.39, and 0.03 per 100 person-years for CHD, stroke, and the other CVD, respectively, a 45% diabetes prevalence and a 38% hypertension prevalence in the 4372 participants, the smallest HR for developing CHD (stroke) that we can detect in Phase V is 1.17 (1.40) for diabetes and 1.18 (1.41) for hypertension, which are smaller than the observed HRs based on currently available data, i.e., 1.92 (2.05) for diabetes and 1.86 (2.49) for hypertension (these HRs are slightly different from those give in Table 4 for CHD because the competing risks model is used here). Thus, we have adequate power to estimate the associations of stroke/CHD/other CVD incidence with the potential risk factors. The above power analysis is based on the competing risks model, which is more conservative than that based on the analysis of each individual manifestation. Therefore, our sample size will provide adequate power for the association analyses of individual sub-types of CVD.

**Question c.** What are the relations between quantitative measures of systemic atherosclerosis, cardiac hypertrophy and cardiovascular dysfunction (e.g., LV mass, carotid plaque, carotid wall...
thickness) and CVD incidence and mortality? Are these potential predictors related to other established risk factors such as diabetes?

For the echocardiographic and carotid ultrasound variables measured at the second and third exams, the average follow-up times to May 2009 are approximately 14 and 10 years, respectively. The Cox proportional hazards model will be applied to the CVD-free time (from the 2nd/3rd exam to 5/2009) to study its association with potential echocardiographic and carotid ultrasound variables and identify the significant variables with and without adjustment for other risk factors. By including other established CVD risk factors, we can study their relationships with the echocardiographic and carotid ultrasound variables. Because of the page limitations in this application, we will use LV mass and presence of carotid plaque as examples of risk factors and MI and stroke as examples of CVD outcomes to illustrate the adequacy of our sample size.

There were 2934 CVD-free participants who had measurements of LV mass in the second exam, and 25% of them were classified as having “high-risk” LV mass because of left ventricular hypertrophy. Based on an average follow-up of 14 years, an annual reduction rate of 1.69%, and estimated incidence rates of 0.81 and 0.44 per 100 person-years for MI and stroke, respectively, we can detect HRs as low as 1.41 for MI and 1.58 for stroke, for the high-risk LV mass group vs. the other LV mass group, with 80% power at the 0.05 level of significance. These estimated detectable HRs are smaller than those reported previously (2.0 for both MI and stroke). Similarly, there were 2456 CVD-free participants who had carotid ultrasound measurements at the 3rd exam. The prevalence of carotid plaque was 55%. With an average of 10 years of follow-up, we can detect HRs as low as 1.46 for MI and 1.66 for stroke in participants with and without carotid plaques, with 80% power at a significance level of 0.05. The observed HR associated with carotid plaques was 2.5 for both MI and stroke. Therefore, our sample sizes are adequate to perform longitudinal analyses of the associations of MI or stroke incidence with LV mass and carotid plaque.

**Question d.** What are the incidence rates and major risk factors for specific types of stroke (atherothrombotic, cardioembolic, hemorrhagic, etc), and do they differ by gender or diabetes status?

Eight subtypes of cerebral events (atherothrombotic infarction, cardioembolic infarction, lacunar infarction, other unknown infarction, subarachnoid hemorrhage, intraparenchymal hemorrhage, TIA, and unknown type of stroke) were included in the SHS surveillance. We will determine incidence rates for each subtype by dividing the number of participants who experienced a specific subtype of cerebral event by the total number of person-years of follow-up for participants who had not experienced a cerebral event at the baseline exam. The person-years will be computed from the baseline exam to the time of the cerebral event or last follow-up time. However, in identifying major risk factors, we will not be able to study every subtype individually because of the small number of incident cases for some sub-types. We will focus our analyses on the risk factors related to thrombo-embolic (cardioembolic and atherothrombotic) strokes.
The Cox proportional model with competing risks (thrombo-embolic strokes vs. other types of strokes) will be used to study the association of thrombo-embolic stroke incidence with potential risk factors in men and women, separately and in individuals with and without diabetes. Their risk factor profiles can then be compared.

Based on the current surveillance data, the incidence of thrombo-embolic strokes was 0.26 per 100 person-years. The prevalence proportions of diabetes and hypertension were, respectively, 45% and 38% in the 4372 participants who were CVD-free at baseline. Using these estimates and an average follow-up of 18 years, we will have 80% power at 0.05 significance level to detect an HR (of developing thrombo-embolic strokes) as low as 1.5 for participants with diabetes vs. those without, and an HR as low as 1.52 for participants with hypertension vs. those without, using the competing risks model. These estimates of the SHRs that can be detected with our sample size are less than the HRs (2.33 for diabetes and 2.29 for hypertension) obtained from our preliminary analyses.

**Question e.** What factors are significantly related to long-term survival? Are there differences in the factors that predict longevity between individuals with and without diabetes at baseline?

For the SHS population, which is known for high rates of the early death, we consider surviving to 75 or more years of age “long-term” survival. To identify factors that are related to “long-term” survival, we will use the data from participants in the 57-74 age group at the baseline exam (those who were 45-56 years of age at baseline will not reach 75 after 18 years of follow-up). We will apply the Cox proportional hazards model to the survival time of these participants, and use the stepwise selection method to identify significant factors that are associated with their “long-term” survival. The same method can be applied to groups of individuals with or without diabetes in the same age group at the baseline exam. The risk profiles so obtained can then be compared.

There were 1914 participants aged 57-74 at the baseline exam. These participants will be followed up for an average of 18 years to May 2009. Based on the current surveillance data, the incident rate of death before 75 was 2.74 per 100 person-years. With a sample size of 1914 participants, 51% prevalence of diabetes and 33% prevalence of micro/macro albuminuria in this group, we have 80% power at the 0.05 level of significance to detect an HR for dying before 75 among diabetics as low as 1.19 compared to non-diabetics, and as low as 1.2 for individuals who have micro/macro albuminuria compared to those without. These estimated SHRs are less than the observed HRs, 1.36 for diabetes and 1.89 for micro/macro albuminuria, based on currently available surveillance data.

**Question f.** What are the age-specific CVD incidence and mortality rates and all-cause mortality rates including rates of premature CVD death in American Indians in the three SHS geographic areas? Do these rates differ significantly among the three areas and if they do, what are the explanations for the differences?
We will determine the age-specific CVD incidence and mortality rates and all-cause mortality rates in the three geographical areas and center-specific risk factors using the same methods as we used in the past. To compute the age-specific mortality rates, the sum of the observation times in person-years of the participants in a specific age group at baseline (45-54, 55-64 and 65-74) will be the denominator and the number of deaths at those ages the numerator. To calculate the age-specific CVD incidence rates, we will use all of the participants who were CVD-free at the baseline examination. For each of these participants, we will compute the CVD-free time, i.e., the time from the baseline examination to the date of CVD death, or the date of first CVD event, or last follow-up, whichever comes first. The sum of the CVD-free-times will be the denominator and the numerator will be the number of new CVD cases identified until the last follow-up. The premature CVD death (before age 55 for men and 65 for women) rate will be calculated using participants who were, at baseline exam, 45-64 years old for women and 45-54 years old for men. The Cox proportional hazards model will be applied to survival/disease-free time to identify significant risk factors, including SHS center, for each age group. If the center effect is significant in an age group, the proportional hazards model and the stepwise selection method will be applied to center-specific data in that age group to obtain a risk factor profile for each center. The resultant risk factor profile for each center can then be used to explain differences in the center-specific incidence rates and to better understand the impact of risk factor prevalence and interaction within a center on disease risk. The power analysis provided for Question a. above is also applicable here. We have adequate power to answer this question.

**Question g.** Have age-specific mortality rates and proportional mortality ratios for CVD and other causes of mortality changed over the 20 years of the SHS follow-up (1989-2009)? Do changes differ between individuals with diabetes and without diabetes? If they do, what are the explanations for the differences?

First, the annual age-specific mortality rates (45-49, 50-54, 55-59, 60-64, 65-69, 70-74) for CVD and other causes will be calculated starting from 1990. To calculate these rates, the number of participants in the age group during that year will be the denominator and the number of CVD (or other causes) deaths in that age group during that year will be the numerator. The age-specific annual mortality rates of CVD and of other causes so obtained will be used to calculate the proportional mortality ratios for CVD and other causes of mortality. The annual mortality rates and proportional mortality ratios will be examined for possible variations and trends. Similar rates and ratios will also be calculated separately for individuals with and without diabetes and compared over the 20 years of the SHS. If meaningful differences are found between the two groups in any age group, attention will be focused on that age group; the proportional hazards model will be applied to the time to CVD death, and its relationship to diabetes and other risk factors, including interaction terms of diabetes and other variables (e.g., systolic blood pressure), will be examined. The coefficient for diabetes will be used to test if the difference between the two diabetes status groups is significant, and the coefficients for interaction terms such as diabetes by SBP and non-diabetes by SBP will be used to test the difference in the effect of SBP in the two diabetes status groups. The power analysis for Question a above is also applicable here. We have adequate power to answer this question.
Question h. What are the health-adjusted life expectancies (HALE) of the American Indians in the three geographic areas? What is the impact of chronic health conditions such as CVD, diabetes, obesity and renal disease on the HALE?

Health-adjusted life expectancies (HALE) will be calculated by combining the weighted age-specific health related quality of life (HRQOL) and age-specific mortality. The weighted age-specific HRQOL will be estimated from the MOS 36-item short-form health survey (SF-36) we collected in the second examination. Mortality information will be obtained from the mortality surveillance. One way to calculate HALE is to first use the life-table methods for gender-age-specific mortality rates for the participants with and without the health condition (e.g., obesity and diabetes) of interest. Second, by including the prevalence of the health condition of interest with the life-table method, the modified Sullivan method can be used to estimate the years of life lived free of the specific health condition. HALE can then be estimated by the years of life lived, weighted with the gender-age specific HRQOL indices. We will calculate the overall HALE and the center-specific HALE for various health conditions, including CVD, diabetes, obesity and renal disease, and compare them with national and respective state data and data from other ethnic groups.

b. Methods for analyses of data from the re-examination of family members (Specific Aim #3a)

One of the opportunities provided by the re-examination of family members (participants in the Phase IV exam) is the assessment of changes from the Phase IV to Phase V exam in key CVD risk factors that are the focus of the linkage analysis. Most of these key risk factors are continuous and these include BMI, LDL-C, HDL-C, TG, SBP, fasting glucose, insulin and fibrinogen. These risk factors and disease outcomes of interest measured from members of a family are considered related observations. We will use the marginal model designed for related observations to analyze the data. First, we will calculate the difference for each continuous risk factor observed between the Phase IV and Phase V exams for each participant. To compare the difference for a continuous risk factor (e.g., fasting glucose) among participants in different subgroups (e.g., different age, gender, center or diabetes status) or to study the association of the difference with other risk factors, we will use the marginal model with the identity link function. For example, to study the association of the difference in fasting glucose with other factors, we will apply the marginal model with the identity link function and include the effects of age, gender, household, center, diabetes status, time between Phase IV and Phase V, BMI, albuminuria, etc. and interactions of these risk factors in the model. In this model we assume that the observed differences in fasting glucose between Phases IV and V from members in the same family are correlated according to an unstructured familial correlation structure.

c. Methods for analyses of longitudinal data from the SHS examinations

For data collected from the SHS cohort members in Phases I – III exams and even for some participants in Phases IV and V, we will use marginal models, which take into consideration the relatedness issue of the observations from a participant at different follow-up
exams. These data include disease outcomes such as hypertension, diabetes, and clinical conditions such as LVH and presence of carotid plaques, and potential risk factors that were collected only at the exams (not included in the surveillance). The analyses will include examination of the relationships of the cumulative incidence of a disease outcome to risk factors, the changes of a risk factor observed at different exams, and the association of a risk factor with other risk factors. For a categorical disease outcome, such as HTN or the presence of carotid plaques, we will use the marginal model with the logit link function to study its association with potential risk factors. To study the changes in a continuous risk factor, such as LV mass, at different follow-up exams, or to examine the association of a continuous risk factor with other risk factors, we will use the marginal model with the identity link function. We assume that the disease outcome or risk factor measurements obtained from a participant at different follow-up exams are related according to an unstructured correlation structure. Included in the model will be exam effects and other covariates to examine how a risk factor, e.g., LV mass, changed over repeated follow-up exams, after adjusting for other covariates. These studies will allow us to address developmental issues related to changes in risk factors over time.
1.7 PUBLICATION POLICY

The SHS Steering Committee appointed the following members to form a Publications and Presentations Committee (P&P Committee):

Dr. Elisa Lee (Chair)
Dr. Lyle Best
Dr. Richard Fabsitz
Dr. Barbara Howard
Dr. Jean MacCluer
Dr. Mary Roman

The P&P Committee shall review and approve/disapprove all paper and thesis proposals. When the P&P Committee does not reach a consensus on a proposal, or when issues concerning a proposal (or other publication matters) are particularly problematic, the matter will be referred to the SHS Steering Committee (SC). The P&P Committee will present the issues and any of its recommendations to the SC, which shall have final authority for approval or disapproval of the paper or thesis proposal (or other publication matters).

The P&P Committee shall meet or discuss by telephone, monthly, or as needed, proposals submitted for a paper or a thesis (and any other publication matters).

1.7.1 Submission of a Paper Proposal

I. Proposal

A formal paper proposal (see Appendix 8 below - this form can be downloaded – see SHS website: http://strongheart.ouhsc.edu) must be submitted to the Chair of the P&P Committee (Elisa T. Lee, PhD at elisa-lee@ouhsc.edu) at least one week prior to the P&P meeting. Upon review for completeness (including preliminary review of the analysis plan by a statistician), the proposal will be added to the agenda of the next P&P Committee meeting for action. The Chair is responsible for distributing copies of the proposal to the members of the Committee.

A formal paper proposal must include the following as a minimum:

1. Title (To maintain a cohesive body of literature, each publication using SHS data should include the phrase "Strong Heart Study" in its title and listed as a keyword whenever possible. Titles not meeting this guideline must be justified at the time of manuscript proposal submission.)

2. Primary author's name, contact information including fax and e-mail, and affiliation. Via distribution of P&P Committee minutes, the P&P Committee will periodically report its
decisions to the SHS Steering Committee (SC), and SC may nominate additional co-authors for any papers that have been approved by the P&P Committee.

3. **Suggested co-authors**

4. **Suggested key words**

5. **A detailed outline which includes:**
   a) Introduction (rationale)
   b) Methods
   c) General analysis plan

6. **Analysis responsibility (authors or Coordinating Center, CC)**

7. **References** (the timeliness and originality of a proposal should be supported by the supplied references).

8. **When submitting a proposal**, authors are encouraged to send a copy of any journal articles that would support their choices for methods of statistical analysis. This will simplify the review process on the part of the statistician performing the preliminary review of the proposal.

9. **Prior to submission**, all proposals must be approved by an SHS P.I. For manuscripts written by investigators outside of SHS, the SHS PI co-author or the SHS PI who is closely affiliated with any of the authors must advise the P&P Committee during the review of the manuscript proposal whether the penultimate manuscript should be sent to the P&P Committee for review prior to submission to a journal. Additionally, if no SHS PI is a co-author and if the analysis was not performed by the CC, the final manuscript must be sent to CC for statistical review.

II. **Review of Paper Proposal by the P&P Committee**

   The P&P Committee shall review all formal proposals and make the following decisions:

1. **Approval** (or approval with recommendation), deferral, or disapproval (with reasons).

2. **Upon approval**, the paper is given an SHS Paper Approval Number.

3. **In the event a proposal does not receive full approval** (approved with recommendations or disapproved), the P&P Committee will supply the author with a complete explanation and recommendations for re-submission, when applicable.

4. **The decision of the P&P Committee** will be forwarded to the submitting author.
5. Along with an approval memo from the Chair, the author of each approved manuscript proposal will receive an Agreement for Data Distribution/Paper/Thesis Proposal form (an SHS author/investigator agreement must be signed by the author obtaining SHS data for a paper), Request for Data form, a Request for Data Analysis form, and a Data Analysis Monitoring System form (Data Request/Analysis forms are to be used by the author as needed). For maintaining better tracking, each form will be marked with the assigned SHS Paper Approval Number (see forms in Appendix 8 below). The author needs to complete, sign, and return the forms to the P&P Committee. CC (or the appropriate SHS PI) then provides required data to the authors. All primary authors must sign an agreement form before CC or the appropriate PI will provide the data.

6. The P&P Committee recommends that authors requesting data from the CC understand that a clear and concise rationale for data extraction is imperative. Representatives of the CC are well capable of streamlining the extraction of the database and analysis processes when supplied with this rationale.

7. If data analysis from the Coordinating Center (CC) is requested, the CC will assign a statistician to work with the primary author after the proposal is approved and all the required forms are returned to P&P Committee by the author. The paper may then be given a priority score if analyses are to be done by the CC. For those authors who choose to analyze their own data, CC representatives will be available for consultation.

8. For manuscripts written by investigators outside of SHS, the SHS PI co-author or the SHS PI who is closely affiliated with any of the authors must advise the P&P Committee during the review of the manuscript proposal whether the manuscript should be sent to the P&P Committee for review prior to submission to a journal. The Chair will send the paper to 2 or more reviewers, and the comments of the reviewers will be communicated to the submitting author. Additionally, if no SHS PI is a co-author and if the analysis was not performed by the CC, the final manuscript must be sent to CC for statistical review.

9. Prior to submission to a journal, the paper must be submitted by the author to NHLBI for review and to the IHS Area IRBs and the tribes for review and approval (see details in section IV below). Please note that as an integral part of the manuscript approval process, the IHS IRBs in the three centers require that all SHS manuscripts must contain the following disclaimer (verbatim): “The opinions expressed in this paper are those of the author(s) and do not necessarily reflect the views of the Indian Health Service.” A cover letter must be attached, requesting review and approval. The paper may not be submitted to a journal until the authors have received the NIH review (normally within one month of submission to NIH). The primary author is responsible for making sure that all Tribal/IHS approvals have been obtained prior to publication by contacting the responsible individual at each of the three field centers (see section IV below).

10. Minutes from the P&P Committee are circulated to the Steering Committee.
III. Analysis

If CC is responsible for the analysis, CC will assign a statistician to work with the author upon receiving the completed and signed "Request for Data Analysis Form" from the author. The statistician is the CC representative to the writing group. Whenever the workload for CC is heavy, CC will work with the investigators in analyzing the data according to the priority scores assigned by the P&P Committee.

Guidelines for authors to use in dealing with CC are:

1. Communicate with the CC representative on the writing group and discuss the objectives of the paper, appropriate statistical methods to be used, format of presentation (tables and figures), etc.

2. Determine a timetable with the CC representative. Be sure that analysis requests are made clearly and in writing (using the "Request for Data Analysis" form) and in a way that will allow sufficient time to complete the analyses.

3. If CC falls behind, the investigator should inform the P&P Committee; if there is a problem, deadlines can be changed.

4. For manuscripts written by investigators outside of SHS, if no SHS PI is a co-author and if the analysis was not performed by the CC, the final manuscript must be sent to CC for statistical review.

IV. Summary of Paper Publication Process

1. An author submits a paper proposal in standard format (see form in Appendix 8 below) to the P&P Committee Chair. (Note: the phrase "Strong Heart Study" should be included in the title and listed as a keyword whenever possible).

2. The P&P Chair notifies the author of the committee decision.

3. Prior to submission, all proposals must be approved by an SHS P.I. For manuscripts written by investigators outside of SHS, the SHS PI co-author or the SHS PI who is closely affiliated with any of the authors must advise the P&P Committee during the review of the manuscript proposal whether the manuscript should be sent to the P&P Committee for review prior to submission to a journal. The Chair will send the paper to 2 or more reviewers, and the comments of the reviewers will be communicated to the submitting author.
4. For manuscripts written by investigators outside of SHS, if no SHS PI is a co-author and if the analysis was not performed by the CC, the final manuscript must be sent to CC for statistical review.

5. Prior to submission to a journal, the paper must be submitted by the author to NHLBI for review (to be returned to the author within 1 month of submission) and to the IHS Area IRBs and the tribes with a lay summary and an attached cover letter requesting review and approval. These approvals are obtained through the following procedures:

   a. The primary author will first send the paper to the co-authors for their input. **When the primary author feels the paper is ready for NIH review and IHS Institutional Review Board (IRB) and Tribal approval, he/she will send a copy of the manuscript (including a Tribal/lay summary) simultaneously to the following with the clear designation that the paper is being sent for such approval:**

   1) Dakota Center: LaVonne Looking Elk  
      Strong Heart Study - Dakota Center  
      P.O. Box 9010  
      Rapid City, SD  57709  
      Phone: (605) 355-2377  
      Fax: (605) 355-2502  
      email: LaVonne.LookingElk@ihs.gov

   2) Oklahoma Center: Lee Keesee  
      Univ of Oklahoma Health Sciences Center  
      CHB 112  
      P.O. Box 26901  
      Oklahoma City, OK  73190  
      Phone: (405) 271-3090  
      Fax: (405) 271-4390  
      email: Lee-Keesee@ouhsc.edu

   3) Arizona Center: Nanette Taho  
      Aztec Building - Ste 250  
      1616 E Indian School  
      Phoenix, AZ  85016  
      Phone: (602) 277-0488  
      Fax: (602) 277-5979  
      email: Nanette.W.Taho@MedStar.net  
      with cc to: Marie Russell, MD  
      Director, MedStar Phoenix Field Center  
      email: Marie.Russell@MedStar.net
4) NHLBI: NHLBI has an electronic means for submission of manuscripts for NHLBI review, and authors are instructed to use this system for NHLBI REVIEW. Comments will be returned to the email address provided by the author in the submission process. All manuscripts need to be submitted to the following email address for NHLBI Review: ebpdocs@nhlbi.nih.gov

NOTE: Please cc Dr. Richard Fabsitz, Project Officer-Strong Heart Study, (FabsitzR@nhlbi.nih.gov) when emailing your manuscript to the above NHLBI email address.

The three individuals listed in 1-3 above are responsible for sending the manuscript for approval by Indian Health Service IRBs and the Tribes.

b. The author must include a Tribal/lay summary for all manuscripts, since such summaries are essential for obtaining Tribal and IHS IRB approval. The Tribal/lay summary should be no longer than one page of easily understandable text. One or two graphics illustrating major points could be included. Such summaries are critical to ensure tribal understanding of research results, and, hopefully, maintain tribal support for SHS research. The intended journal should be mentioned for all papers in the cover letter/memo.

c. The paper may not be submitted to a journal until the authors have received NIH review (see #4 above). Authors must check with the Oklahoma, Arizona, and Dakota Centers (see contact info in #1-3 above) to ensure that IHS IRB and Tribal approvals have been obtained; this should be done at the time when the author receives reviewers’ comments from the journal and is in the process of making final revisions. The primary author is responsible for making sure that all approvals have been obtained prior to publication.

d. The manuscript must include the following disclaimer (verbatim) (usually in the Acknowledgments or in a footer on the first page): “The opinions expressed in this paper are those of the author(s) and do not necessarily reflect the views of the Indian Health Service.”

The intention of this multi-step procedure is to ensure that all principal investigators are aware of the status of publications and also to ensure that appropriate review by NIH and approval by IHS and the Tribes occur prior to publication.

6. After the article is published, the primary author must send at least one reprint of the published article to the NHLBI Project Officer:

Richard Fabsitz, PhD
Project Officer-Strong Heart Study
Two Rockledge Center-Rm 8164
6701 Rockledge Dr. MSC
and to each of the three persons designated in the field centers (as listed above in #1-3), who
will then distribute the published articles to Tribes and IHS IRBs for their centers. The
primary author should also send reprints of the published article to all co-authors.

7. **NOTE:** Papers that are likely to result in press coverage or substantial press/media interest
require notice in advance to the NHLBI (contact Dr. Fabsitz) so that the staff and public
information office can be prepared.

8. The P&P Chair will maintain a list of published and in press SHS papers (posted on the SHS
website: [http://strongheart.ouhsc.edu](http://strongheart.ouhsc.edu)) and papers in various stages of preparation. In order
to help update the status of papers in the SHS publication list, authors are required to notify
the P&P committee by sending the cover letter each time when submitting their papers to the
NIH/IRBs and to a journal. Also, authors are required to notify the P&P when papers are
accepted by a journal for publication and when published. If using electronic transmission to
submit papers, authors need to copy Dr. Momotaz Begum (momotaz-begum@ouhsc.edu).

9. To track the progress of approved paper proposals, the P&P Committee distributes a status
survey of the approved papers by emailing a Paper Tracking Status Form every six months.
The authors must fill out the respective space regarding the progress/current status of their
paper(s) and return the form to the committee.

10. If the P&P Committee determines that progress on a manuscript is taking an unduly long
time, the Chair will communicate with the author, asking for a plan of action for completing
the paper or for the author(s) to release the topic for authorship by someone else.

11. In rare cases, the P&P Committee may need to make a recommendation to the Steering
Committee regarding reassignment of a paper topic.

**NOTE:** It must be recognized that any step of this approval process may entail requested
revisions and re-submissions by the authors.

V. **Approval of Abstracts** *(Please note that authors must submit a Lay Summary along with
the abstract, as required by the IHS IRB of the Dakota Center)*

1. It is assumed that all SHS abstracts will have at least one SHS PI as a co-author. The PI co-
author is responsible for ensuring that the abstract abides by SHS standards and guidelines.
If none of the PIs is a co-author, the abstract must be approved by the PI who works most
closely with the authors. The title of the abstract should include the phrase "Strong Heart Study" whenever possible.

2. Abstracts must be submitted for NHLBI review. NHLBI has an electronic means for submission of abstracts and manuscripts for NHLBI review, and, PRIOR TO submission of SHS abstracts to a conference, all authors MUST submit their materials using this NHLBI REVIEW system. Comments will be returned to the email address provided by the author in the submission process. All abstracts need to be submitted to the following email address for NHLBI review: ebpdocs@nhlbi.nih.gov

3. Abstracts must also be sent to the Dakota Center for approval by their IRB. (The Oklahoma and Arizona Centers do not have this requirement.) In addition to the abstract, please include a brief LAY SUMMARY of the work to be presented. Please specify that the abstract is being forwarded for Dakota Center IRB approval, include information about the meeting or other venue intended for the presentation, and send the abstract to:

   LaVonne Looking Elk  
   Strong Heart Study - Dakota Center  
   P.O. Box 9010  
   Rapid City, SD 57709  
   Phone: (605) 355-2377  
   Fax: (605) 355-2502  
   email: LaVonne.LookingElk@ihs.gov

4. Prior to presenting the paper, the presenting author should verify (if notice has not been received) that the NHLBI review and Dakota Center IRB approval have been received.

VI. Summary of Thesis/Dissertation Approval Process

1. A college student who wishes to use SHS data for a thesis or dissertation must submit a thesis proposal to the P&P Committee Chair. (See Thesis Proposal Form below in Appendix 8 below - also, the form can be downloaded – see SHS website: http://strongheart.ouhsc.edu)

2. The Thesis/Dissertation Proposal must include the Prospectus for the Doctoral Thesis/Dissertation or an Outline for a Masters/Bachelor Thesis. If a prospectus is not required by the doctoral degree program, the student needs to submit a detailed outline.

3. A thesis/dissertation proposal (see Appendix 8 below) must include: Title of Thesis/Dissertation, Name of Degree Candidate, Type of Degree, Candidate Affiliation including the contact information (full address, telephone, fax and email) and name of the Primary Mentor, including the same type of contact information.
4. Upon approval, the thesis/dissertation is given an SHS Thesis/Dissertation Number, and the P&P Chair notifies the student of the committee decision. The student is provided with the Agreement for Data Distribution/Paper/Thesis/Dissertation Proposal form, Request for Data form, and a Data Analysis Monitoring System form to complete, sign, and return to the P&P Committee (see forms in Appendix 8 below). CC (or the appropriate SHS PI) then provides required data to the student.

5. As part of the agreement, the student agrees to write at least one paper based on the approved thesis/dissertation proposal. At the time the student is ready to develop a paper for publication, the student must submit a separate paper proposal to the P&P Chair and follow all of the P&P paper approval procedures described above.

VII. Forms for Paper and Thesis/Dissertation Proposals

Appendix 8 below contains the desired formats for paper and thesis/dissertation proposals submitted to the P&P Committee. Also, the forms can be downloaded from the Internet – see SHS website: http://strongheart.ouhsc.edu. Additionally, upon receiving requests from the SHS authors, these forms will be transmitted electronically by email. For the electronic forms, email or word processing software may be easily implemented for form completion and submission. "Cut and Paste" or other electronic means may be used to download the proper form, to fill it in (electronically expanding the space as much as needed for each section), and to submit the form to the P&P Chair by email, or more traditional means if desired. An electronic file containing the SHS Publication Policy will also be included with the proposal form to make the prospective authors aware of the rules and procedures of the SHS P&P Committee.

The SHS P&P paper, thesis/dissertation, and ancillary study proposal forms (see Appendix 8 of this Volume) are:

1. Strong Heart Study Paper Proposal
2. Sample of paper proposal approval Memo
4. Sample of thesis/dissertation proposal approval Memo
6. Strong Heart Study Request For Data
7. Strong Heart Study Request For Data Analysis
8. Strong Heart Study Data Analysis Monitoring System
9. Strong Heart Study Ancillary Study Proposal Form

10. Strong Heart Study Data and/or Materials Distribution Agreement Form

11. Strong Heart Study Annual Update on Ancillary Study
1.8. ANCILLARY STUDIES POLICY

(SHS home page: http://strongheart.ouhsc.edu/)

1.8.1 General Policy

To enhance the value of the Strong Heart Study (SHS) and to ensure the continued interest of the investigators, the Steering Committee (SC) welcomes proposals from individual investigators to carry out ancillary studies and to promote the advancement of science. Nevertheless, to protect the integrity of SHS and the privacy of its participants, such ancillary studies, before their inception, must be reviewed and approved by the SC and by the NHLBI through its SHS Observational Study Monitoring Board (OSMB). In general, ancillary studies require outside (non-SHS) funding.

1.8.2 Definition of Ancillary Study

The SHS Steering Committee has defined an ancillary study as a project that imposes additional burdens on the SHS participants, or is outside of the goals of SHS, or has commercial aspects (patents, profit, etc). Ancillary studies require tribal and IRB approvals and separate consent forms. An SHS substudy is one that is consistent with the goals of SHS as stated in the consent form, involves no additional participant burden, and has no commercial aspects. For some SHS centers, the investigators submit protocol modification paperwork to their IRBs regarding substudies. The tribes are informed of substudies through the SHS newsletter, community meetings, or similar means.

An ancillary study is one based on information from SHS participants in an investigation that is not described in the SHS protocol and involves data collection or data analyses that are not included as part of the routine SHS dataset or data analyses. The core Strong Heart Study includes the use of blood, DNA, and urine stored for additional studies not described within the original protocol, but within the scope of the participant consents and approved by the SC; these are considered to be SHS substudies, not ancillary studies. In general, ancillary studies require external (non-SHS) funding. Funding must cover any costs incurred by the SHS lab(s) (Penn Medical Laboratory (PML) and/or the Southwest Foundation for Biomedical Research (SWF), e.g., cost reimbursement for sample handling & shipping), the Cornell Ultrasound/EKG Reading Center (RC) (e.g., any customized selection or reading of clinical material), and by the Coordinating Center (CC) (e.g., for tasks such as sample selection, preparing and documenting analysis files, participating in statistical analysis, and integrating the new ancillary data back into the combined SHS database). No funds for this purpose are available within the Study.

1.8.3 Requirements for Approval of an Ancillary Study

An ancillary study must receive SHS/OSMB approval (see 1.8.05 below) before a grant application to support it is submitted. Approval will be based on finding that the ancillary study will have scientific merit but will not do any of the following:

a. Interfere with the completion of the main objective of SHS.
b. Adversely affect participant cooperation.

c. Create a serious diversion of study resources (personnel, equipment or study samples), either locally or centrally.

d. Jeopardize the public image of SHS and/or the Study relationship with the tribes.

e. Use SHS grant resources without reimbursement.

1.8.4 Preparation of Request for Approval of an Ancillary Study

For approval of an ancillary study, a written request on the SHS Ancillary Study Proposal Form (see Appendix 8 below) must be submitted to the Steering Committee (via the SC Chair, Barbara V. Howard, PhD). The Ancillary Study Proposal Form collects the following information:

a. Identifiers:
   - Title of proposal
   - Initiating investigators (with PI contact information) and collaborators
   - Planned starting date
   - Funding plans and estimated cost

b. Design and methods:
   - Brief background and rationale
   - Study questions or hypotheses
   - Methods, data to be collected
   - Proposed statistical analyses
   - List of all analytes to be measured using SHS biological samples (for novel analytes, document that the (within person plus laboratory variability)/(total variability) is acceptably low)
   - Analyte assay methods to be used
   - Burden on SHS participants
   - Burden on SHS CC, PML, SWF, RC, and Field Center Staffs - summary of tasks involved for each of the SHS centers and how each would be reimbursed (e.g., by subcontracts, with amounts approved by the participating SHS centers)
   - Impact on the main Study and potential utility of the new data for collaboration with other investigators

c. Data or specimen requirements:
   - Data needed from SHS datasets
   - Specimens needed from SHS repositories, specifying SHS Phase # (I, II, etc), type, and amount
   - Address whether previously thawed specimens are adequate

d. Handling of SHS data and specimens:
   - Disposition of stored samples from main Study and those processed by ancillary study
   - Disposition of ancillary study data at the conclusion of the ancillary study
1.8.5 Review of Ancillary Study Proposals

The Steering Committee, often in consultation with the SHS Sample Committee, will review and will approve, reject or request modification of ancillary study proposals in a timely manner (generally 8 weeks plus the time needed for OSMB review (in case of participant burden) and/or IRB review (e.g. for use of stored samples)). Ancillary studies using stored biological samples must be recommended for approval by the SHS Sample Committee. Exceptions to the need for OSMB approval may be granted by the SHS SC Chair in case of studies with no participant risk or burden.

At least one SHS investigator must be included as a co-investigator in each proposal. This investigator, collaborating with the ancillary study PI, will facilitate preparation of the ancillary study proposal, its submission to the SHS SC, and subsequent communications between the collaborating studies. Other SHS investigators may request to become collaborators on a proposal. The key criteria for approval of proposals are scientific merit and impact on SHS. In addition, the plan for reimbursing SHS components for all ancillary study-related costs must be adequate.

Formal IRB approval will be required, if such studies require further interaction with SHS participants (e.g., interviews or additional procedures). The principal investigator (PI) of the ancillary study, working with the three SHS field centers, is responsible for obtaining approval from the Indian communities, the grantee institution IRBs, and the three area IHS IRBs.

Proposals related to cardiovascular and pulmonary diseases and their risk factors, which include measurements (even of stored samples) that are not specifically described in the original SHS protocol must obtain approval from some SHS IRBs. If the SHS Steering Committee feels that the ancillary study will result in a major change in the protocol, the PI will be required to seek IRB approval prior to conducting the study. Any ancillary study that is not related to cardiovascular or pulmonary diseases or their risk factors will require IRB and tribal approval.

1.8.6 Amendments of Ancillary Study Proposals

Amendments to ancillary study proposals (e.g., adding analytes to be measured) require approval via submission of a revised proposal with a note describing the changes. It should be noted that such amendments may require further review and approval by the SHS IRBs.

1.8.7 Yearly Progress Report for Ancillary Study

Following approval of an ancillary study, a yearly progress report must be submitted to the Coordinating Center PI (Elisa T. Lee, PhD). CC will include these annual progress reports in its annual report to the OSMB regarding overall progress of SHS. The annual ancillary study progress report should follow the format contained in the Strong Heart Study Annual Update on Ancillary Study form (see Appendix 8 below).

1.8.8 Analysis and Publication of Results of Ancillary Studies
The goals of this policy are to provide participant protection (ensure use of data does not exceed informed consent), coordination of efforts to avoid duplication of work, and to minimize barriers to publication of Ancillary Studies.

The PI or other representative of the ancillary study, and if necessary the SHS SC, will consult with the CC during data analysis to ensure that all study data used in analysis of ancillary results are consistent with data in the main SHS database. Manuscript proposals must be approved in advance by the SHS Publications and Presentations Committee (P&P). This procedure is necessary to establish authorship and prevent overlap in the publication effort. Approval of manuscript proposals is sought by submitting the proposal using standard SHS format (see SHS Paper Proposal form in Appendix 8 below) to the P&P. The ancillary study PI will be required to sign an Ancillary Study Proposal Form (see form in Appendix 8 below). This agreement stipulates that the ancillary study investigators agree to submit paper proposals for approval by the SHS P&P and to submit draft manuscripts for review by the NHLBI and approval by the IHS IRBs and the tribes (see section 1.7 above). Additionally, abstracts for presentations at meetings require review by the NHLBI and approval by the Dakota Center IHS IRB (see section 1.7 above). The investigator who assumes lead responsibility for the ancillary study shall generally be listed as an author. Whenever possible, the phrase "Strong Heart Study" should be included in the manuscript title and listed as a key word. Manuscripts will also contain an acknowledgment section listing individual SHS investigators and staff as deemed appropriate. Upon publication, reprints must be distributed as specified above in section 1.7.

1.8.9 Feedback of Results of Ancillary Studies to Participants

Results of ancillary studies shall be reported to participants and/or their physicians if such reporting is medically useful and approved by the relevant IRBs and SHS. Once approved, such reporting should follow standard SHS protocol for notification of participants. Overall results of ancillary studies shall be reported to participating tribes via lay language articles in the SHS Newsletter and/or by oral presentations of results at tribal meetings.

1.8.10 Handling of SHS Data and Specimens

At the time of distributing SHS specimens and/or data, the SHS Collaborating Investigator, with help from SHS CC and Lab (PML and/or SWF), makes explicit arrangements with the ancillary study PI for:

1. security of these study materials
2. completion of the SHS Ancillary Study Proposal Form and the SHS Data and/or Materials Distribution Agreement Form (See Appendix 8 below)
3. documentation of IRB approval
4. final disposition of study materials at the conclusion of the ancillary study

The safety and confidentiality of the SHS data at the collaborating institution are the responsibility of the ancillary study PI, as is the appropriate disposition of data and remainders of SHS samples after the ancillary study has been completed. Leftover DNA and any other types of laboratory specimens must be returned to PML or SWF. Files of SHS data must be returned or deleted, as established and agreed at the outset of the collaboration. An archival copy of the
newly collected data and/or laboratory results must be sent in a secure manner to the SHS CC one year after the conclusion of the data cleaning and closure or one year after acceptance of the primary publication, whichever comes first. This should allow sufficient time for publication of the main (ancillary) study hypothesis. This transfer is the responsibility of the SHS representative(s) collaborating with the ancillary study. The data from the ancillary study will be included in the SHS dataset for distribution according to procedures currently under negotiation with the involved tribes, the IHS, and the NHLBI.

The SHS Steering Committee monitors the development of the ancillary studies, receipt of funding, initiation dates, and progress. A written progress report on ancillary studies is to be made annually to the SHS CC, who will include the summary in the annual report to the SHS OSMB (NHLBI). This annual report should include a list of data collected and/or analytes measured. For the convenience of the collaborators, a shell document for these reports (SHS Annual Update on Ancillary Study form) is provided in Appendix 8 below.

The ancillary study PI will send the completed SHS Data and/or Materials Distribution Agreement Form to the SHS Coordinating Center PI (Dr. Lee) (see contact info immediately below). The CC will review the agreement, sign the agreement on behalf of SHS, and forward the agreement to the SHS NHLBI Program Officer (Richard Fabsitz, PhD). A file copy with all required signatures will be retained by CC, and a copy will be returned to the Ancillary Study Principal Investigator.

Elisa T. Lee, PhD, SHS PI and CC Director
SHS Ancillary Study Correspondence
Center for American Indian Health Research
College of Public Health
University of Oklahoma Health Sciences Center
PO Box 26901 - Room CHB-112
Oklahoma City, OK 73190

Express Svc:
Center for American Indian Health Research
College of Public Health
801 NE 13th St, Room CHB-100
Oklahoma City, OK 73104

Phone: 405-271-3090
Fax: 405-271-4390
Email: Elisa-Lee@ouhsc.edu

1.8.11 Ancillary Studies Using DNA or Other Stored Samples

SHS represents a unique public resource to be used by the American Indian communities in conjunction with clinical, public health, and scientific entities to better understand the etiology and epidemiology of cardiovascular and pulmonary diseases and their risk factors, and clinical
sequelae. The SHS investigators are committed to managing the stored biologic material for the good of this endeavor in the manner agreed to and expected by the participating tribes and Study participants. This resource includes blood, DNA, and other biological samples obtained from the SHS participants and stored at PML or SWF for future studies of scientific merit related to cardiovascular and pulmonary diseases and their risk factors that are proposed by SHS or collaborating investigators.

With respect to use of DNA, proposals will need to be as specific as possible, describing the genetic hypothesis of interest, the specific genes or chromosomal sequence to be analyzed, the laboratory method, the primary dependent variable (if applicable), endpoint or risk factor of interest, preferred sampling design, and sources of funding. Strong Heart Study DNA samples are maintained by PML and the SWF. PML maintains DNA samples from the original cohort collected in Phases I through III. SWF maintains DNA samples on participants in the Strong Heart Family Study (SHFS) collected in phases III, IV, and V. For studies requiring genotyping of polymorphisms, STRONG preference will be given to those investigators who agree to have PML (through Children’s Hospital and Dr. Joseph Devaney) and/or SWF perform the genotyping (as opposed to studies requesting DNA for genotyping at the collaborating institution). For genotyping polymorphisms at collaborating institutions, if the identity of the variant is known a priori, it should also be included in the proposal. If the identity of the variant is not known a priori, such information should be transmitted to the SHS Steering Committee for approval prior to genotyping, and certainly before data analysis and publication. The CC maintains a database of single nucleotide polymorphisms (SNPs) typed (or being typed) on SHS cohort participants, and SWF maintains the same for SHFS participants. Interested investigators must contact Dr. Lyle Best, Chair of the SHS Sample Committee or Dr. Shelley Cole of SWF to inquire about particular SNPs:

Lyle Best, M.D.
Principal Investigator, Strong Heart Study-Dakota Center
Chair, SHS Sample Committee
PO Box 9010
Rapid City, SD 57709

Express Service, change last two lines to:
3200 Canyon Lake Drive
Rapid City, SD 57702
Office: (605) 355-2401
FAX: (605) 355-2502
E-mail: sbest@utma.com

Shelley A. Cole, Ph.D.
SHS Co-Investigator
Director, Genetics Core Laboratory
Southwest Foundation for Biomedical Research
P.O. Box 760549
San Antonio, TX 78245-0549
Express Service, change last two lines to:
7620 NW Loop 410
San Antonio, TX 78227-5301
Office: (210) 258-9688
FAX: (210) 670-3344
E-mail: scole@darwin.sfbr.org

In general, all costs attributed to this ancillary study are the responsibility of the originating investigator. The proposal will be reviewed by the SHS Sample Committee to assess scientific merit and possible overlap with existing activities. The Sample Committee will recommend approval or disapproval to the SC. In the event that a study is disapproved, the investigator will be notified of the reason for the decision.

When a study is approved, the SHS CC has the responsibility of generating a list of SHS participant IDs to be included, which is consistent with the approved design and objectives of the ancillary study. In general, it is better for ancillary studies to take advantage of case-control, case-cohort, and other contrasts that have already been generated and investigated for other analyses or hypotheses. In addition, preference will be given to proposals focusing on polymorphisms with documented functional significance. Due to the limited resources of PML and SWF that can be devoted to servicing requests of collaborating studies, it is STRONGLY suggested that investigators request genotyping by PML (though Children’s Hospital) and/or SWF as opposed to requesting DNA samples for genotyping at the collaborating institution. In this way, the work can be carried out quickly and efficiently without wasting DNA and time spent on the aliquoting, shipping, genotyping, and monitoring the return of sample remainders to SWF, and logging and re-storage of DNA. The resulting genotype data will be provided to the investigator by PML and/or SWF, and the other SHS data needed to address the approved hypotheses will be provided by CC. There should be no loss of the originating investigator’s proprietary (if any) or publication rights. In the rare instances when SC approves genotyping by a collaborating lab, PML or SWF will aliquot DNA into 96-well plates. The amount of DNA to be supplied will be determined by the SHS Sample Committee. In general, no more than 50 ng will be provided for typing six to ten polymorphisms.

All costs of the approved ancillary DNA study are the responsibility of the initiating investigators. SWF and PML (through Children’s Hospital) will work with the investigators to supply accurate information about charges for genotyping and other aspects of needed support, and such charges will closely reflect best estimates for actual costs to be incurred. Any sub-contractual arrangements will need to be made in coordination with NHLBI staff, the involved SHS Centers, and the participating collaborating institution(s). Resulting data from the ancillary study must be made available in a timely manner to the SHS CC, as specified above. In this way the value of SHS resources will continue to grow as the foundation database enlarges in size and scope, and analyses can be verified when necessary.
APPENDIX 1

THE STRONG HEART STUDY V
CARDIOVASCULAR DISEASE IN AMERICAN INDIANS

PRINCIPAL AND CO-INVESTIGATORS

Arizona Center and Core Laboratory

Barbara Howard, Ph.D.
Principal Investigator
MedStar Research Institute
6495 New Hampshire Ave., Ste. 201
Hyattsville, MD 20783
Office: (301) 560-7302
FAX: (301) 560-7307
Pager & Cell: (301) 602-0125
Phoenix Office: (602) 277-0488
E-mail: Barbara.V.Howard@medstar.net

Marie Russell, M.D.
Coordinator, Field Studies
The Aztec Building – Strong Heart Study
1616 E. Indian School, Ste. 250
Phoenix, AZ 85016
Office: (602) 277-0488
FAX: (602) 277-5979
Cell: (602) 363-7554
E-mail: Marie.Russell@Medstar.net

Jason G. Umans, M.D., Ph.D.
Director, Penn Medical Laboratory
MedStar Research Institute
108 Irving Street, NW, Annex 2
Washington, DC 20010-2933
Office: (202) 877-7352
FAX: (202) 877-7342
E-mail: Jason.G.Umans@MedStar.net

Dakotas Center

Lyle Best, M.D.
Principal Investigator
Strong Heart Study-Dakota Center
PO Box 9010
Rapid City, SD 57709
Express Service, change last two lines to:
3200 Canyon Lake Drive
Rapid City, SD 57702
Office: (605) 355-2401
FAX: (605) 355-2502

William James Howard, M.D.
Adjudicator, Mortality Review
Vice President for Academic Affairs
Washington Hospital Center
110 Irving Street, NW, Rm. 6A-126
Washington, DC 20010
Office: (202) 877-5285
FAX: (202) 877-8024
E-mail: Wm.James.Howard@medstar.net

Home Address/Info:
R.R. 1, Box 88
Rolette, ND 58366
(For Express Service add: Missouri Breaks
Industries Research Inc., 1 Airport Rd. to
address)
Home: (701) 246-3884
FAX: (701) 246-3698 (notify by phone before
sending fax)
E-mail: sbest@utma.com

Helaine Resnick, Ph.D.
Co-Investigator
Director, Department of Epidemiology
MedStar Research Institute
6495 New Hampshire Ave., Ste. 201
Hyattsville, MD 20783
Office: (301) 560-7615
FAX: (301) 560-7615
E-mail: Helaine.E.Resnick@medstar.net

Jeffrey A. Henderson, M.D., M.P.H.
Co-Investigator
President and CEO
Black Hills Center for American Indian Health
701 St. Joseph Street, Suite 204
Rapid City, SD 57701
Office: (605) 348-6100  
FAX: (605) 348-6990  
E-mail: jhenderson@bhcaih.org

Thomas K. Welty, M.D., M.P.H.  
Co-Investigator  
Strong Heart Study-Dakota Center  
5990 East Jeremy Lane  
Flagstaff, AZ 86004  
Home: (928) 526-0955  
Cell: (208) 989-0340  
FAX: (928) 526-9059  
E-mail: twelty@earthlink.net

Ellie Zephier, R.D., M.P.H.  
Principal Investigator  
Strong Heart Dietary Study  
Acting Diabetes Consultant  
Area Diabetes & Dietetics Program  
Aberdeen Area Indian Health Service  
Federal Building, Rm. 309  
115 4th Avenue, SE  
Aberdeen, SD 57401  
Office: (605) 226-7238  
FAX: (605) 226-7733  
E-mail: ellie.zephier@ihs.gov

Oklahoma Center and Coordinating Center

Elisa T. Lee, Ph.D.  
Principal Investigator  
Center for American Indian Health Research  
University of Oklahoma Health Sciences Center  
P. O. Box 26901, Rm. CHB100  
Oklahoma City, OK 73190  
Express Service, change last two lines to:  
801 NE 13th St., Rm. CHB100  
Oklahoma City, OK 73104  
Office: (405) 271-3090  
FAX: (405) 271-4390  
E-mail: elisa-lee@ouhsc.edu

ECG and Ultrasound Reading Center

Richard B. Devereux, M.D.  
Principal Investigator  
Director, Laboratory of Echocardiography  
Weill Medical College of Cornell University  
New York – Presbyterian Hospital  
525 East 68th Street, Rm. K415 (Box 222)  
New York, NY 10021  
Office: (212) 746-4655  
FAX: (212) 746-8561  
E-mail: rdevere@med.cornell.edu or  
Home: cdevereux@mem.po.com

Mary J. Roman, M.D.  
Co-Investigator
SHS Family Study Center

Jean W. MacCluer, Ph.D.
Principal Investigator
Department of Genetics
Southwest Foundation for Biomedical Research
P.O. Box 760549
San Antonio, TX 78245-0549
Express Service, change last two lines to:
7620 NW Loop 410
San Antonio, TX 78227-5301
Office: (210) 258-9490
FAX: (210) 670-3317
E-mail: jean@darwin.sfbr.org

Laura Almasy, Ph.D.
Co-Investigator
Department of Genetics
Southwest Foundation for Biomedical Research
P.O. Box 760549
San Antonio, TX 78245-0549
Express Service, change last two lines to:
7620 NW Loop 410
San Antonio, TX 78227-5301
Office: (210) 258-9444
FAX: (210) 670-3317
E-mail: xwang@darwin.sfbr.org

NHLBI Program Office

Richard R. Fabsitz, Ph.D.
Project Manager
APPENDIX 3

THE STRONG HEART STUDY V
CARDIOVASCULAR DISEASE IN AMERICAN INDIANS

ORGANIZATIONAL STRUCTURE

STEERING COMMITTEE

Chairperson: Barbara V. Howard, Ph.D., Principal Investigator – Arizona Center
Members: Lyle Best, M.D., Principal Investigator – Dakota Center
Linda D. Cowan, Ph.D., Co-Investigator – Oklahoma Center
Richard B. Devereux, M.D., Principal Investigator – ECG and Carotid Ultrasound Reading Center
Richard R. Fabsitz, Ph.D., Project Manager – NIH/NHLBI/DECA
Jeffrey Henderson, M.D., Co-Investigator – Dakota Center
Elisa T. Lee, Ph.D., Principal Investigator – Oklahoma Center
Jean W. MacCluer, Ph.D., Principal Investigator – Genetics Center
Helaine Resnick, Ph.D., Co-Investigator – Arizona Center
Everett R. Rhoades, M.D., Co-Investigator – Oklahoma Center
Marie Russell, M.D., Director – Phoenix Field Center
Jason G. Umans, M.D., Ph.D., Co-Investigator – Director, Core Lab
Thomas K. Welty, M.D., Co-Investigator – Dakota Center
Jeunliang Yeh, Ph.D., Co-Investigator – Coordinating Center
Ellie Zephier, RD, MPH, Co-Investigator – Dietary Study Center
APPENDIX 4

THE STRONG HEART STUDY V
CARDIOVASCULAR DISEASE IN AMERICAN INDIANS

SUBCOMMITTEES

Data Committee
Linda D. Cowan, Ph.D.
Richard B. Devereux, M.D. – Chair
Richard R. Fabsitz, Ph.D.
Barbara V. Howard, Ph.D.
Elisa T. Lee, Ph.D.
Helaine E. Resnick, Ph.D. – Co-Chair
Jeunliang Yeh, Ph.D.
Ying Zhang, M.D., Ph.D.

Ethics Committee
Sonja Antone
Lyle Best, M.D.
Jeffrey A. Henderson, M.D., M.P.H.
Bert Lewis
Jean W. MacCluer, Ph.D.
Everett R. Rhoades, M.D. – Chair

Genetics Committee
Lyle Best, M.D.
Shelley A. Cole, Ph.D.
Jeffrey A. Henderson, M.D., M.P.H.
Jean W. MacCluer, Ph.D. – Chair
Everett R. Rhoades, M.D.
Francine Welty, M.D.

Infectious Disease Committee
Lyle Best, M.D. – Chair
Linda D. Cowan, Ph.D.
Richard Devereux, M.D.
Jeffrey A. Henderson, M.D., M.P.H.
Thomas K. Welty, M.D., M.P.H.

Morbidity Review Committee

Reviewers
Jonathan Bella, M.D.
Lyle Best, M.D.
Richard B. Devereux, M.D. – Chair
Richard Rodeheffer, M.D.
Marie Russell, M.D.

Stroke Morbidity
Jorge Kizer, M.D.

Mortality Review Committee

First Review
Maurice Sievers, M.D.

Second Review
James M. Galloway, M.D.
Jeffrey A. Henderson, M.D., M.P.H.
Dorothy A. Rhoades, M.D., M.P.H.
Everett R. Rhoades, M.D.
Thomas K. Welty, M.D., M.P.H.

Adjudicator
Wm. James Howard, M.D.

Stroke Mortality
Jorge Kizer, M.D.

Coordinating Center
Linda D. Cowan, Ph.D.
Amir Butt, M.D.
Jeunliang Yeh, Ph.D.
Nutrition Committee
Sigal Eilat-Adar, Ph.D., R.D.
Barbara V. Howard, Ph.D.
Catherine Loria, Ph.D.
Claudia Mattil, M.S., R.D., C.C.R.C-Chair
Helaine Resnick, Ph.D.
Susan Xu, Ph.D.
Ellie Zephier, M.P.H., R.D.

Quality Control Committee
Tauqeer Ali, M.D., Ph.D.
Richard R. Fabsitz, Ph.D.
Callen Hull, R.N.
Marcia O’Leary, R.N.
Jason G. Umans, M.D., Ph.D.
Jeunliang Yeh, Ph.D. – Chair

Psychosocial Committee
Jan Beals, Ph.D.
Lyle Best, M.D.
Gary Leonardson, Ph.D.
Marcia O’Leary, R.N.
Linda Poolaw
Everett R. Rhoades, M.D.
Thomas K. Welty, M.D., M.P.H. – Chair

Renal Disease Committee
Andrew Narva, M.D.
Jason G. Umans, M.D., Ph.D. - Chair
Thomas K. Welty, M.D., M.P.H.
Susan Xu, Ph.D.
Jeunliang Yeh, Ph.D.

Publications and Presentations Committee
Lyle Best, M.D.
Richard R. Fabsitz, M.A.
Barbara V. Howard, Ph.D.
Elisa T. Lee, Ph.D. - Chair
Jean W. MacCluer, Ph.D.
Mary J. Roman, M.D.

Sample Committee
Lyle Best, M.D. – Chair
Shelley A. Cole, Ph.D.
Linda D. Cowan, Ph.D.
Jean W. MacCluer, Ph.D.
Jason G. Umans, M.D., Ph.D.
APPENDIX 5
THE STRONG HEART STUDY V
CARDIOVASCULAR DISEASE IN AMERICAN INDIANS

OTHER KEY PERSONNEL

Arizona
Field Center
The Aztec Building – Strong Heart Study
1616 E. Indian School, Ste. 250
Phoenix, AZ 85016
Phone: (602) 277-0488
FAX: (602) 277-5979

Callen Hull, R.N.
Research Nurse
E-mail: Callen.P.Hull@Medstar.net

Bert Lewis
Personnel Manager
Cell: (602) 363-9405
E-mail: Allen.B.Lewis@MedStar.net

Mary E. Rybka, RHIT
Outcomes Coordinator-Arizona
E-mail: Mary.E.Rybka@MedStar.net

Maurice Sievers, M.D.
Consultant, Mortality Surveillance

Nanette Taho
SHS Phlebotomist
E-mail: Nanette.W.Taho@MedStar.net

Sharon Taho
Sr. Administrative Assistant
E-mail: Sharon.E.Taho@MedStar.net

Damon Davis, R.N.
SHS Field Nurse/SANDS Coordinator
E-mail: Damon.W.Davis@medstar.net

Tanya Molina
Clinical Research Assistant/SHS Phlebotomist
E-mail: Tanya.R.Molina@medstar.net

Bernadette Cooper
SHS Field Recruiter/Clerk
E-mail: Bernadette.M.Cooper@medstar.net

Kristina Thomas
Data Entry Clerk
E-mail: Kristina.K.Thomas@medstar.net

Rosinna Briones
Sonographer
E-mail: Rosinna.Briones@MedStar.net

Christina Chiago, L.P.N., Coordinator
NIH-NIDDK
Hu Hu Kam Hospital
Sacaton, AZ 85247
Office: (602) 528-1225

Laboratory Center
Penn Medical Laboratory
MedStar Research Institute
108 Irving Street NW, Annex 2
Washington, DC 20010-2933
FAX: (202) 877-7342

Jianhui Zhu
Technical Director
Office: (202) 877-5630
FAX: (202) 877-7342
E-mail: Jianhui.Zhu@medstar.net

Sophia Rushton-Reid, M.T. (ASCP)
Research Study Coordinator/Clinical Research Associate, PML
Office: (202) 877-8379
FAX: (202) 877-7342
E-mail: Sophia.Rushton-Reid@medstar.net

MedStar Research Institute
MedStar Research Institute
6495 New Hampshire Ave., Ste. 201
Hyattsville, MD 20783
Main Phone: (301) 560-7395
FAX: (301) 560-7307

Angela Silverman, M.S.N., CANP
Administrative Director
Office: (301) 560-7304
E-mail: Angela.Silverman@medstar.net

Sabrina Smith
AVP, Office of Contracts and Grants
Management
Office: (301) 560-7383
FAX: (301) 560-7388
E-mail: Sabrina.L.Smith@medstar.net

Michele Clements, B.A.
Executive Assistant
Office: (301) 560-7303
FAX: (301) 560-7307
E-mail: Michele.L.Clements@medstar.net

Dakotas Center
Strong Heart Study
Missouri Breaks Industries Research, Inc.
HCR 64, Box 52
Timber Lake, SD 57656
(Address good for Express Service)
Office: (605) 964-3418 or 964-3419
FAX: (605) 964-3415

Timber Lake Area
Marcia O’Leary, R.N.
Coordinator
E-mail: mol@mbiri.com

Marie Gross, R.N.
Morbidity/Mortality Coordinator
Office: (605) 865-3526
FAX: (605) 865-3615
E-mail: mgross@mbiri.com

Jason “Jay” Kunf
Data Manager
E-mail: jdkunf@mbiri.com

Sue Sherwood
Administrative Officer
E-mail: ssherwood@mbiri.com

Eagle Butte Area
Eagle Butte Clinic
239 Elm Street
Eagle Butte, SD 57625
Office: (605) 964-1260
FAX: (605) 964-1263

Wendy Lawrence, R.N.
Supervisor
E-mail: wlawrence@mbiri.com

Lillian Brown
Senior Recruiter

Kendra Enright
Research Assistant

Melissa Kunf, R.N.
Ultrasound Technician
E-mail: makunf@mbiri.com

Millie Afraid of Hawk, R.N.
Research Assistant
E-mail: mafraidofhawk@mbiri.com

Kyle Clinic
Dakota SHS
Lakota Fund Building, Suite 105
P.O. Box 515
Kyle, SD 57752
Office: (605) 455-2955
FAX: (605) 455-2568

Laurie Bickel, R.N.
QA/QC
Supervisor, Pine Ridge Clinic
E-mail: lbickel@mbiri.com

Lois Bettelyoun
Recruiter/Interviewer
E-mail: lbettelyoun@mbiri.com

Mary Merrival
Recruiter

Earline “Arie” Shiroma
Research Assistant
E-mail: ashiroma@mbiri.com

Rapid City Area
Strong Heart Study – Dakota Center
PO Box 9010
Rapid City, SD 57709-9010
FAX: (605) 355-2502

Express Service, change last two lines to:
3200 Canyon Lake Drive
Rapid City, SD 57702
FAX: (605) 355-2502
LaVonne J. Looking Elk
Executive Assistant
Office: (605) 355-2377
E-mail: LaVonne.LookingElk@ihs.gov

Sheila Romero
Research Assistant
Phone: (605) 355-2346
E-mail: Sheila.Romero@ihs.gov

Oklahoma Center (OC) and Coordinating Center (CC)

Oklahoma City Area:
Center for American Indian Health Research
University of Oklahoma Health Sciences Center
P.O. Box 26901, Rm. CHB100
Oklahoma City, OK 73190
Express Service, change last two lines to:
801 N.E. 13th Street, Rm. CHB100
Oklahoma City, OK 73104
FAX: (405) 271-4390

Momotaz Begum, M.P.H., M.B.B.S.
P&P Coordinator
Office: (405) 271-3090 x46734
E-mail: momotaz-begum@ouhsc.edu

Jill Black
Sr. Administrative Manager
Office: (405) 271-3090 x46726
E-mail: jill-black@ouhsc.edu

Amir Butt, M.D.
SHS Morbidity & Mortality Coordinator
Office: (405) 271-3090 x46778
E-mail: amir-butt@ouhsc.edu

Debra Gates
Data Manager
Office: (405) 271-3090 x46724
E-mail: debra-gates@ouhsc.edu

Lee Keesee
Administrative Secretary
Office: (405) 271-3090 x46712
E-mail: lee-keesee@ouhsc.edu

Karen Kimbley
Data Abstractor
Office: (405) 271-3090 x46732
E-mail: karen-kimbley@ouhsc.edu

Carl F. Schaefer, Ph.D.
Office: (405) 271-3090 x46725
E-mail: carl-schaefer@ouhsc.edu

Martha Stoddart, M.S.
Quality Control Analyst
Office: (405) 271-3090 x46721
E-mail: martha-stoddart@ouhsc.edu

Mushfiq Tarafder
Graduate Research Assistant
Office: (405) 271-3090 x46887
E-mail: mushfiq-tarafder@ouhsc.edu

Wenyu Wang, Ph.D.
Senior Statistician
Office: (405) 271-3090 x46723
E-mail: wenyu-wang@ouhsc.edu

Yiming Wang, M.S.
Webmaster
Office: (405) 271-3090 x46728
E-mail: yiming-wang@ouhsc.edu

Susan Xu, Ph.D.
Statistician
Office: (405) 271-3090 x46722
E-mail: susan-xu@ouhsc.edu

Fawn Yeh, Ph.D.
Epidemiologist
Office: (405) 271-3090 x46733
E-mail: fawn-yeh@ouhsc.edu

Ying Zhang, Ph.D.
Statistician
Office: (405) 271-3090 x46731
E-mail: ying-zhang4@ouhsc.edu

Anadarko Area:
Strong Heart Study
Anadarko Indian Health Clinic
Box 828, 115 NE Old Town Dr.
Anadarko, OK 73005
FAX: (405) 247-6653
Stephanie Gomez, L.P.N.
Medical Assistant
Office: (405) 247-9053
E-mail: stephanie-gomez@ouhsc.edu

Linda Poolaw
Tribal Liaison
Office: (405) 247-2458 x249
E-mail: linda-poolaw@ouhsc.edu

Donna Smith
Tribal Liaison
Office: (405) 247-9053
E-mail: donna-l-smith@ouhsc.edu

**Lawton Area:**
Strong Heart Study
USPHS Indian Hospital
1515 NE Lawrie Tatum Rd.
Lawton, OK 73507
Office: (580) 353-0350 x350
FAX: (580) 531-9107

Kathe Samuelson, R.N.
Supervisor
E-mail: kathe.samuelson@ouhsc.edu

Verna Cable
Tribal Liaison
E-mail: verna-cable@ouhsc.edu

Tristian Ferguson
Medical Assistant
E-mail: tristian-ferguson@ouhsc.edu

Carmen Klinekole
Medical Assistant
E-mail: carmen-klinekole@ouhsc.edu

Shelby Santos
Office Clerk
E-mail: shelby-santos@ouhsc.edu

**ECG and Ultrasound Reading Center**
Weill Medical College of Cornell University
New York - Presbyterian Hospital
525 East 68th Street, Rm. K415 (Box 222)
New York, NY 10021
FAX: (212) 746-8561 or 746-8451

Virginia Burns, B.A.
Secretary
Office: (212) 746-4655 or 746-4650
E-mail: vmburns@med.cornell.edu

Dawn Fishman
Office: (212) 746-2198 or 746-4654
E-mail: dfishman@med.cornell.edu

Jennifer Liu, M.D.
Office: (212) 746-2437
E-mail: jeliu@med.cornell.edu

Mary Paranicas, B.S.
Office: (212) 746-4654 or 746-4655
E-mail: mfparan@med.cornell.edu

Lily S. Yee, B.S.
Office: (212) 746-4655 or 746-4650

**Family Study Center**
Department of Genetics
Southwest Foundation for Biomedical Research
P.O. Box 760549
San Antonio, TX 78245-0549
Express Service, change last two lines to:
7620 NW Loop 410
San Antonio, TX 78227-5301
SFBR Main Phone: (210) 258-9400

Teresa Cantu
Office: (210) 258-9651
FAX: (210) 258-9883
E-mail: tcantu@darwin.sfbr.org

Vanessa Fehlinger
Office: (210) 258-9562
FAX: (210) 258-9883
E-mail: vfehling@darwin.sfbr.org

**SHS Nosologist**
Karl E. Wise
Nosologist
4705 Glen Coe St., Royal Highlands
Leesburg, FL 34748-2327
Office: (352) 728-2172
Sandra Laston, R.N., Ph.D.
Office: (210) 258-9536
FAX: (210) 670-3317
E-mail: slaston@darwin.sfbr.org

Charlotte Wenger
Office: (210) 258-9691
FAX: (210) 670-3317
E-mail: cwenger@darwin.sfbr.org

Kari North, Ph.D.
Assistant Professor
Department of Epidemiology
University of North Carolina Chapel Hill
Bank of America Center
137 E. Franklin St., Suite 306
CB#8050
Chapel Hill, NC 27514-3628
Office: (919) 966-2148

FAX: (919) 966-9800
E-mail: kari_north@unc.edu

Dietary Study
Strong Heart Dietary Study
Aberdeen Area Indian Health Service
Federal Building, Rm. 309
115 4th Ave., S.E.
Aberdeen, SD 57401

Alice Seversen
Data Entry
Office: (605) 226-7491
FAX: (605) 226-7733
E-mail: alice.severson@ihs.gov
APPENDIX 6
THE STRONG HEART STUDY V
CARDIOVASCULAR DISEASE IN AMERICAN INDIANS

CONSULTANTS

Jan Beals, Ph.D.
Co-Director of Research
Division of American Indian & Alaska Native Programs
Box A011-13, UNP
4455 E. 12th Avenue
Denver, CO 80220
Office: (303) 315-9229
FAX: (303) 315-9579
E-mail: jan.beals@uchsc.edu

Bennett Dyke, Ph.D.
Southwest Iconics
98 Cliffside Dr.
San Antonio, TX 78231
Office: (210) 492-5947
FAX: (210) 492-1813
E-mail: bdyke@satx.rr.com

Jonathan Bella, M.D.
Bronx Lebanon Hospital
Division of Cardiology, 12th Floor
1650 Grand Concourse
Bronx, NY 10457
Office: (718) 518-5222
FAX: (718) 518-5585
E-mail: jonnabella@earthlink.net

Paul Enright, M.D.
Respiratory Sciences
University of Arizona
1501 N. Campbell Ave., Rm. 2342
Tucson, AZ 85724
Office: (520) 626-6114
Home: (520) 577-8254
E-mail: lungguy@aol.com

Peter Bennett, M.D.
NIH-NIDDK
1550 E. Indian School
Phoenix, AZ 85014
Office: (602) 200-5219
FAX: (602) 200-5225
E-mail: sbennett@email.arizona.edu

Stephen Epstein, M.D.
Executive Director
Cardiovascular Research Institute
110 Irving St., N.W. 4B-1
Washington, DC 20010
Office: (202) 877-5977
FAX: (202) 877-2715
E-mail: Stephen.Epstein@medstar.net

Mark Daniels, Ph.D.
Department of Psychology
University of South Dakota
414 East Clark St., SDU 107
Vermillion, SD 57069
Office: (605) 677-5353
FAX: (605) 677-6604
E-mail: mdaniels@usd.edu

James Everhart, M.D., M.P.H.
Chief, Epidemiology and Clinical Trials Branch
Division of Digestive Diseases and Nutrition
National Institute of Diabetes and Digestive and Kidney Diseases
Two Democracy Plaza, Room 673
6707 Democracy Boulevard, MSC 5450
Bethesda, MD 20892-5450
Office: (301) 594-8878
FAX: (301) 480-8300
E-mail: je17g@nih.gov

Michael Davidson, M.D.
4132A W. 88th St.
Anchorage, AK 99502
Phone: (907) 245-0508
E-mail: mdavidso@cesmail.net
James M. Galloway, M.D.
Director, Native American Cardiology Program
University Medical Center
1501 N. Campbell Ave., Rm. 4612
Tucson, AZ 85724
Office: (520) 694-7000
FAX: (928) 694-6712
Use following address when sending M&M charts:
1215 N. Beaver St., Ste. 201
Flagstaff, AZ 86001
Office: (928) 214-3920
FAX: (928) 214-3924
E-mail: jgalloway@umcaz.edu
Send cc to Diane if pressing:
dsteuart@umcaz.edu

Betty Jarvis, R.N.
6869 E. Kelton Lane
Scottsdale, AZ 85254
Phone: (480) 951-5079
Cell: (602) 448-1774
E-mail: bjarvis27@qwest.net

Dr. George King
Research Division
Joslin Diabetes Center
1 Joslin Place
Boston, MA 02215
Office: (617) 732-2622
FAX: (617) 732-2637

Jorge Kizer, M.D.
Division of Cardiology
Weill Medical College of Cornell University
New York – Presbyterian Hospital
525 East 68th Street, Rm. K415 (Box 222)
New York, NY 10021
Office: (212) 746-4655
FAX: (212) 746-8561 or 746-8451
E-mail: JOK2007@med.cornell.edu

William Knowler, M.D., Dr. P.H.
Chief, Diabetes & Arthritis Epidemiology Section
NIH-NIDDK
1550 E. Indian School
Phoenix, AZ 85014
Office: (602) 200-5206
FAX: (602) 200-5225
E-mail: wknowler@phx.niddk.nih.gov
Alternate E-mail: wk2c@nih.gov

Brett Koplin, M.D.
Department of Child and Adolescent Psychiatry
Mayo Clinic
200 First Street S.W.
Rochester, MN 55905
Office: (507) 284-1043
FAX: (507) 284-4158
E-mail: koplin.brett@mayo.edu

Andrea Kriska, Ph.D.
University of Pittsburgh
Graduate School of Public Health
Department of Epidemiology
130 DeSoto St., Room 505B
Pittsburgh, PA 15261
Office: (412) 624-3996
FAX: (412) 624-7397
E-mail: aky+@pitt.edu

Gary Leonardson, Ph.D.
55 Rodeo Trail
Dillon, MT 59725
Office: (406) 683-6424
FAX: (406) 683-4090
E-mail: gary@avicom.net

Andrew Narva, M.D.
Indian Health Service
801 Vassar Drive, N.E.
Albuquerque, NM 87106
Office: (505) 248-4018
FAX: (505) 248-7697
E-mail: anarva@albmail.albuquerque.ihs.gov
APPENDIX 7

THE STRONG HEART STUDY V

CARDIOVASCULAR DISEASE IN AMERICAN INDIANS

Confidentiality Pledge

I, ___________________________________________, understand that data obtained for subjects of research projects are confidential.

I will not reveal to unauthorized persons any patient’s name or any identifying information or any other information obtained from subjects of the project entitled, “Cardiovascular Disease in American Indians (The Strong Heart Study)”.

I will not allow any persons who are not authorized members of the Strong Heart Study staff to have access to any information collected from or about the subjects.

I will properly store the data forms, computer printouts and other documents in locked file cabinets or drawers to protect confidentiality.

I understand that breach of this confidentiality pledge is grounds for dismissal from employment on the Strong Heart Study.

I will return all data to the Principal Investigator when my employment terminates.

_________________________________________
Staff Member

_________________________________________
Principal Investigator

_________________________________________
Date
STRONG HEART STUDY
PAPER PROPOSAL

Title of Paper: (include the phrase “Strong Heart Study” whenever possible)

Name of Primary Author:

Author Affiliation:

Suggested Co-Authors:

Suggested Key words:

Outline of Paper:
  a) Introduction (Rationale)
  b) Methods
  c) General analysis plan

Analysis Responsibility: (authors or Coordinating Center)

Note: 1) If no SHS PI is a co-author and if the analyses are not performed by the CC, the authors must agree to submit the penultimate (next to final) draft to the Coordinating Center for statistical review.

2) Authors must comply and respond regularly to the status survey on their approved paper proposals conducted by the SHS P&P Committee twice a year.

3) Papers lacking a PI as a co-author. P&P will advise the primary author whether a near final draft will need to be sent to the P&P committee for review by at least two (2) reviewers (selected by the Chairperson). This review is the first step that must be completed prior to review of the penultimate draft by NHLBI/Tribes/IHS.

Submitted by: (Corresponding author, address, telephone, fax and e-mail for correspondence)

Date:
STRONG HEART STUDY

FAX TRANSMITTAL

TO: FAX NO.: 

FROM: P&P Committee FAX NO.: (405) 271-4390

DATE:

SUBJECT: Paper proposal entitled:

SHS P&P Committee decision:

_____ Approval with recommendations as listed below:

_____ Disapproval (Please see recommendations below)

_____ Deferred (Please see recommendations/comments below)

Recommendations:

Assigned paper no.:

Please fill out and return all the forms attached with this memo. Refer to the above number for all correspondence about this paper. If no SHS PI is a co-author and if the analyses are not performed by the CC, the penultimate (next to final) draft must be sent to the Coordinating Center for statistical review. Please inform us when this paper is approved by the NIH or accepted by a journal and if there is a change of the title. It is very important that you respond promptly during our ‘Paper Progress Survey’ done twice a year.

PLEASE NOTE: A Lay Summary is required when submitting the completed draft for NHLBI review and IHS IRB and tribal approvals. Also, the IHS IRBs require that all SHS manuscripts must contain the following disclaimer (verbatim): “The opinions expressed in this paper are those of the author(s) and do not necessarily reflect the views of the Indian Health Service.”

Center for American Indian Health Research, College of Public Health, University of Oklahoma Health Sciences Center, P.O. Box 26901, Oklahoma City, OK 73190
Phone: (405) 271-3090
STRONG HEART STUDY

THESIS/DISSERTATION PROPOSAL FORM

Title of Thesis/Dissertation:

Name of Degree Candidate:

Type of Degree:

Candidate Affiliation:

Primary Mentor:  (With e-mail, telephone and fax numbers, and address for correspondence)

Descriptions of Thesis/Dissertation Plan:

1. Prospectus - for Doctoral Thesis/Dissertation (if prospectus is not required by your degree program submit a detailed outline).

2. Outline - for Masters/Bachelor Thesis

Submitted by:  (Corresponding candidate, with telephone and fax numbers and address for correspondence)

Date:
FAX TRANSMITTAL

TO: 

FROM: Elisa T. Lee, PhD
SHS P&P Committee

DATE: 

SUBJECT: Thesis/Dissertation proposal entitled:

SHS P&P Committee decision:

_____ Approval with recommendations as listed below:

_____ Disapproval

Recommendation:

Assigned thesis/dissertation approval no.: T

Please fill out and return all forms attached with this memo to SHS P&P Committee. Please include the above thesis/dissertation approval number in all correspondence with us about this thesis/dissertation. Also, be advised that you need to write a paper for publication based on the SHS data used for this thesis/dissertation, and you must submit a paper proposal to the SHS P&P Committee prior to writing that paper.

NUMBER OF PAGES _____ (INCLUDING COVER SHEET)

College of Public Health, P.O. Box 26901, Oklahoma City, OK 73190, Phone: (405) 271-3090
Agreement for Data Distribution/Paper/Thesis/Dissertation Proposal*

To: Strong Heart Study Coordinating Center

From: __________________________________________ (Principal Investigator / First Author)

Institution/Address: __________________________________________

Name of the associated SHS PI / Mentor: _____________________________

Title of Study, Paper, Thesis or Dissertation: __________________________

Paper/Thesis/Dissertation Number (if known): ______________

I agree to read and follow the SHS protocol with regard to distribution and analysis of Strong Heart Study data that I request or that I generate in my research/paper/thesis/dissertation. I have attached a research protocol or a paper/thesis/dissertation proposal describing how I will use these data to better understand cardiovascular disease and its related diseases in American Indians.

I agree to protect the confidentiality and privacy of the SHS participants and the security of the data. I am not to transfer or disclose any individually identifiable information about the SHS participants at any time. Violation of the confidentiality agreement is considered a breach of confidentiality and may leave requesting investigator liable to legal action on the part of Study participants and their families. I also agree that the SHS data provided to me by the SHS Coordinating Center or SHS investigators are to be used only for the research protocol or the paper/thesis/dissertation approved by the SHS P&P Committee or the Steering Committee. I further agree not to distribute SHS data to anyone else.

For each paper I wish to write using any SHS data, I agree to comply with the SHS Publication Policy and to submit a paper proposal for review and approval of the SHS P&P Committee. Further approvals from the NHLBI, IHS, and the participating tribes will be needed prior to submission to any journal for publication. If approval from the SHS P&P Committee, the NHLBI, IHS, or the participating tribes is not granted, I agree not to publish these results.

I understand that the SHS P&P Committee or Steering Committee will assist me in revising my paper in such a way that will make it acceptable for publication. I agree to include at least one of the SHS investigators as a co-investigator and a co-author. I will send a reprint of my published article to the NHLBI Program office, and all other as detailed in the SHS P&P Publication Policy.

Signed: ___________________________ Date: ___________________________

* Each requesting investigator must complete this agreement separately.
STRONG HEART STUDY

REQUEST FOR DATA

(Please fill out and request the data needed for ONLY this approved paper/thesis. Data requested in excess will not be honored)

Title of paper or thesis:

Number assigned for this paper or thesis:

Primary author:

Purpose (Please check one):

___Paper
___Thesis
___Abstract for professional conference
___Invited talk
___Pilot data for grant or contract submission
___Quality control or local monitoring
___Other

Date Needed: mm / dd / yy (please allow 1-2 weeks from data request received)

Data for Study Period: (mark ONLY the phase you need the data from)

Phase-I Phase-II Phase-III Phase-IV

Center:

Arizona Oklahoma South/North Dakota All 3 centers

Variables Needed: (List ONLY the variables you need for this approved paper/thesis)

COORDINATING CENTER USE ONLY:

Date Received:

Date Data Delivered:
STRONG HEART STUDY

REQUEST FOR DATA ANALYSIS

Title of project:

Major hypotheses: 1)

2)

3)

4)

5)

Purpose: Paper
Abstract for professional conference
Invited talk
Pilot data for grant or contract submission
Quality control or local monitoring
Other

Investigator(s):

Expected date of completion: __________ / __________ / __________

mm dd yy

Variables to used: (List all the variables)
Statistical methods to be used (check all that apply):

- Summary statistics and frequencies
- Simple correlation and partial correlation
- Regression analyses
- t-test, ANOVA, and multiple comparison
- Logistic regression
- Other
  (Specify)

Comments:
STRONG HEART STUDY
DATA ANALYSIS MONITORING SYSTEM

When authors/researchers request Strong Heart Study (SHS) data for any purpose, the Strong Heart Study Coordinating Center (SHS-CC) would like to know how you manage and analyze the data. By answering the following questions, the SHS-CC is better able to track SHS data utilization patterns and to provide needed information for quality control. Thank you for your cooperation.

1. Do you use any of the following statistical package(s) for data analyses?
   Check all applicable.
   ___ a. SAS
   ___ b. SPSS
   ___ c. BMDP
   ___ d. S+
   ___ e. Statistic
   ___ f. StatXact
   ___ g. Other, specify: ________________________________

2. Other than the routine SHS derived variables, do you plan to derive any variables for your analysis purposes?
   ___ Yes.   ___ No.

3. If you plan to derive your own variable(s), will you consult with the SHS-CC?
   ___ Yes.   ___ No.

   If you derive certain variables for your analysis purpose, please attach a copy of the algorithm that you will use to define your variable(s) and the program to generate the variable(s)
4. Do you usually use any of the following procedures in your statistical analyses?
Check all applicable.

___ a. Multiple regression
___ b. Logistic regression
___ c. Time-related variables analysis
___ d. Modeling
___ e. Simulation
___ f. Other, specify: ____________________________

5. What training does your statistician(s) have?
Check all applicable.

___ b. Doctoral degree in other field but with quantitative training.
___ c. Master degree in statistics/biostatistics/math statistics.
___ d. Master degree in other field but with quantitative training.
___ e. Bachelor degree in statistics/biostatistics/math statistics.
___ f. Bachelor degree in other field but with quantitative training.
___ g. Other, specify: ____________________________

Feedback:

Please return to Strong Heart Study-Coordinating Center either by email or fax as soon as possible.
STRONG HEART STUDY

ANCILLARY STUDY PROPOSAL FORM

I. Basic Study Information and Projected Impact on SHS
1. Title of ancillary study:

2. Ancillary study PI(s) contact information (name, address, phone and fax numbers, e-mail address):

3. Proposed collaborators (must include at least one SHS investigator):

4. Summary of tasks involved for SHS Centers (NA=not applicable)

<table>
<thead>
<tr>
<th>Center</th>
<th>Enroll or examine participants (N)</th>
<th>Assay samples (N participants)</th>
<th>Provide samples (N participants)</th>
<th>Provide data (yes/no)</th>
<th>Analyze data (yes/no)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHS CC</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PML (Central Lab)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>SWF (DNA Lab)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>NA</td>
</tr>
<tr>
<td>RC (Ultrasound &amp; ECG Reading Center)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AZ Field Center (FC)</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>SD/ND FC</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>OK FC</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
</tr>
</tbody>
</table>
5. SHS participant and staff involvement:

A. Participants:
   Describe number of participants needed; special characteristics of participant group(s); age, and
gender distribution. Will participants be contacted, interviewed, or examined? If so, describe
participant involvement. Will biologic samples be collected from participants? If so, provide
details. Estimate time required of each participant.

B. Stored SHS specimens:
   Describe materials to be used (e.g., stored plasma, urine, DNA).
   i. SHS Phase(s) (e.g., I, II, etc) from which samples are to be obtained
   ii. Sample type (e.g., serum, EDTA, citrate, DNA)
   iii. Ability to use previously thawed samples
   iv. Sample volume (or weight for DNA) (conservation of samples is a critical factor in obtaining
       approval)
   v. Efforts to integrate sample needs with those of other studies to conserve sample and/or limit
      freeze-thaw cycles.

C. SHS Field Centers:
   Describe effort (and estimated time) required of SHS staff at each participating FC.

D. SHS Coordinating Center/US&EKG Reading Center/SWF Genetic Center/PML:
   Describe effort (and estimated time) required of SHS staff. Specifically:
   i. Will the Coordinating Center be involved in data collection, tracking, or preparation of
      forms or software? If so, provide details. If not, will these tasks be completed locally by
      the ancillary study, and a data file sent to the CC?
   ii. If the Reading Center, Genetic Center, or PML is involved, will data be sent directly
       from that entity to the CC for processing (dataset construction, data cleaning, data
       analyses, etc) or will processing be done by the ancillary study?
   iii. Will data analyses be done by the ancillary study or by the CC? If analyses will be done
       locally, the CC must verify analyses prior to publication of resulting paper(s).
   iv. How many ancillary study papers do you estimate will be written?

6. Variables/measurements from the SHS database to be analyzed:

7. Genomic information (defined as any data from a participant’s DNA):

   A. Does your proposal include any genomic materials? (please check one)
      __ No (go to question 8)  __ Yes (see question 7B)

   B. Name the gene(s), chromosomal regions, genotypes, SNPs to be investigated.

   C. Is genetic information used to address a primary aim or secondary aim of SHS?
      (please check one or both)
      __ Primary aim (heart/vascular/pulmonary disease)
Secondary aim (other health conditions)
List the conditions addressed: ______________________________________

8. Proposed starting and ending dates:

9. Estimated cost by year; number of years:

10. Source of funding; projected date of submission:

11. Please indicate whether this ancillary study involves the support or collaboration of a for-profit corporation; if it does, please affirm that the data will not be use to patent or profit from any process, aspect or outcome of the analysis (as stipulated in the SHS consent forms).

12. What is/are the advantage(s) to American Indian communities, to the participants, to SHS, and to yourself of conducting the study within the SHS cohort versus another population?

13. Possible/probable impact on ongoing SHS studies (SHS or other ongoing SHS ancillary studies):

14. Provide the following assurances (answer each):

   (1) Who (name, position (and contact information, if different from above)) will provide the annual progress reports? (ancillary study PI or designate)

   (2) How will confidentiality of SHS participants be maintained?

   (3) The ancillary study PI will be given the first and exclusive opportunity to analyze, present and publish data collected by the ancillary study. Recipient agrees that an archival copy of the newly collected data and/or laboratory results, with documentation, will be sent in a secure manner to the SHS CC one year after the conclusion of the data cleaning and closure or one year after acceptance of the primary publication, whichever comes first. This should allow sufficient time for publication of the main (ancillary) study hypothesis. This transfer is the responsibility of the SHS representative(s) collaborating with the ancillary study. The data from the ancillary study will be included in the SHS dataset for distribution by the SHS CC and/or the NHLBI according to procedures currently under negotiation with the involved tribes, the IHS, and the NHLBI. Recipient agrees that it is the responsibility of the ancillary study PI to state in writing to the SHS SC any special circumstances that might warrant an exception to these guidelines for data sharing. In the spirit of encouraging collaboration, reasonable and justified requests for limiting such routine SHS access to the data will be honored, or some compromise will be worked out.

15. Ancillary Study PI signature:
I agree to comply with the SHS Ancillary Studies Policy (for current version, see the SHS website at: http://strongheart.ouhsc.edu/T)

Signed: ___________________________ Date: ____________________

(Ancillary Study PI)
II. Abbreviated Ancillary Study Proposal

Please provide a brief (2 to 4 page) description of the proposed study. Include the following:

Purpose:

Background:

Hypothesis(es):

Experimental Design (include sample size justification):

Methods, including:
- Participant involvement (if any)
- Data and/or samples to be collected by the ancillary study (attach questionnaires and forms)
- Statistical and Laboratory Analysis methods

Literature cited:

Please send (electronically or by surface mail) the completed proposal to:

Barbara V. Howard, PhD
Chair, SHS Steering Committee
President, MedStar Research Institute
6495 New Hampshire Ave Suite 201
Hyattsville, MD 20783
Phone: 301-560-7302
Fax: 301-560-7309
PAGER AND CELL PHONE: 301-602-0125

For Coordinating Center Use Only

Approved? _____ Date ______________ If approved, ancillary study # ________________
The undersigned parties hereby enter into this Distribution Agreement as of the date specified in the signature section at the end of this agreement.

PRELIMINARY STATEMENT
The National Heart, Lung, and Blood Institute (NHLBI), in collaboration with the SHS Investigators, the participating American Indian tribes, and the IHS, has supported collection of biological samples and clinical and other data from participants in the SHS. This clinically and genetically well-characterized population is a valuable scientific resource that is maintained under the joint stewardship of the SHS Investigators, the NHLBI, the participating tribes, and the IHS. Promoting use on a national scale of such a resource will require a large and concerted effort, which may involve investigators not currently part of the SHS. The NHLBI and the researchers it supports have a responsibility to the American Indian communities, the general public, and the scientific community to encourage as rapid scientific progress as possible using these resources, subject to appropriate terms and conditions. In order to take full advantage of such resources and maximize their research value, it is important that samples and data collected with public funds be made available, under appropriate terms and conditions, to the largest possible number of qualified investigators in a timely manner.

Biological samples and clinical and other data collected by the SHS have been stripped of all personal identifiers. However, identification of individual participants is made less difficult due to 1) public awareness of the Indian communities who participate in SHS, 2) public knowledge of the geographic areas from which the SHS participants were drawn, and 3) the wealth of data available in the SHS datasets. To protect the confidentiality and privacy of these participants and their families, investigators granted access to these data and materials must adhere to the requirements of this Distribution Agreement. Failure to comply with this Agreement will result in denial of further access to SHS samples and data. Violation of the confidentiality requirements of this agreement may leave requesting investigators liable to legal action on the part of the SHS participants and/or the US Government (NHLBI and/or IHS).

The SHS investigators have made a substantial long-term contribution in establishing, maintaining, and expanding the database and samples. The NHLBI, the tribes, the IHS, and the SHS investigators seek to encourage appropriate collaborative relationships by outside investigators with the SHS investigators, which will increase the value of this research in improving the health of American Indians and people in general. In all such collaborations, the contributions of the tribes and of the SHS investigators must be appropriately acknowledged.

The NHLBI, the tribes, the IHS, and the SHS Investigators further seek to promote the development of valuable discoveries and inventions beneficial to the public health based upon use of the SHS data and samples repositories. Sample materials from the SHS field centers are
DEFINITIONS

For purposes of this agreement,

"Data" refers to data and associated records collected and recorded from SHS participants through periodic examinations conducted pursuant to the SHS Investigators' contract with the NHLBI,

“Materials” refers to “biological materials” of participants (urine, blood samples, and products thereof, including extracted DNA) collected and prepared in SHS pursuant to the SHS Investigators' contract with the NHLBI;

"Genetic Analysis Data" refers to "molecular genetic data", which consists of data derived from analyses of DNA samples contained in biological materials including but not limited to genotyping analysis, anonymous marker polymorphisms, DNA sequence information, mutation analysis, and other genetic analyses.

RECIPIENT

_____________________________________, a non-profit organization governed by the laws of the State of _______________________________

with a principal address at ________________________________________________________________

("Recipient") requests access to SHS data and materials at its sole risk and at no expense to the SHS or NHLBI.

AGREED TERMS AND CONDITIONS

It is mutually agreed as follows:

1. Data. The SHS, the NHLBI, the participating tribes, and the IHS agree to provide Recipient with SHS data described as follows for use by the Recipient's principal investigator named below ("Principal Investigator"): ____________________________________________________________________________________________

2. Materials. The SHS, the NHLBI, the participating tribes, and the IHS agree to transfer to Recipient SHS materials described below to conduct the research described in paragraph 3 below. These materials (including numbers of samples) are described as follows: ____________________________________________________________________________________________

Strong Heart Study V  70/01/2006; rev:  12/19/2006   I- Appendix 8-17  P&P Forms
3. Research Project.
3.1 The SHS materials and/or data will be used by Recipient's Principal Investigator solely in connection with the following research project ("Research Project"), specifically described below or in an attached Exhibit A:

3.2 The Research Project involves the following SHS investigator(s) as co-investigator(s). His/her/their name(s) and the work he/she/they will perform are described below or in an attached Exhibit B:

3.3 This Distribution Agreement covers only the above-described Research Project. Recipient must complete and submit a separate Data and/or Materials Distribution Agreement for each research project for which SHS data and/or materials are requested.

3.4 Recipient agrees to follow the current SHS Ancillary Studies Policy in conducting this Research Project. As the ancillary study progresses, the Principal Investigator will find the most up-to-date SHS policy by visiting the SHS website at: http://strongheart.ouhsc.edu.

4. Non-transferability. This Agreement is not transferable. Recipient agrees that substantive changes made to the Research Project described above, and/or appointment by Recipient of another Principal Investigator to complete the Research Project, require execution of a new Agreement in which the new Principal Investigator and/or new Research Project are designated.

5. Publication. Prompt publication of the results of the Research Project is encouraged. The PI or other representative of the ancillary study, and if necessary the SHS SC, will consult with the CC during data analysis to ensure that all study data used in analysis of ancillary results are consistent with data in the main SHS database. Manuscript proposals must be approved in advance by the SHS Publications and Presentations Committee (P&P). This procedure is necessary to establish authorship and prevent overlap in the publication effort. Approval of manuscript proposals is sought by submitting the proposal using standard SHS format (see SHS Paper Proposal Form in Appendix 8 of SHS V Manual of Operations, Volume 1) to the SHS P&P. The SHS Ancillary Study Proposal Form (form can be found in Appendix 8), as signed by the ancillary study PI, stipulates that the ancillary study investigators agree to submit paper proposals for approval by the SHS P&P and to submit draft manuscripts for approval by the NHLBI, the IHS IRBs, and the tribes (for full details see section 1.7 of SHS V Manual of Operations, Volume 1). Additionally, abstracts for presentations at meetings require approval by the NHLBI and the Dakota Center IHS IRB (see section 1.7). The investigator who assumes lead responsibility for the ancillary study shall generally be listed as an author. Whenever possible, the phrase "Strong Heart Study" should be included in the manuscript title and listed as a key word. Upon publication, reprints must be distributed as specified in section 1.7.
6. Acknowledgments. Recipient agrees to acknowledge, as deemed appropriate, the contributions of the participating tribes, the IHS, the SHS investigators and staff, and the NHLBI in any and all oral and written presentations, disclosures, and publications resulting from any and all analyses of the SHS data and/or materials.

6.1 Collaborations/Acknowledgments. Recipient will acknowledge the SHS co-investigators as co-authors, as appropriate, on any publication and will use the acknowledgment printed below.

"The Strong Heart Study (SHS) is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with the SHS Investigators, the participating tribes, and the Indian Health Service (IHS). This manuscript has been reviewed by the SHS and NHLBI for scientific content and consistency of data interpretation with previous SHS publications and significant comments have been incorporated prior to submission for publication. Additionally, this manuscript has been reviewed by the participating tribes and the IHS IRBs involved. The opinions expressed in this paper are those of the author(s) and do not necessarily reflect the views of the Indian Health Service."

7. Non-Identification. Recipient agrees that SHS materials and/or data will not be used, either alone or in conjunction with any other information, in any effort to determine the individual identities of any of the participants from whom data and/or materials were obtained.

8. Use Limited to Research Project. Recipient agrees that SHS data and/or materials, their progeny, and unmodified or modified derivatives thereof will not be used in any experiments or procedures that are not disclosed and approved as part of the Research Project.

9. Compliance with Participant’s Informed Consent and HIPAA form. Recipient agrees that SHS data and/or materials, their progeny, and unmodified or modified derivatives thereof will not be used for any purpose contrary to a participant’s applicable signed informed consent document(s). It is the responsibility of the Recipient’s Principal Investigator to consult with the SHS investigators and ascertain, specifically and in detail, the terms and conditions of applicable SHS informed consent documents. In keeping with the Health Insurance Portability and Accountability Act (HIPAA) guidelines, participants have consented to having their data used by investigators outside of SHS to study the causes of cardiovascular disease, lung disease, and their risk factors and to conduct genetic research regarding cardiovascular disease, lung disease, and their risk factors. Recipient will make no attempt to access Protected Health Information including, but not limited to, their identities, their medical records, family information, employer or insurance information, or previous medical or genetic conditions not included as part of the limited data set or materials provided to the Recipient by the SHS or obtained by the Recipient in conducting IRB-approved procedures of the ancillary study.

10. No Distribution, Avoidance of Waste, Return of Materials. Recipient agrees to retain control over SHS data and materials, their progeny, and unmodified or modified derivatives thereof, and further agrees not to transfer data and/or materials, their progeny, or unmodified or modified
derivatives thereof, with or without charge, to any other entity or any individual. Recipient agrees, in handling the SHS biological materials, to make reasonable efforts to avoid contamination or waste of the samples. When the Research Project is completed, or three (3) years have elapsed from the effective date of this Distribution Agreement, whichever occurs first, the SHS materials will be either returned to the SHS lab (PML or SWF, as appropriate) or disposed of as mutually agreed upon by the SHS Investigators, NHLBI, IHS, participating tribes, and Recipient, unless an extension of this Agreement is obtained. A record of how SHS biological materials have been handled and stored during that time must be submitted to the SHS lab (PML or SWF, as appropriate) prior to the end of the Recipient’s project to facilitate decisions as to appropriate handling of the remaining SHS materials.

11. Ancillary Study Annual Reports. Recipient agrees to provide the SHS Coordinating Center (and thereby the NHLBI) with a report every twelve (12) months during the term of this Agreement containing a summary of findings derived by Recipient in the performance of the Research Project. Such report will summarize all data derived by Recipient up to six (6) months before the reporting date.

12. Costs/No Warranties. Costs for DNA or other material distribution (retrieval, processing, and shipping) will be borne by the Recipient at no cost to SHS. NO WARRANTIES, EXPRESS OR IMPLIED, ARE OFFERED AS TO THE MERCHANTABILITY OR FITNESS FOR ANY PURPOSE OF THE DATA AND/OR MATERIALS PROVIDED TO RECIPIENT UNDER THIS AGREEMENT.

13. Recipient's Responsibility for Handling Biological Materials. Recipient acknowledges that biological materials may carry viruses, latent viral genomes, and other infectious agents. The Recipient agrees to treat Biological Materials as if they are not free of contamination, and that SHS biological materials will be handled only by trained persons under laboratory conditions that afford adequate biohazard containment. By accepting SHS biological materials, Recipient assumes full responsibility for their safe and appropriate handling.

14. Non-Endorsement, Indemnification. Recipient agrees not to claim, infer, or imply Governmental endorsement of the Research Project, the entity, and personnel conducting the Research Project except as described in paragraph 6 above. To the extent permitted by law, Recipient agrees to hold the United States Government, the SHS investigators, and all other investigator(s) who generated SHS data and materials, and the agents and employees of each of them, harmless and to defend and indemnify all such parties for all liabilities, demands, damages, expenses, and losses arising out of Recipient's use for any purpose of SHS data and materials, their byproducts, or modified or unmodified derivatives.

15. Accuracy of Data. The United States Government and the SHS investigators are not responsible for the accuracy of SHS data or materials provided.

16. Recipient's Compliance with IRB Requirements. Recipient acknowledges that the conditions for use of the SHS data and/or materials have been approved by the Recipient's Institutional
Review Board (IRB) in accordance with Department of Health and Human Services regulations (45 CFR Part 46). Recipient agrees to comply fully with all such conditions and with the participants' informed consent and HIPAA documents, and any additional conditions that may be imposed by the SHS Centers’ IRBs (IRBs of all grantee institutions and the IHS IRBs for each of the 3 field centers) and relevant HIPAA approval bodies. Recipient agrees that any proposed change in the approved ancillary study protocol must be submitted to SHS for approval, and that such change may require review and approval of all of the aforementioned SHS IRBs. Furthermore, any unanticipated problems involving risks to subjects or others must be promptly reported to SHS and the SHS IRBs. Recipient remains subject to all applicable State and local laws and regulations and institutional policies that provide additional protections for human subjects.

17. Amendments. Amendments to this Distribution Agreement must be made in writing and signed by authorized representatives of both parties.

18. Conflict of Interest. The Recipient agrees to disclose promptly any direct or indirect conflicts of interest, such as affiliation(s) with any organization with any financial interest in the subject matter of the proposed research employing SHS data and/or materials. Examples of such affiliations are employment, consultancies, expert testimony, honoraria, stock, or retainers that may affect the work being considered.

19. Termination. The SHS Steering Committee, in consultation with the NHLBI, may terminate this Distribution Agreement if Recipient is in default of any of its conditions and such default has not been remedied within 30 days after the date of written notice to the Recipient by the SHS of such default. Upon termination of this Distribution Agreement, Recipient agrees to return all SHS data and materials to SHS.

20. Disqualification, Enforcement. Failure to comply with any of the terms specified herein may result in disqualification of Recipient from receiving additional SHS data and/or materials. The United States Government (on behalf of NHLBI and/or IHS), the SHS investigators, and the participating tribes shall have the right to institute and prosecute appropriate proceedings at law or in equity against the Recipient for violating or threatening to violate the confidentiality requirements of this agreement, the limitations on the use of SHS data and/or materials provided, or both. Proceedings may be initiated against the violating party, or legal representatives, and assigns, for a restraining injunction, compensatory and punitive damages, mandamus, and/or any other appropriate proceeding in law or equity, including obtaining the proceeds from any intellectual property or other rights that are derived in whole or in part from the breach of the confidentiality requirements or use limitations of this agreement. In addition, Recipient acknowledges that a breach or threatened breach of the confidentiality requirements or use limitations of this agreement may subject Recipient to legal action on the part of SHS participants, their families, or both.

21. Accurate Representations. Recipient certifies that to the best of his/her knowledge and belief the contents of any statements made or reflected in this document are truthful and accurate.
22. Prior Distribution Agreements. The following stipulation applies only to Recipients who have entered into a previous Distribution Agreement with SHS or NHLBI: Execution of this Distribution Agreement is contingent upon Recipient's compliance with all terms and conditions of existing Distribution Agreements with SHS or NHLBI.

23. Recipient's Ancillary Study Results to be Provided to SHS. The ancillary study PI will be given the first and exclusive opportunity to analyze, present and publish data collected by the ancillary study. Recipient agrees that an archival copy of the newly collected data and/or laboratory results, with documentation, will be sent in a secure manner to the SHS CC one year after the conclusion of the data cleaning and closure or one year after acceptance of the primary publication, whichever comes first. This should allow sufficient time for publication of the main (ancillary) study hypothesis. This transfer is the responsibility of the SHS representative(s) collaborating with the ancillary study. The data from the ancillary study will be included in the SHS dataset for distribution by the SHS CC and/or the NHLBI according to procedures currently under negotiation with the involved tribes, the IHS, and the NHLBI. Recipient agrees that it is the responsibility of the ancillary study PI to state in writing to the SHS Steering Committee any special circumstances that might warrant an exception to these guidelines for data sharing. In the spirit of encouraging collaboration, reasonable and justified requests for limiting such routine SHS access to the data will be honored, or some compromise will be worked out.

In any instance when genotyping is performed in the collaborating lab (as opposed to SWF), as soon as the variant information is known (before genotyping and certainly before data analysis and publication), the SNP information must be conveyed to the SHS Steering Committee. Recipient will provide Genetic Analysis Data, indexed by genotyping ID number in the precise electronic format specified by the SHS CC and/or SWF and/or NHLBI. When genotyping has been conducted, DNA marker names and allele names will be provided for each individual subject as indexed by the SHS participant ID number. Descriptive information about each typed marker, preferably obtained from public databases, when applicable, must be provided, including any Human Genome SNP database information (nlm/ncbi) and the SNP database name, the chromosomal physical map location and source of map location, gene name (if relevant) and location in gene relative to transcriptional start site, and surrounding DNA sequence or PCR primers used.

Recipient agrees that the safety and confidentiality of the SHS data and/or materials at the collaborating institution are the responsibility of the ancillary study PI, as is the appropriate disposition of data and remainders of SHS samples after the ancillary study has been completed. Leftover DNA and any other types of laboratory specimens must be returned to PML or SWF. Files of SHS data must be returned to the SHS CC or deleted, as established and agreed at the outset of the collaboration.
This Distribution Agreement is entered into as of ____________ (effective date)

RECIPIENT:
Name of Recipient Entity:

Name and Title of Recipient's Authorized Representative:

_________________________ Date: ____________

PRINCIPAL INVESTIGATOR:
Principal Investigator's Name and Title:

Principal Investigator's Surface Mail Address:

Principal Investigator's Email Address:

Principal Investigator's Telephone Number:

Principal Investigator's Fax Number:

_________________________ Date: ____________

STRONG HEART STUDY:
Name and Title of SHS Authorized Representative:

_________________________ Date: ____________

NHLBI:
Name and Title of NHLBI's Authorized Representative:

_________________________ Date: ____________
STRONG HEART STUDY

ANNUAL UPDATE ON ANCILLARY STUDY

Please provide an update on the SHS Ancillary Study under your direction by answering the questions listed below. The information will be provided to the SHS Steering Committee and will also be included in the annual report to the SHS Observational Study Monitoring Board.

Send the completed report to:

Elisa T. Lee, Ph.D.
SHS PI and Director of the SHS Coordinating Center
Center for American Indian Health Research
College of Public Health
University of Oklahoma Health Sciences Center
PO Box 26901 - Room CHB-112
Oklahoma City, OK 73190

Express Svc:
Center for American Indian Health Research
College of Public Health
801 NE 13th St, Room CHB-100
Oklahoma City, OK 73104

Phone: 405-271-3090
Fax: 405-271-4390
Email: Elisa-Lee@ouhsc.edu

Ancillary Study Title: ___________________________________________ (Please use the title as supplied in your study proposal, which was approved by SHS.)

Ancillary Study #: ___________________________________________ (This is the study number which was listed on your SHS agreement form and assigned to your ancillary study by SHS when approved)

Date (day/month/year) ________________________________ (This report should be submitted to Dr. Lee on or about the anniversary date of the SHS approval of your ancillary study)

Current Status:
– Not yet started.
– In Progress.
– Data collection completed.
– Publication Stage.

Please give comments to explain any conditions that are noteworthy (such as funding not yet received, but expected in next 6 months, funding received and hiring in progress, problems encountered, etc)

Please provide a one-paragraph summary of your progress and findings, which can be reported to the SHS SC and OSMB.
Does this ancillary study or its parent study have a website? □ Yes  □ No  If yes, the URL is:

Please answer the following questions:

1. Any change in study status since your last update? (e.g., in data collection, data analysis, manuscript preparation, etc - or are all activities now completed?)

2. Was new funding received since your last update? What is the new funding source?

3. What is the status of each publication and presentation derived from the study? Provide full citations for published papers.

4a. Did you complete data collection and cleaning more than one year ago or was the primary manuscript accepted for publication more than one year ago?

4b. If yes to Q4a, have you sent the dataset to the SHS Coordinating Center? _____  If so, when? _____________
   If not, when do you plan to do so? ____________________________

   A dataset containing the important analytic variables should be sent to the SHS Coordinating Center (see the reminder below). Please send the data & documentation separately from this progress report. Please contact the CC regarding appropriate format, manner of transmission of data, data security, etc prior to sending your data.

5. Have there been any changes in your mailing address, phone and fax numbers, and e-mail address since you submitted your last annual report? _________  If so, please provide your current contact information:

A reminder regarding Publications and Presentations:
You should be aware that publication of SHS ancillary study data requires review by the SHS Publications and Presentations Committee, the NHLBI, the tribes, and the IHS IRBs. All manuscripts must be preceded by an approved manuscript proposal. Abstracts and presentations must be based on an approved manuscript proposal and must be approved by the NHLBI and the Aberdeen Area IHS IRB (please see section 1.7 of the SHS V Manual of Operations, Volume 1 for details).

A reminder regarding Ancillary Study data:
SHS policy requires that data collected by the ancillary study must be provided, with documentation, to the SHS Coordinating Center for integration into the main database. The ancillary study PI will be given the first and exclusive opportunity to analyze, present and publish data collected by the ancillary study. One year after data cleaning is complete or one year after the primary manuscript has been accepted for publication, whichever comes first, ancillary study data will be made available for additional uses by other SHS investigators. It is the responsibility of the ancillary study PI to state in writing to the SHS Steering Committee any special circumstances that would warrant an exception to these guidelines for data sharing. In the spirit of encouraging collaboration, reasonable and justified requests for limiting Steering Committee access to the data will be honored or some compromise will be worked out.