IS SICKLE CELL TRAIT A RISK FACTOR FOR STROKE AND CEREBRAL SMALL VESSEL DISEASE?

Melissa Champlin Caughey

A dissertation submitted to the faculty at the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Epidemiology in the Gillings School of Global Public Health.

Chapel Hill
2014

Approved by:

Laura Loehr

Kari North

Rebecca Gottesman

Nigel Key

Gerardo Heiss

ABSTRACT

Melissa Champlin Caughey: Is sickle cell trait a risk factor for stroke and cerebral small vessel disease?
(Under the direction of Laura Loehr and Kari North)

We have recently shown an association between sickle cell trait (SCT) and ischemic stroke in the Atherosclerosis Risk in Communities (ARIC) study. The etiology of stroke in this population is unclear, however. Though not considered a hematological disorder, the SCT phenotype is nonetheless associated with hypercoagulability, vasculopathy, and possibly hypoperfusion. To further understand the cerebrovascular pathophysiology, we examined a subset of African Americans in the ARIC study (N=844, mean age=62, female=64%) who were prospectively imaged by cerebral MRI in 1993-1995, and 470 (56%) who returned for a follow up MRI in 2004-2006. White matter lesions (WML) and subclinical brain infarctions (SBI) in participants with no prior history of stroke were detected by cerebral MRI. Associations between SCT and WML prevalence and severity were analyzed using ordinal logistic and linear regression. Similarly, associations between SCT and the prevalence and progression of SBI were analyzed using logistic regression. Models were adjusted for age, sex, cerebrovascular risk factors, and 10 principal components of ancestry. SCT was identified in 56 (6.6%) participants at the first MRI. Individuals with SCT had more prevalent (86% vs. 79%), and more severe (mean score 1.5 vs. 1.3) WML than individuals without SCT. SCT was also associated with a 20% increased odds of WML prevalence (POR 1.2, 95% CI: 0.7 – 2.0), and an adjusted mean severity score that was 0.2 (-0.1 – 0.5) points higher; however neither of these estimates was statistically significant. Likewise, SCT was not associated with prevalent SBI (POR = 0.7; 95% CI: 0.3 – 1.8), or

incidence of new infarctions by the follow up exam (OR = 1.4; 95% CI: 0.6 – 3.1). In conclusion, we observed no statistically significant associations between SCT and cerebral small vessel disease. There was a trend for greater WML prevalence and severity among those with SCT, as well as a higher 11-year incidence of SBI; however, the estimates were imprecise and inconclusive.

ACKNOWLEDGEMENTS

I am very grateful to the participants and staff of the Atherosclerosis Risk in Communities Study, without whom none of this would be possible. I am also thankful to Vimal Derebail and Abhi Kshirsagar, for permitting me the use of genotype data from their ancillary study. I am indebted to Misa Graff for guiding me through the GWAS and Exome chip data from the ARIC study. Last but not least, I would like to acknowledge my dissertation committee: Laura Loehr, Kari North, Nigel Key, Rebecca Gottesman, and Gerardo Heiss, for their valuable insight and instruction.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

ACAS:	Asymptomatic Carotid Atherosclerosis Study
ACPC:	Anterior commissure / posterior commisure
ARIC:	Atherosclerosis Risk in Communities Study
CSSCD:	Cooperative Study of Sickle Cell Disease
CHS:	Cardiovascular Health Study
CT:	Computed tomography
DNA:	Deoxyribonucleic acid
ECG:	Electrocardiogram
FLAIR:	Fluid attenuated inversion recovery
FMD:	Flow mediated dilation
GWAS:	Genome-wide association study
HWE:	Hardy-Weinberg equilibrium
ICD-9:	International classification of disease 9
LADIS:	Leukoaraiosis and Disability Study
MAF:	Minor allele frequency
MRI:	Magnetic resonance imaging
RFLP:	Restriction fragment length polymorphism
SBI:	Subclinical brain infarction
SCA:	Sickle cell anemia
SCT:	Sickle cell trait
SNP:	Single nucleotide polymorphism
sVCAM:	Soluble vascular cell adhesion molecule
TIA:	Transient ischemic attack
VADAS:	Vascular dementia assessment scale

WML: White matter lesions

CHAPTER 1: LITERATURE REVIEW

1. INTRODUCTION

African Americans are disproportionately burdened by stroke, a leading cause of disability and death in the United States. We hypothesize that sickle cell trait is a genetic risk factor cerebrovascular disease in African Americans. The following literature review will describe the genetics and pathophysiology of sickle cell trait, and its associations with hypercoagulability and vasculopathy. Following this, the pathophysiology and epidemiology of stroke, subclinical brain infarctions, and white matter lesions will be described, highlighting the genetic risk factors and etiological associations with thrombosis and vasculopathy. We will conclude by summarizing observational studies of sickle cell trait and stroke, and the public health implications of the hypothesized association.

2. SICKLE CELL TRAIT

2.1 Overview

Sickle cell trait (SCT) is the heterozygous form of sickle cell anemia (SCA), a Mendelian hemoglobinopathy characterized by misshapen red blood cells, acute chest syndrome, and vaso-occlusive crises[1]. Unlike sickle cell anemia, SCT is generally considered a benign carrier state, rather than a disease[1]. However, under certain conditions, the erythrocytes of SCT carriers are known to sickle, impeding blood flow through the

microcirculation and increasing coagulation. The expression of SCT is influenced by several genetic factors and the coinheritance of other hemoglobinopathies, as described below.

2.2 The Sickle Mutation

Adult hemoglobin consists of 2 α-globin chains, 2 β-globin chains, and 4 heme groups, which bind to and transport oxygen throughout the circulatory system[2]. Hemoglobinopathies, genetic mutations affecting the hemoglobin, either decrease α or β globin production, or generate mutated, functional variants of the β-globin chains[1]. Alpha globin is encoded by 2 genes located on the 16th chromosome: hemoglobin, alpha 1 (*HBA1*), and hemoglobin, alpha 2 (*HBA2)*; however *HBA2* produces 2-3 fold more protein than *HBA1*[3]. Beta globin is encoded by a single gene (hemoglobin, beta) located on the 11th chromosome. The majority of hemoglobinopathies affecting African Americans arise from mutations to the hemoglobin, beta (*HBB*) gene[1].

Figure 1: The human beta globin gene like cluster.
LCR = Locus Control Region; HBE1 = Hemoglobin Epsilon;
HBG2 = Hemoglobin Gamma G; HBG1 = Hemoglobin Gamma A;
HBBP1 = Hemoglobin Beta Pseudogene 1;
HBD = Hemoglobin Delta; HBB = Hemoglobin Beta

The *HBB* gene-like cluster (Figure 1) extends over 80,000 bases, and includes 6 genes: hemoglobin, beta (*HBB*); hemoglobin, delta (*HBD*); hemoglobin, beta pseudogene 1 (*HBBP1*); hemoglobin, gamma A (*HBG1*); hemoglobin, gamma G (*HBG2*); and hemoglobin, epsilon 1 (*HBE1*), which are regulated by a locus control region (LCR)[4]. Each gene is structurally similar, arising from gene duplications during the early evolution of

invertebrates[5]. With the exception of the pseudogene *HBBP1*, each gene produces functional globin proteins[4]. As the names imply, *HBB*, *HBD*, *HBG1*, *HBG2*, and *HBE1*, produce beta globin, delta globin, gamma globin, gamma globin, and epsilon globin; respectively[6]. The sickle mutation is an adenine to thiamine substitution occurring in the 6th codon of the human beta globin gene (*HBB*)[7]. Transcription of the mutated DNA produces a functional variant of the β-globin protein, constructed with valine in the place of glutamic acid[7]. Sickle β-globin combines with α-globin to form a protein product commonly referred to as hemoglobin S. While the hemoglobin of individuals with sickle cell anemia is almost entirely hemoglobin S, heterozygotes produce a combination of normal and mutated hemoglobin. The concentration of hemoglobin S in SCT carriers varies considerably, ranging from 25-45%[8]. In both SCT carriers and individuals with SCA, hemoglobin S production is known to be influenced by genetic modifiers, described subsequently in sections 2.5 and 2.8.

2.3 Erythrocyte Sickling

With deoxygenation, the erythrocyes of individuals with SCA become sickled and deformed[1]. To a lesser extent, sickling is also observed in SCT carriers[1]. Hemoglobin transports oxygen throughout the systemic arterial circulation[2]. The oxygen is released in the microcirculation, and supplied to the tissues by diffusion through the capillary walls[2]. When hemoglobin S releases oxygen, the aberrant valine within the mutated beta globin chains becomes exposed[9]. Unlike glutamic acid, found in normal beta globin, valine is a nonpolar, hydrophobic amino acid[9]. Valine fits within a hydrophobic pocket formed by glutamic acid and phenylalanine at the 88th and 85th residues, allowing it to bind to neighboring hemoglobin molecules[9]. This activates a polymerization cascade, forming long chains of hemoglobin molecules that eventually stretch and distort the erythrocytes into a characteristic

sickle cell shape[9]. At rest, approximately 30-60% of the erythrocytes in the deoxygenated venous circulation are sickled in individuals with sickle cell disease, compared to 1% in SCT carriers[10]. However, under experimental conditions, the percentage of sickled red blood cells in SCT carriers has been shown to increase from 1.0 ± 1.0% at rest, to 8.5 ± 7.1%, with exercise at a simulated altitude of 4000 meters[11]. Even under controlled conditions, a wide range of sickling (1% – 25%) among SCT carriers is observed[11]. In addition to oxygen saturation, *in vitro* sickling is also influenced by temperature, pH, and the intracellular concentration of hemoglobin S[12-14]. After several deoxygenation cycles, erythrocytes of individuals with SCA eventually become permanently sickled[1]. However, sickling in SCT carriers is generally reversible, as red blood cells regain normal morphololgy with oxygenation in the arterial circulation[1]. Irreversible sickling, evidenced by sickled red blood cells in the arterial circulation, is observed under experimental conditions in less than 1% of red blood cells in SCT carriers subjected to exercise testing[11].

2.4 Prevalence of Sickle Cell Trait

The sickled deformation of red blood cells is known to deter cytoadherance of *Plasmodium falciparum*, the parasite responsible for malaria[15]. As a result, individuals producing hemoglobin S have fewer episodes of malaria or high-density parasitemia, defined as an excess of 10,000 parasites per μL of blood[16]. While sickle cell anemia is a debilitating disease with an average life span of 44 years in the United States[17], the sickle mutation confers a selective survival benefit to heterozygotes in areas with endemic malaria, even with adequate access to antimalarial treatments. This is evidenced by a longitudinal study following 1022 children from a Kenyan birth cohort, which reported a decreased risk of all-cause mortality in children with SCT, during the critical period between 2 and 16 months,

when clinical immunity to malaria has not yet developed. Of note, all study participants with a detected fever and positive blood film were treated with sulfadoxine-pyrimethamine, an antimalarial drug[16]. Due to selective pressure, the sickle mutation, though deleterious in its homozygous form, has a higher prevalence in the gene pools of populations arising from tropical regions. In 2010, the world birth prevalence of SCT was estimated to be 5,476,000 (IQR: 5,291,000 – 5,679,000) [18]. In support of the historic and possibly ongoing selective advantage of SCT in areas with endemic malaria, high allelic frequencies (>10%) of the sickle mutation were almost exclusively located in sub-Saharan Africa[18]. In the United States, Americans with African ancestry have a much greater prevalence of SCT (8%) than those with European ancestry (0.3%)[19].

2.5 Sickle Haplotypes

Haplotype blocks, long stretches of alleles linked to one another on a single chromosome, are often used to trace genetic ancestry. While the DNA of unrelated individuals will differ at every 1000-2000 base pairs, identical polymorphisms within a haplotype block imply common ancestry[20]. Five distinct haplotypes of the sickle mutation, Cameroon, Benin, Bantu (Central African), Senegal, and Arab, have been discovered, indicating the sickle mutation arose multiple times over the course of human history[20]. The sickle haplotypes are named for their regions of discovery and geographical distribution. Sickle haplotypes are detected by the presence or absence of restriction fragment length polymorphisms (RFLPs) around the beta globin gene cluster[20]. When present, RFLPs are cleaved by the endonucleases Xmn I, Hind III, Hinc II, and Hpa I[21]. While the Benin haplotype extends for 130 kilobase pairs, the Senegal and Bantu haplotypes extend for 76 kilobase pairs[21]. In the United States, the most common sickle haplotypes are Benin, Bantu, and Senegal[22]. Benin haplotypes are by far the

most prevalent, found in 61% of African Americans with SCA.[23] By comparison the Bantu haplotype is found in 21%, and the Senegal haplotype in 10%[23]. These haplotypes are associated with varying expression of the sickle mutation and hemoglobin S production, with Bantu being the most severe, and Senegal the mildest[21]. Variability in severity is related to hemoglobin F production, described below.

2.5a Fetal Hemoglobin and Sickle Haplotypes

Fetal hemoglobin, or hemoglobin F, suppresses the production and polymerization of hemoglobin S, mitigating the sickle phenotype[1]. Hemoglobin F has a greater oxygen affinity than adult hemoglobin, and is structurally distinct, composed of 2 alpha globin chains and 2 gamma globin chains (in the place of beta globin)[1]. At birth, the concentration of hemoglobin F is 55-85%, which normally decreases to less than 1% in adults[1]. However, multiple mutations, both within the *HBB* gene-like cluster and external to it, influence hemoglobin F production[24]. Gamma globin is encoded by *HBG2* and *HBG1*, both located on chromosome 11, within the *HBB* gene like cluster[24]. Unlike the Bantu and Benin haplotypes, the Senegal haplotype block contains a SNP (rs7482144), located 158 base pairs upstream from HBG2, which is cleaved by endonuclease Xmn1 [24]. This polymorphism is thought to be in linkage disequilibrium with functional elements increasing *HBG2* expression; however the exact mechanisms of genetic modification are unknown[24]. There is considerable variability in hemoglobin F and hemoglobin S production among the sickle haplotypes[24]. The Bantu haplotype is considered the most severe, with hemoglobin S comprising 90% of the total hemoglobin in individuals with SCA[21]. The Benin haplotype is intermediate to the Bantu and Senegal haplotypes, with hemoglobin S making up 85% of the total hemoglobin in individuals with SCA, and hemoglobin F 10%.[21] The Senegal haplotype results in the

mildest phenotype of SCA, with hemoglobin S comprising 75% of the total hemoglobin, and hemoglobin F concentrations approaching 20%[21]. However, hemoglobin F is not influenced solely by sickle haplotypes. Quantitative trait loci within the *HBS1L-MYB* intergenic region (rs9399137), and B-cell lymphoma gene (*BCL11A*, rs766432), also influence gamma globin production, explaining 30-50% of the variation in hemoglobin F concentrations among individuals with SCA[24]. The impact of sickle haplotypes and fetal hemoglobin expression in SCT carriers has not been widely published in the literature.

2.6 Sickle Mutation and Hypercoagulability

Both SCA and SCT are associated with increased factors of coagulation, when compared to African Americans with normal hemoglobin[25]. This is likely due to red blood cell asymmetry and disordered cell membranes expressing phosphatidylserine[25, 26]. Under hypoxic conditions, such as heavy exertion, dehydration, or high altitudes, the hemoglobin S of SCT carriers polymerizes, causing erythrocytes to sickle[11, 27, 28]. As the red blood cells sickle, phosphatidylserine, a cellular recognition signal, is pulled from the interior of the phospholipid monolayer to the outside of the cell membrane surface[29]. Phosphatidylserine serves as a docking site for enzymes which facilitate prothrombin activation, coagulation and clot formation[30]. While red cell sickling is far less prevalent in SCT carriers than with SCA, evidence for hypercoagulability in SCT carriers has been indicated by blood assays[26]. Compared to African Americans with normal hemogloblin, SCT carriers have elevated levels of prothrombin fragment 1+2, thrombin-antithrombin complexes, and d-dimer[26]. These biomarkers signify coagulation activation. Prothrombin fragment 1+2 is a byproduct released by the conversion of prothrombin to thrombin, an integral step in the coagulation cascade. D-dimer is a metabolite of fibrinolysis, indicating degradation of fibrin clots, and

an indirect biomarker of hypercoagulability[26]. To some extent, the hypercoagulability in SCT carriers may be modified by environment or behavior. This is evidenced by differing levels of procoagulant activity in monozygotic twins with SCT[31].

Hypercoagulability increases risk of thrombosis, and a greater prevalence of thrombotic infarctions has been observed in individuals with SCT[32-34]. In an autopsy series of 128 SCT carriers with a mean age of 47, obvious visceral infarcts were observed in 18%, after excluding cases with cardiac thrombi, bacterial endocarditis, and atherosclerotic emboli[32]. By comparison, visceral infarcts were observed in less than 1% of similarly aged African Americans with normal hemoglobin. The spleen was the most common site of infarction in SCT cases, followed by the kidneys, brain, and lung[32]. A higher prevalence of thrombotic events has also been observed in SCT carriers in epidemiological studies. In an observational study, 65,154 consecutively admitted African American males in 13 Veterans Administration hospitals were tested for SCT[33]. Pulmonary embolism was noted in 2.2% of SCT carriers, compared to 1.5% with normal hemoglobin[33]. Furthermore, in a case control study of 1070 hospitalized African Americans, SCT carriers were observed to have twice the risk of venous thromboembolism[34]. Certain exposures may influence the risk of thrombosis in SCT carriers. For example, oral contraceptive use has been associated with deep vein thrombosis in female SCT carriers of child bearing age[35].

2.7 Sickle Mutation and Vasculopathy

Vasculopathy, a general term denoting disorders of the vasculature, has been described in individuals with both SCA and SCT. The arterial vasculature is regulated by the endothelial cells, which make up the tunica intima, or endothelium, lining the inner layer of

the arterial wall[2]. When injured or stimulated by irritants, the endothelium becomes activated, releasing cellular signaling molecules that regulate vasomotor tone, inflammatory responses, coagulation, and vascular remodeling[36]. Vascular remodeling may be either dilatative, as is the case with dolichoectasia, or occlusive, the result of hyperplasia, or thickening, or the intimal layer of the arterial wall. Endothelial dysfunction and vascular remodeling associated with the sickle mutation will be described in greater detail next.

2.7a Sickle Mutation and Endothelial Dysfunction

Endothelial dysfunction is detectable by biomarkers and noninvasive testing[37], and has been observed in both individuals with SCA and SCT[38]. Hemolysis, bursting red blood cells, creates endothelial dysfunction by depleting the bioavailability of nitric oxide, a potent vasodilator produced by the endothelium, with anti-athrogenic and anti-thrombotic effects[39]. Sickled red blood cells are particularly friable, and when ruptured, release free hemoglobin and arginase into the blood plasma[39]. Free hemoglobin binds to and scavenges nitric oxide, reducing its bioavailabiliy, while arginase converts L-arginine, a substrate required for nitric oxide production, into ornithine, decreasing nitric oxide formation[39]. Among its many other effects, nitric oxide represses the expression of vascular cell adhesion molecules on the surface of endothelial cells[40]. Accordingly, elevated cellular adhesion molecules can be considered biomarkers of endothelial dysfunction. In laboratory assays, patients with SCA show elevated levels of soluble vascular cell adhesion molecules (sVCAM) compared to African Americans with normal hemoglobin, an indication of endothelial dysfunction[39].

In contrast to homozygous SCA, the erythrocytes of SCT carriers are less prone to rupture[41]. However, after a single sickling episode provoked by hypoxia, SCT erythrocytes

have demonstrated hemolysis *in vitro*.[41] While biomarkers of endothelial dysfunction have not been examined in SCT carriers, endothelial dysfunction in SCT carriers has been suggested by noninvasive tests. One such test, flow mediated dilation (FMD) testing, measures the brachial artery response to induced hyperemia. Briefly, the forearm blood flow is occluded by a tourniquet, and reactive hyperemia, or turbulent, high velocity blood flow, is induced upon release of the tourniquet. The sudden increase in arterial wall sheer stress, caused by the hyperemic blood flow, stimulates the endothelium to release nitric oxide. Endothelial function is gauged by the degree of arterial dilation, measured by high resolution ultrasonography. FMD testing has demonstrated endothelial dysfunction in SCA patients, and a dilation response in SCT carriers that is intermediate to SCA patients and African American controls with normal hemoglobin[38].

2.7b Sickle Mutation and Intimal Hyperplasia

Endothelial dysfunction and overexpression of vascular adhesion molecules induce intimal thickening, or hyperplasia in individuals with sickle hemoglobinopathies. When expressed, the cellular adhesion molecules on the endothelial cell surface recruit white blood cells, and initiate their uptake into the vascular wall[42]. Leukocytes secrete cytokines and growth factors, which induce smooth muscle cells to proliferate and migrate from the tunica media to the intima[43]. The resulting vascular intimal hyperplasia can become stenotic, blocking the passage of blood by protruding into the arterial lumen[42]. Additionally, the irregular surfaces of intimal lesions facilitate thrombus formation, as activated platelets adhere to sites of vascular injury[44]. Arterial intimal hyperplasia with superimposed thrombi have been noted in deceased SCA patients, by post-mortem examinations. An autopsy series 12 SCA patients with fatal stroke at a mean age of 21 reported arterial stenosis and intimal hyperplasia in 7,

and stenosis with occlusive thrombus in 6[45]. Few post-mortem examinations of vasculopathy in individuals with SCT have been published. However, in one report of 20 consecutive patients who died of various causes (15 with SCA autopsied at mean age of 39, and 5 with SCT autopsied at a mean age of 51), intimal hyperplasia and fibrosis was observed in the pulmonary artery of all 20[46]. Thus, intimal hyperplasia may be a common finding with SCT, at least in the pulmonary artery. Unfortunately, no comparisons can be made with African Americans with normal hemoglobin, as none were autopsied in this case series.

2.7c Sickle Mutation and Dolichoectasia

Another form of vasculopathy, dilatative arteriopathy, has been detected by cerebral MRI in both patients with SCA and SCT. Dolichoectasia, or dilatative arteriopathy, is characterized by elongated cerebral vessels following a tortuous, or sinuous course, with marked thinning of the arterial walls and dilation[47]. Bidirectional blood flow within ectatic, or dilated, segments, results in blood stasis and intraluminal thrombus formation, which can lead to downstream infarctions in the cerebral circulation[47]. In older individuals, dolichoectasia is most commonly associated with hypertension and atherosclerosis[47]; however, wall shear stress, induced by acute increases in blood volume and turbulent flow, has precipitated dolichoectasia in experimental animals[48]. Chronic and uneven wall shear stress is thought to denude the endothelium, and degrade the internal elastic lamina of the arterial wall by stimulating the release of metalloproteinases[49]. Sickle cell anemia is characterized by high cardiac output, as the heart compensates for anemia and oxygen demand by pumping a greater blood volume[1], which may be related to the increased prevalence of dolichoectasia observed with SCA. In a retrospective study of 47 children with SCA who were imaged by magnetic resonance angiography, 37% demonstrated ectasia of the basilar artery, compared

to 2% of controls without SCA, who were imaged to assess brain tumor[50]. While heterozygous SCT is not characterized by anemia or high output states[1], a prospective imaging study comparing SCT children to normal hemoglobin siblings reported a dolichoectasia prevalence of 19% in those with SCT, compared to none of the controls[51]. Thus, both the homozygous and heterozygous sickle mutation may be associated with dolichoectasia; however the mechanisms causing dolichoectasia in SCT carriers have not been elucidated.

2.8 Sickle Cell Trait and Other Hemoglobinopathies

Several hemoglobinopathies, or mutations affecting alpha or beta globin production, are observed in populations of African descent. These include alpha thalassemia, beta thalassemia, and Hemoglobin C (Table 1). The SCT expression is influenced by co-inherited hemoglobinopathies. While alpha thalassemia mitigates the phenotype, concomitant heterozygosity for SCT and beta thalassemia or SCT and hemoglobin C produces a disease state phenotypically similar to sickle cell anemia, exacerbating the impact of the SCT variant[1]. The hemoglobinopathies affecting African Americans and their coinheritance with SCT are described in greater detail next.

Table 1: Hemoglobinopathies and their prevalence in African Americans[1, 19, 52, 53]

Hemoglobinopathy	Gene(s)	Mutations	Presentation	Prevalence
Alpha Thalassemia (-- α / α α) or (--α / -- α)	*HBA2* Chromosome 16	Δ3.7 kb deletion (heterozygous or homozygous)	Asymptomatic to mild microcytic anemia	30%
Beta Thalassemia	*HBB* Chromosome 11	-88 (C-T) or -29 (A-G)	Asymptomatic (heterozygotes) to severe anemia (homozygotes)	1%
Sickle Cell Trait	*HBB* Chromosome 11	Glu6Val (heterozygous)	Carrier state	8%
Sickle Cell Anemia	*HBB* Chromosome 11	Glu6Val (homozygous)	Severe hemolytic anemia	0.14%
Hemoglobin C Trait	*HBB* Chromosome 11	Glu6Lys (heterozygous)	Asymptomatic	2%
Hemoglobin C Disease	*HBB* Chromosome 11	Glu6Lys (homozygous)	Mild hemolytic anemia	0.017%
Hemoglobin SC Disease	*HBB* Chromosome 11	Glu6Val, Glu6Lys	Severe hemolytic anemia	0.08 - 0.12%
Sickle β Thalassemia	*HBB* Chromosome 11	Glu6Val, β thalassemia	Mild to severe hemolytic anemia, depending on β^0 or β^+ mutation	0.02% - 0.05%
Sickle Cell Trait Alpha Thalassemia	*HBB, HBA2*	Glu6Val and Δ3.7 deletion	Attenuated expression of sickle cell trait	2.4%
Sickle Cell Anemia Alpha Thalassemia	*HBB, HBA2*	Homozygous Glu6Val and Δ3.7 deletion	Milder form of sickle cell anemia	0.042%

(-- α / α α) = 3 functioning alpha globin genes, (-- α / -- α) = 2 functioning alpha globin genes
Glu6Val = glutamic acid to valine mutation, Glu6Lys = glutamic acid to lysine mutation
β^0 = no beta globin production, β^+ = diminished beta globin production

2.8a Alpha Thalassemia

Alpha thalassemias result from deletions to the alpha globin encoding genes (*HBA2* or *HBA1*), or mutations affecting their expression. The alpha globin gene cluster spans 30 Kb, and contains 5 genes: hemoblobin, zeta (*HBZ*); hemoglobin zeta pseudogene 1 (*HBZP1*); hemoglobin, mu (*HBM*, also referred to as hemoglobin alpha pseudogene 2, or *HBAP2*); hemoglobin alpha pseudogene 1 (*HBAP1*); hemoglobin alpha 2 (*HBA2*); hemoglobin alpha 1 (*HBA1*); and hemoglobin theta 1 (*HPQ1*), as shown in Figure 7.

Over 120 mutations are associated with alpha thalassemia; however, the $\Delta^{3.7}$ deletion

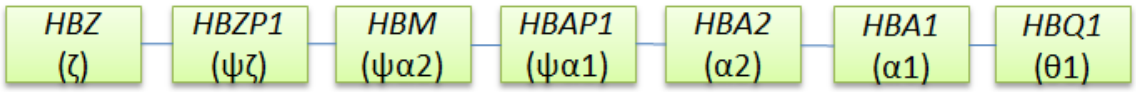


Figure 2: The alpha globin gene cluster contains 6 loci
HBZ = Hemoglobin Zeta; *HBZP1* = Hemoglobin Zeta Pseudogene 1; *HBM* = Hemoglobin Mu; *HBAP1* = Hemoglobin Alpha Pseudogene 1; *HBA2* = Hemoglobin Alpha 2; *HBA1* = Hemoglobin Alpha 1; *HBQ1* = Hemoglobin Theta

(rs63751476) of the *HBA2* gene is most common in African Americans. African Americans have a high prevalence (30%) of alpha thalassemia, defined by either a single (-- α / α α) or double (-- α / -- α) alpha globin gene deletion, denoting either heterozygosity or homozygosity for the $\Delta^{3.7}$ mutation[52]. The clinical presentation of alpha thalassemia trait ranges from asymptomatic to mild microcytic anemia, characterized by small red blood cells[1]. The more severe forms of alpha thalassemia, Hemoglobin H (-- -- / -- α) and Barts Hydrops Fetalis (-- -- / -- --) are exceedingly rare in African Americans, and as a result the prevalence has not been estimated[52].

When coinherited with SCT, alpha thalassemia decreases hemoglobin S production[1]. Because alpha thalassemia minimizes the available pool of alpha globin chains, normal and

sickle beta globin must compete for available alpha chains to form hemoglobin tetramers[1]. Normal beta globin dimerizes twice as effectively as sickle beta globin, leaving sickle beta globin without available alpha chains to form hemoglobin S[1]. In fact, hemoglobin S concentrations in SCT carriers depend on the number of functioning alpha globin genes. SCT carriers with all 4 alpha globin genes have hemoglobin S concentrations of about 40%, compared to 35% in SCT carriers with 3 functioning alpha globin genes, and 30% in those with 2 alpha globin genes[52].

The prevalence of alpha thalassemia in patients with SCA appears to be comparable to the general population of African Americans[52, 54, 55]. Alpha thalassemia has been estimated to have a prevalence of 30% in African Americans[52]. Similarly, a 1 or 2 alpha gene deletion was detected in 33% of SCA patients enrolled in the Cooperative Study of Sickle Cell Disease[54], but was slightly higher (40%) in infants with SCA enrolled in the Baby Hug trial[55]. Alpha thalassemia is inherited independently of sickle haplotype groups as well[21]. Assuming a random coinheritance of alpha thalassemia with respect to sickle mutations or haplogroups, the expected prevalence of SCT and alpha thalassemia would be the product of each muation's prevalence in African Americans, or 2.4%.

2.8b Beta Thalassemia

Beta thalassemias are hemoglobinopathies that diminish (β^+) or abrogate (β^0) production of the beta globin chains. When coinherited with SCT, beta thalassemia increases the concentration of hemoglobin S. This is because beta thalassemia decreases the number of normal beta globin chains, leaving a predominance of sickle beta globin chains[1]. Sickle beta thalassemia, the coinheritance of SCT and β^+ thalassemia, results in mild hemolytic anemia[1].

On the other hand, SCT that is coinherited with β^0 thalassemia has a clinical presentation that is almost indistinguishable from sickle cell anemia[1]. Over 200 mutations are known to cause beta thalassemia. In African Americans, the -88 (C→T) and -29 (A→G) mutations, affecting promoter regulatory elements of the *HBB* gene, are the most frequent cause of beta thalassemia[56]. In African American birth cohorts, the prevalence of beta thalassemia is 1%[53], while the prevalence of coinherited SCT and beta thalassemia is 0.02 – 0.05%[19, 53].

2.8c Hemoglobin C

Similar to hemoglobin S, hemoglobin C is a functional variant of the beta globin gene (rs33930165), caused by a single nucleotide polymorphism of the 6th codon of the hemoglobin gene[1] (Figure 8). While sickle beta globin contains a glutamic acid to valine substitution, hemoglobin C produces a mutated beta globin protein with lysine in the place of glutamic acid[1]. The heterozygous genotype, or hemoglobin C trait, is asymptomatic. The homozygous genotype, known as hemoglobin C disease, is characterized by mild hemolytic anemia. In African Americans, the prevalence of hemoglobin C trait is 2%, while the prevalence of hemoglobin C disease is much rarer, at 0.017%[19]. Concomitant heterozygosity of SCT and hemoglobin C, however, results in hemoglobin SC disease, a hemolytic anemia with symptoms of vaso-occlusive crises and acute chest syndrome[57]. In African American birth cohorts, the reported prevalence of hemoglobin SC disease in ranges from 0.008 – 0.12%[19].

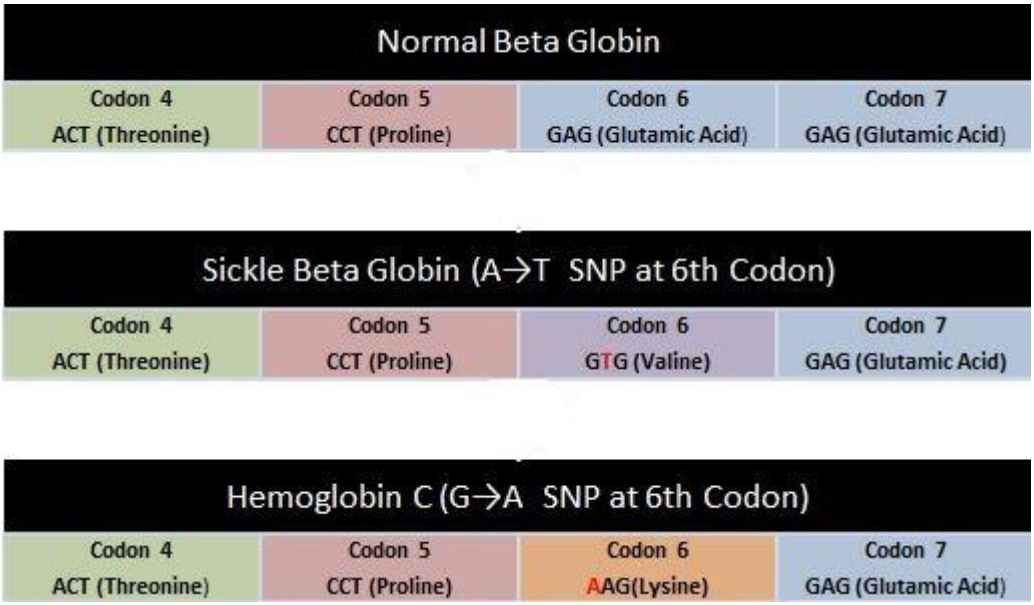


Figure 3: Single nucleotide polymorphisms on 6th codon of beta globin gene

2.9 Summary of Primary Exposure: SCT

Sickle cell trait, the heterozygous form of sickle cell anemia, affects approximately 8% of African Americans. Laboratory assays and observational studies suggest greater risk of adverse vascular events associated with SCT. Under hypoxic conditions, the hemoglobin S polymerizes, increasing hypercoagulability and risk of venous thromboembolism in SCT carriers. The sickle mutation is also associated with vasculopathies, and a greater degree of endothelial dysfunction, intimal hyperplasia, and dolichoectasia. However, the sickle mutation is affected by several genetic factors. Hemoglobin F production, influenced by polymorphisms within the sickle haplotype blocks, and coinheritance of alpha thalassemia, reduce the production of hemoglobin S, mitigating the severity of the SCT phenotype. On the other hand, a concomitant heterozygosity of SCT and Hemoglobin C, or the coinheritance of SCT with β^0 thalassemia, results in a phenotype that more closely resembles sickle cell

anemia. Thus, the hypothesized associations between SCT and cerebrovascular disease may be influenced by concomitant hemoglobinopathies and haplotype.

3. STROKE

3.1 Overview

Stroke, or cerebrovascular accident, is an acute neurologic syndrome provoked by interrupted blood supply to the brain[58]. The cerebral neurons lack glycogen; as a result, blood flow cessation precipitates rapid energy failure followed by brain tissue death within 4 – 10 minutes[58]. Blood flow disturbances originate from vascular obstructions and ruptured blood vessels, and occur in both the large and small cerebral vessels[58]. The large intracranial arteries include the internal carotid arteries, vertebral arteries, and the middle, anterior, and posterior cerebral arteries, which make up the Circle of Willis. The small vessels, which typically range from 400 – 900 μ in diameter, branch from the large vessels deep into the interior of the brain[58]. While most cerebral vascular accidents involve the large vessels, the small vessels are implicated in 20%[58]. Stroke more often involves the arterial circulation; however, cerebral vascular accident is also provoked by disrupted blood flow in the cerebral veins and sinuses[58].

3.2 Ischemic Stroke

The majority of strokes are ischemic, accounting for 87% of all cerebrovascular accidents[59]. Ischemic strokes are caused by vascular obstructions from clot or atheroma, with origins that are more often embolic than in-situ[58]. Common sources of emboli include the heart and proximal extracranial arteries, and to a lesser extent, the venous circulation[58]. Cardioembolic stroke, provoked by embolized blood clots originating in the heart, account

for 20% of all ischemic strokes[58]. Clotting is induced by stasis of the left atrium, due to low ejection fraction, atrial fibrillation, and left atrial dilation. Another source of embolic stroke is the extracranial carotid arteries[58]. The bifurcation of the internal carotid is particularly vulnerable to atherogenesis. Thrombus forming on the plaque surface can dislocate, or the plaque itself can rupture and break loose, forming emboli that eventually become lodged in the intracranial vessels. A third source of emboli is the venous circulation, causing ischemic stroke by so-called paradoxical embolization[58]. Under normal circumstances, blood from the systemic venous circulation is returned to the right atrium, and ejected by the right ventricle to the lungs. However, in 15% of the general population, blood communicates from the right to left atrium, through a small opening in the atrial wall, or atrial septal defect[58]. This passage allows embolized clot from the venous circulation, often from deep vein thrombosis, to cross the cardiac circulation and become ejected into the systemic arterial circulation. Compared to embolic stroke, vascular obstruction by in-situ thrombosis is less frequently observed, and usually occurs in either the small deep penetrating arteries, or the cerebral veins and sinuses[58]. Stenotic plaques, typically in the carotid or vertebral arteries[58], also provoke ischemic stroke by inducing hypoperfusion to the downstream cerebral circulation. In the United States, the 30-day case fatality rate of ischemic stroke is 7.8%[60]

3.3 Hemorrhagic Stroke

Cerebrovascular accidents that arise from ruptured blood vessels, rather than vascular obstructions, are termed hemorrhagic strokes. Hemorrhagic strokes are particularly deadly, with a case fatality rate of 50%[58]. The most common type of hemorrhagic stroke is hypertensive intraparenchymal hemorrhage, invoked by spontaneous rupture of the small, deep penetrating arteries[58]. Following arterial rupture, blood seeps into the ventricular spaces

of the brain, causing hydrocephalus and increasing intracranial pressure perniciously[58]. Hemorrhagic stroke is sometimes characterized by bleeding into the subarachnoid space between the brain tissue and the pia mater membrane surrounding the brain[58]. Subarachnoid hemorrhage may result from ruptured aneuryms, weakened areas of the arterial wall, and generally arise from the large vessels comprising the Circle of Willis[58]. Other causes of subarachnoid hemorrhage include trauma, and congenital arteriovenous malformations[58]. Arteriovenous malformations allow abnormal communication between the arteries and veins, bypassing the capillary circulation, and are prone to rupture[58].

3.4 Stroke Symptoms

Symptoms and severity of stroke depend on type of obstruction, size of vessel impacted, extent of collateralization, and brain region supplied by the arterial distribution[58]. Symptoms are almost always contralateral to the brain hemisphere affected, with functional disability reflecting the brain region impacted[58]. Limb paralysis (hemiplegia), limb weakness (hemiparesis), impaired gait (ataxia), and clumsy hand suggest damage to the sensory and motor cortex[58]. Aphasia (language deficit) implies damage to Broca's or Wernicke's area of the cerebral cortex, while homonymous hemianopia (visual field loss) indicates injury to the cerebral visual cortex[58]. The stroke syndrome may also present with severe memory loss, stupor, coma, loss of consciousness, and seizure[58].

Infarction of the small, deep penetrating arteries, or lacunar stroke, is typically less severe than stroke related to the large vessels. In fact, small vessel infarcts are almost always ruled out by presentation of severe memory loss, stupor, coma, aphasia, monoplegia, homonymous hemianopia, loss of consciousness, or seizures[61]. The vascular distributions of

the small vessels are less extensive than the large vessels, minimizing areas affected by infarction. Lacunar infarcts are also more often due to thrombi, which occasionally lyse, restoring the passage of blood flow[58]. Episodes of stroke symptoms lasting less than 24 hours, considered transient ischemic attacks (TIA), are often provoked by blood clots of the small vessels that quickly lyse, allowing functional recovery[58]. However, TIAs are not without risk; in 10-15% of TIA cases, major stroke occurs within 3 months[58].

3.5 Stroke Diagnosis

Acute stroke is suspected by abrupt changes in mental status with presentation of neurological symptoms, and diagnosed by imaging studies and assessment of cerebral spinal fluid by lumbar puncture[58]. Hemorrhagic stroke is distinguished from ischemic stroke by the presence of blood in the cerebral spinal fluid, and hemorrhaging in the absence of infarction. Computed tomography (CT) is generally preferred for acute stroke diagnosis, due to its speed and lower cost than magnetic resonance imaging (MRI)[58]. CT imaging reliably detects intracranial hemorrhaging, but may not visualize infarctions until 24-48 hours after the event[58]. Small infarctions along the cortical surface of the brain or within the posterior fossa may also be missed, due to bone artifact obstructing the image[58]. As an alternative to CT, cerebrovascular accident may be diagnosed with MRI. While MRI has less sensitivity visualizing hemorrhagic bleeding, it is a superior imaging modality for the detection of infarctions in all regions of the brain[58]. One limitation to stroke diagnosis, either by CT or by MRI, is the patient's ability to seek medical care. Stroke symptoms are frequently painless, or cause anosagnosia, the lack of perception that anything is wrong[58]. As a result, patients who live alone may not have the foresight to call emergency medical services for help.

3.5a. Stroke Diagnosis in Epidemiological Studies

In large observational studies, stroke cases are often classified by administrative claims records, hospital surveillance, chart reviews, or diagnostic algorithms. The 9th International Classification of Disease (ICD-9) codes, submitted to insurance carriers for billing purposes, provide a facile but imperfect record of patients discharged with stroke. To ascertain stroke diagnoses, epidemiological studies sometimes examine abstracted medical records, including the physician notes and diagnostic imaging reports[60]. One disadvantage of retrospective data collection after hospital discharge is the possibility of missing data relevant to the study aims. Unlike prospective study designs, which rigorously record clinical data related to pre-specified outcomes, retrospective surveillance, or "cold pursuit", is limited to recorded data deemed relevant to the patients' clinical care[62]. For this reason, stroke diagnoses by chart reviews often include measures of uncertainty. For example, stroke may be considered "definite" when classified by clinical symptoms, imaging studies, and lumbar puncture. On the other hand, stroke may be classified as "probable" when the clinical picture suggests stroke, but the diagnosis does not meet certain criteria for the epidemiological definition, for example if imaging studies are inconclusive, or lumbar puncture is not performed[60]. Agreement between ICD-9 diagnoses and chart reviews varies, based on the stroke subtype and the number of ICD-9 codes analyzed. In a study of 206 patients hospitalized with stroke, ischemic stroke was more accurately classified by considering all ICD-9 codes in the discharge record, while intracerebral and subarachnoid hemorrhage were better classified by considering only the primary ICD-9 discharge code[63]. Physician-reviewer validation of ICD-9 codes has also been ascertained in the ARIC study, which reported greater agreement between ICD-9 codes and subarachnoid hemorrhage (86%), than

ICD-9 codes and ischemic stroke (78%)[60]. In some epidemiological studies, reviewing the medical record may be too labor intensive or cost-prohibitive. As an alternative, a validated computer algorithm may be utilized to diagnose stroke cases, based on neurological signs and symptoms[64]. In the ARIC study, which examined 538 stroke hospitalizations, agreement between computer algorithm and physician reviewer diagnoses was 78%[60].

3.5b. Cerebral Magnetic Resonance Imaging

Cerebral small vessel disease is often detected in clinical settings and observational, prospective screenings by magnetic resonance imaging. MRI scanners use magnetic field sequences to manipulate hydrogen protons, and convert the signals from protons returning to baseline energy states into images. Hydrogen atoms are present in 99.98% of all human tissues[65]. Positive subatomic particles, or protons, continuously rotate within the nuclei of atoms, creating a magnetic field. When atoms are subjected to an external magnetic field, the protons start to rotate around the axis of the magnetic field, in a direction that is either parallel or antiparallel to the field[65]. This interaction creates magnetic resonance. After the MRI scanner aligns atoms within a magnetic field, a second magnetic field, or excitation pulse, is applied, pulling the atoms to a specified angle that is transverse to the first magnetic field. The excitation pulse is then switched off, and protons are monitored as they return to the original alignment[65]. The return time is dependent on the tissue. For example, the relaxation time for hydrogen protons in fat is very brief, while the relaxation time in fluids is very long[65]. These return time signals are converted into voxels, or 3D pixels, by fast Fourier transform, allowing a visual representation of the tissues. For image optimization, the magnetic resonance signal can be filtered, or weighted, by manipulating the repetition time and echo time of the excitation pulse[65]. Two sequences, called T1 and T2 weights, may be

used, depending on the structures being visualized. When both the T1 and T2 weights are removed, MRI is considered proton density weighted, meaning the magnetic resonance signals are proportional to the water concentration[65]. Brain imaging is also optimized by fluid attenuated inversion recovery, or FLAIR, which filters signals arising from the cerebral spinal fluid[65].

3.5c. Subclinical Stroke Detection by MRI

Silent stroke is unaccompanied by neurological symptoms, but is detectable by a presentation of white matter lesions or subclinical brain infarctions on MRI.[66] Severity of white matter lesions is either quantified volumetrically, or graded qualitatively from 2-dimensional images[67]. Subclinical brain infarctions, which present on MRI as fluid-filled lacunes, are quantified by measuring the largest diameters[68]. In highly selected populations, such as patients with Alzheimers[69] or Binswangers disease[70], MRI-detected WML and lacunes have been validated by autopsy. However, silent stroke detected by MRI in the general population is not likely to be validated by autopsy, and few cohorts incorporate post-mortem examination with prospective MRI. Despite this limitation to validity, the detection of subclinical cerebrovascular disease by MRI is reproducible (discussed further in sections 4.1 and 5.1). Reliability is typically tested among randomly selected duplicate images, to determine intra- and inter-reader reproducibility[67, 71, 72]. An advantage of MRI is its suitability for prospective imaging in observational studies[58]. Unlike CT, which exposes a patient to 3-5 mGy of radiation for a routine brain exam, MRI does not utilize ionizing radiation[58]. However, MRI is time consuming and sensitive to motion, and children under the age of 10 typically require conscious sedation to remain immobile for the duration of the test[58]. In addition, metal objects must be removed prior to MRI examination[65]. Study

participants with metallic implants, such as aneurysm clips, hearing aids, cardiac pacemakers, spinal cord stimulators or other internal electrical devices must be excluded. Other considerations include extreme obesity and claustrophobia. As the MRI scanning involves bodily insertion into the magnet bore, patients unable to fit, or tolerate, confined spaces may need to be excluded.

3.6 Stroke Prevalence

In the United States, approximately 7 million people over the age of 20 have had a stroke[59]. The Southeast has the greatest density of stroke cases, clustered within the so-called stroke belt: Alabama, Arkansas, Georgia, Indiana, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, and Virginia[73]. Stroke occurrence increases with age, and is more prevalent in women than men. At the ARIC study baseline, the prevalence of TIA or stroke symptoms in women was 7%, compared to 5% in men ($p<.0001$)[64]. However, the greater prevalence of stroke in women appears to be predominantly due to the longer lifespans of women[74]. African Americans are disproportionately burdened by cerebrovascular disease, and experience first stroke at a younger mean age than European Americans[75]. The prevalence of stroke in African Americans18 years or older is nearly twice that of than non-Hispanic whites (4.0% vs. 2.3%)[76]. Annually, the incidence of stroke in the United States has been estimated to be about 800,000 cases per year[59], and has been decreasing since the 1990s for European Americans. However, this trend has not been observed for African Americans[77]. Over 8 years of follow up, a longitudinal analysis from the ARIC study reported a 62% higher stroke incidence rate in African Americans than European Americans (incidence rate ratio = 1.62; 95% CI: 1.22 – 2.14), after controlling for

age, gender, hypertension, diabetes, center, education, smoking, and history of coronary disease[60].

3.7 Stroke Risk Factors

Risk factors for stroke vary by stroke type; however, hypertension is a strong determinant of both ischemic and hemorrhagic stroke[58]. Additional risk factors associated with ischemic stroke include older age, diabetes, smoking, hyperlipidemia, atrial fibrillation, and recent myocardial infarction. On the other hand, hemorrhagic stroke is correlated with aging, hypertension, trauma, congenital arteriovenous malformation, and amyloid angiopathy, the deposition of amyloid proteins within the brain and arterial walls[58].

Endothelial dysfunction and dolichoectasia may increase the risk of lacunar stroke. While endothelial heterogeneity is known to exist among the various vascular distributions[36], there is evidence that systemic endothelial dysfunction correlates with cerebral small vessel disease and lacunar stroke. In a cross sectional study of 45 Korean patients with symptomatic lacunar stroke undergoing flow mediated dilation testing, greater endothelial dysfunction of the brachial artery was noted in the stroke patients than the age and sex matched controls with hypertension[78]. Lacunar stroke has also been cross sectionally associated with intracellular adhesion molecule (ICAM-1), tissue factor, and thrombomodulin, all biomarkers of endothelial dysfunction[79]. Dolichoectasia appears to be a risk factor for lacunar stroke, as well. In a population-based study of 387 patients with first cerebral infarction, 42% with dolichoectasia were diagnosed with lacunar infarct, compared to 17% without dolichoectasia[80].

3.7a. Genetic Risk Factors of Stroke

Family history and twin studies have predicted both ischemic and hemorrhagic stroke, implying a genetic component of stroke risk[59]. Of the few studies that have reported genome-wide associations between SNPs and stroke, none have been conducted in populations of African ancestry. However, in European populations, variants of the histone deacetylase (*HDAC9*) gene[81, 82], paired-like homeodomain (*PITX2*) gene[82], and Ninjurin (*NINJ2*) gene[83] have been shown to have validated associations with large vessel stroke, at genome-wide statistical thresholds (p-value of 10^{-8} or smaller).

3.8 Sickle Cell Anemia and Stroke

Sickle cell anemia is a well-established genetic risk factor for stroke. The Cooperative Study of Sickle Cell Disease, a longitudinal cohort of 4,082 patients with sickle cell disease (sickle cell anemia, sickle hemoglobin C, and sickle thalassemia), observed that by the age of 45, 24% of patients with sickle cell anemia had experienced a first stroke, defined ischemic stroke, hemorrhagic stroke, or TIA[84]. This association is confirmed by administrative claims data for stroke. The incidence of ischemic stroke in patients with sickle cell anemia aged 35-64 is 740 per 100,000 person years[85], much higher than 270 per 100,000 person years for 35-64 year old African Americans overall[86]. Unlike the general population, ischemic stroke in patients with SCA is often due to arterial narrowing with thrombus that is in-situ, rather than embolic. In an autopsy series of children with SCA and fatal stroke, intimal hyperplasia with superimposed thrombosis were common findings, particularly in the middle and anterior cerebral arteries[45]. In-situ thrombosis of the small cortical veins and the lateral and sagittal sinuses, the site of venous confluence within the

brain, are also observed with SCA[58]. Stroke risk in children with SCA is determined noninvasively by transcranial Doppler, used to detect stenotic lesions and increased blood flow velocities within the Circle of Willis[87].

3.8a Genetic Modifiers of Sickle Cell Anemia and Stroke Risk

The stroke risk associated with sickle cell anemia is heterogenous, and may be influenced by coinheritance of modifier genes. Support for this hypothesis comes from a pediatric sibship study, examining siblings who were all affected by sickle cell anemia[88]. In 42 of the 207 sibships, at least one child experienced clinical stroke, and in 10 sibships, 2 siblings had a stroke. The number of families with 2 children with SCA and stroke exceeded expectations derived by permutation analysis, leading the authors to conclude that coinheritance of genetic modifiers shared among siblings with SCA increased stroke risk[88]. The sickle haplogroups, alpha thalassemia, and candidate genes associated with stroke in the general population have been investigated as potential genetic modifiers of the association between sickle cell anemia and stroke, and are described next.

3.8b Sickle Cell Anemia and Stroke Risk Modification by Haplotype

Stroke risk modification by haplogroup polymorphisms remains uncertain. In a cross sectional study of 221 patients with sickle cell anemia, the Bantu haplotype, the most severe SCA phenotype with the greatest concentration of hemoglobin S, was associated with a 3-fold increase in obliterative sickle vasculopathy, an outcome that included stroke, renal failure, chronic lung disease with cor pulmonale, leg ulcers, and young adult death[21]. However, a similarly sized case-control study that examined stroke as an explicit, rather than composite, outcome, found no differences in stroke prevalence among the 5 haplotype

groups[89]. While statistical power calculations were not described, this study included 130 children with SCA and stroke, and 103 SCA controls without stroke[89]. Stroke risk modification by sickle haplogroups remains an interesting hypothesis, but at the current state of science, is far from established.

3.8c Sickle Mutation and Stroke Risk Modification by Alpha Thalassemia

Concomitant alpha thalassemia may also modify stroke risk in patients with SCA, and appears to protective effect[21, 89, 90]. In a cross sectional study of 300 children with SCA, the prevalence of confirmed alpha thalassemia in children experiencing stroke (21%) was lower than those who were stroke free (38%)[90]. These findings were later replicated by a pediatric case-control study including 233 children with SCA and alpha thalassemia detecting by genotyping[89]. In SCT carriers, stroke risk modification has been investigated less extensively. In a single study examining 355 hospitalized black men with SCT, no differences in stroke risk were observed among quartiles of hemoglobin S concentrations[91]. While alpha thalassemia was not ascertained by genotyping, it was inferred by hemoglobin S concentrations and mean corpuscular volume. However, these biomarkers may be confounded by iron-deficiency anemia[91], leaving the results inconclusive.

3.8d Sickle Cell Anemia Stroke Risk Modification by Candidate Genes

Several candidate genes known to influence the development of vascular disease have been validated as stroke risk modifiers in children with sickle cell anemia. In the Cooperative Study of Sickle Cell Disease (CSSCD), which included 1,398 African Americans with SCA, 108 SNPs in 80 candidate genes were analyzed by Bayesian networks for associations with stroke[92]. From this analysis, 25 SNPs in the adenylate cyclase

(*ADCY9*), annexin (*ANXA2*), bone morphogenic protein (*BMP6),* chemokine ligand (*CCL2*), colony stimulating factor (*CSF2*), endothelin converting enzyme (*ECE1*), v-ets virus oncogene (*ERG*), met-proto oncogene (*MET*), selectin p (*SELP*), Tek tyrosine kinase (*TEK*), transforming bone growth factor beta receptor (*TGFBR3*) and tumor necrosis factor (*TNF-α*) genes were found to be associated with stroke[92]. Genetic associations were then validated in 114 African Americans with SCA not included in the original Bayesian analysis, and were found to predict stroke over a 5 year follow up with an accuracy of 98.2%[92].

Candidate genes have also been associated with large and small vessel stroke subtypes, classified by MRI. In a validation study involving the CSSCD cohort and Stroke Prevention Trial in Sickle Cell Anemia (STOP), variants of tumor necrosis factor (*TNF* -308 G/A) and interleukin 4 (*IL4R* 503 S/P) were found to protect against and increase the risk of large vessel disease, respectively[93]. Genetic polymorphisms have also been associated with small vessel disease, in children with SCA prospectively screened by MRI. In the CSSCD study, variants of the vascular cellular adhesion molecule (*VCAM1* -1594) and the major histocompatibility complex, class II (*HLA-DPB1*) were associated with small vessel infarctions and silent stroke[94]; however, replication studies are needed to confirm these associations.

3.9 Stroke Burden

Stroke is a leading cause of both long-term disability and death in the United States, and African Americans shoulder the greatest burden. Severe to moderate disability has been reported in 43% of elderly stroke survivors participating in the Framingham Study, which quantified activities of functional living using the Barthel scale[95]. However, the Framingham

study did not include African Americans. The National Health and Interview Survey, a more representative sample of 1,613 stroke survivors, revealed greater activity limitations in African Americans than European Americans, after adjusting for age, sex, and income. African Americans were more likely to report difficulty standing for more than 2 hours, pushing or pulling large objects, walking a quarter of a mile, stooping, bending, or kneeling, and walking up 10 steps without resting[96]. While stroke is the fourth overall leading cause of death in the United States[97], this too is disproportionately high for African Americans. In 2009, the age-adjusted stroke death rate per 100,000 was 62 for African American men, compared to 38 for non-Hispanic White men[97]. Racial disparities were also noted for women; the adjusted death rate for African American women was 51, compared to 37 for their non-Hispanic White counterparts[97].

Stroke is not only twice as prevalent in African Americans compared to European Americans, it occurs at a younger age, resulting in substantial loss of potential earnings[75]. When direct and indirect costs are considered, including ambulance services, hospitalization, rehabilitation, outpatient clinics, prescription drugs, nursing homes, informal caregiving, and potential lost earnings, the projected cost of stroke from 2005-2050 is $1.52 trillion for non-Hispanic whites, and $379 billion for African Americans[75]. However, the per capita cost of stroke is projected to be $25,782 for African Americans, substantially higher than $15,597 for European Americans[75].

3.10 Summary of Stroke

Stroke is a leading cause of disability and death in the United States, and African Americans are more often affected. Stroke may be ischemic or hemorrhagic, and may arise

from the large vessels, small deep penetrating arteries, or veins and sinuses. Stenotic atherosclerosis and embolized blood clots from the heart and venous circulation increase stroke risk, while in-situ thrombosis of the small vessels and vasculopathies contribute to lacunar stroke and TIA. Sickle cell anemia is an established genetic risk factor for stroke, associated with intimal hyperplasia and narrowing of the large cerebral vessels, with superimposed in-situ thrombosis. However, the stroke risk associated with SCA may be modified by haplotype group, concomitant alpha thalassemia, and candidate genes associated with vascular disease in the general population. The associations between SCT and stroke, and modification by candidate genes, alpha thalassemia, and haplotypes are largely unexplored.

4. SUBCLINICAL BRAIN INFARCTIONS

4.1 Overview

Lacunar infarcts, or blockages of the small, deep, penetrating arteries of the brain, are often considered subclinical brain infarctions (SBI) when asymptomatic. Upon healing, subclinical brain infarctions leave behind small fluid-filled cavities, or lacunes, of necrosed cerebral tissue. Lacunes typically measure 3 – 20 mm in diameter. In epidemiological studies, SBI are detected by prospective MRI, using T1 and T2 weighted imaging in the gray matter, and T1 weighted imaging in the white matter. While MRI-detected SBI are seldom validated by post-mortem examination, image reliability has been reported, with intra- and inter-reader reproducibility measures of 82% and 79%, respectively[71].

4.2 Etiology of Subclinical Brain Infarctions

The etiology of SBI appears to be heterogeneous, and related to vascular obstruction by atheroma, blood clots, and small artery stenosis[98]. Brain autopsies identifying the source of lacunar infarctions are infrequently done, as this requires uninterrupted serial sections of the basal ganglia, tracing the infarcted area proximally until an occlusion is found[98]. However, from the few published autopsy reports, obstructions of the penetrating arteries appear to be due to embolic particles of atheroma, or to hyalinosis, an adaption to hypertension that results in hypertrophic growth of the arterial walls that eventually occludes the lumen[61]. In lacunar infarcts with no obstructions identified by autopsy, thrombus has been implicated, as blood clots lyse and disappear within days of an embolic event[98]. Occlusion by thrombus is further supported by an association between subclinical brain infarctions and factors of coagulation (d-dimer and von Willebrand factor) in a cross sectional sample of older adults from the ARIC study[99].

4.3 Prevalence of Subclinical Brain Infarctions

The prevalence of SBI appears to increase with advanced age, and may be higher in African Americans than European Americans[71, 100]. During the baseline MRI exam of the ARIC study, 1890 participants aged 55-72 were prospectively imaged by brain MRI. The overall prevalence of SBI was 15%, and increased with age from 8% in 55-59 year olds to 23% in 65-72 year olds[71]. The SBI prevalence was also shown to be higher in African Americans (21%), compared to European Americans (10%). Similar trends were observed in the Cardiovascular Health Study, which prospectively imaged 3,658 participants aged 65 and older, using identical definitions for SBI as the ARIC study. The SBI prevalence increased

from 22% in 65-69 year olds, to 43% for those over the age of 85 years[100]. However, no differences in SBI prevalence was observed between African Americans and European Americans in the CHS study.

4.4 Risk Factors of Subclinical Brain Infarctions

Risk factors for SBI have primarily been inferred from cross-sectional studies of the general population imaged by cerebral MRI. Age and hypertension were the most consistently reported risk factors in a systematic review of 21 cross-sectional and 3 longitudinal analyses of SBI[66]. However, traditional risk factors for stroke, such as sex, cardiovascular disease, smoking, and diabetes have been reported as well[66]. The risk factors for SBI may vary, due to heterogeneous etiologies of microatheroma, hyalinosis, and thrombosis. A cross-sectional analysis from the ARIC study, which examined small (3–7 mm) and large (8-20 mm) lacunes separately, found age, race, hypertension, smoking, and diabetes to predict small lacunes, while age, hypertension, smoking, and LDL cholesterol predicted large lacunes[68]. The authors concluded that hyalinosis is predominantly responsible for the small SBI, and atheroma putatively causing the large lesions[68]. However, lacunes measuring 5 – 20 mm have been detected by autopsies, with thrombus implicated as the cause[98]. As with stroke, cardiac risk factors appear to be associated with SBI; however, the size of the lacunes may not always discriminate the etiology.

4.5 Sickle Cell Anemia and Subclinical Brain Infarctions

Genetic risk factors may also predispose development of SBI. The sickle mutation appears to be a genetic risk factor for lacunes, at least in homozygotes[101]. In the CCSCD study, SBI was the most common neurological manifestation observed in children with

SCA[101]. However, few studies have examined this outcome, due to the cost of magnetic resonance imaging, sedation requirement in small children, and lack of effective interventions[101]. From the few studies that have been conducted, SBI appears to affect 27% of children with SCA under the age of 6, and 37% under the age of 14[101]. While SBI is not accompanied by overt signs of clinical stroke, it is associated with poor academic achievement and future stroke in children with SCA[101]. The prevalence of SBI in adults with SCA is not well established. In a cross sectional study of 149 patients with a mean age of 32 years, lacunes were detected in 13% of SCA patients, compared to 2% of age, sex, and education matched African American controls[102]. However, this study used a more restrictive definition of SBI, requiring lesions to exceed 5 mm in diameter, and to be visualized by both T1 and T2 weighted MRI images[101]. To date, no publications have examined heterozygous SCT as a genetic risk factor for subclinical brain infarctions.

4.6 Burden Associated with Subclinical Brain Infarctions

Though asymptomatic, SBI are associated with subtle neurological deficits and future cognitive decline[103]. In the Cardiovascular Health Study, baseline MRI exams were performed on 1433 participants, with follow-up MRI performed 5 years later. Participants developing incident SBI scored worse on the Modified Mini-Mental State Examination and the Digit-Symbol Substitution test, both validated measures of cognitive performance[103]. These results are consistent with the Rotterdam Scan Study, which prospectively imaged 1015 stroke-free, elderly participants by MRI. Participants with SBI were found to have a steeper decline in cognitive function, and more than twice the risk of developing dementia or Alzheimer's disease[104]. Longitudinal studies have also reported an association between SBI and future stroke. Over an average follow up of 4 years, the Cardiovascular Health Study

reported a hazard ratio for stroke of 1.5 (95% CI: 1.1 – 2.1) in participants with SBI, compared to those without[105]. Stroke risk also appeared to be related to the number of infarctions; participants with multiple SBI were found to have twice the risk of subsequent stroke[105].

4.7 Summary of Subclinical Brain Infarctions

Infarctions of the small, deep penetrating arteries are often asymptomatic, but are associated with future stroke and cognitive decline. The prevalence of SBI increases with age, and may be higher in African Americans. While the etiology of SBI appears to be heterogeneous, thrombus and hypercoagulability likely play a role. Sickle cell anemia is a known risk factor for SBI. While SBI is the most common neurological manifestation affecting children with SCA, its prevalence in adults with SCA is not well defined. Furthermore, no published studies have examined the associations between SBI and the heterozygous form of sickle cell anemia, SCT.

5. WHITE MATTER LESIONS

5.1 Overview

The central nervous system is divided into white and gray matter (Figure 12). White matter owes its whitish color to myelin sheaths, lipid and protein coverings surrounding the neuron axons[2]. The function of myelin sheaths is to insulate the axons and increase the speed of action potentials, or nerve impulses[2]. In the central nervous system, myelin sheaths are produced by oligodendrocytes, a type of glial cell[2]. Leukoaraiosis, also called white matter lesions (WML), are diffuse areas of the cerebral white matter characterized by loss of oligodendrocytes, axon demyelination, formation of vacuoles within and adjacent to the

neurons, and spongiosis, or intercellular edema[106]. In epidemiological studies, WML are primary detected by MRI, using proton density imaging. The severity of WML is typically estimated by the extent of WML, graded either qualitatively by comparison to reference images, or volumetrically. Intra-reader reproducibility of qualitative scores range from 71% - 94.5%, while inter-reader agreement ranges from 68% - 92%[67, 72]. The validity of MRI-detected WML is generally not ascertained, as this would require post-mortem examination.

5.2 Etiology of White Matter Lesions

Hypoperperfusion and ischemia are thought to cause oligodendrocyte cell death and resulting WML[107]. The deep white matter is particularly vulnerable to ischemia and hypoperfusion. Its network of small penetrating arteries and arterioles has a poor anastomatic system of collateral vessels, supplied solely by the leptomeningeal arteries[107, 108]. In animal experiments, cerebral ischemia induced in gerbils[109] and rats[110] has been shown to result in WML. While animal studies are not always generalizable to humans, human brain autopsies have confirmed a 20% decreased concentration of arterioles and afferent capillaries within white matter lesions, compared to healthy white matter tissue[106]. Loss of microvasculature may result from chronic ischemia inducing endothelial cell apoptosis. However, to date, autopsies have not identified the source of cerebral ischemia putatively causing WML[107].

5.3 Prevalence of White Matter Lesions

In elderly populations WML are particularly prevalent, with a greater severity observed in African Americans. During the ARIC study baseline MRI exam, WML were detected in 86% of individuals aged 55 – 72. This is consistent with baseline estimates from

the Cardiovascular Health Study, which reported a WML prevalence of 87% in participants aged 65-95. In the ARIC study, the prevalence of WML was found to be slightly lower in African Americans than European Americans (81% vs. 91%); however the severity of WML in African Americans was greater. After adjusting for age, the prevalence of severe WML in African Americans was 14%, higher than the prevalence observed in European Americans (11%). When examined longitudinally over a 10 year interval, African Americans were also more likely to experience worsening of WML severity by at least one grade, from the baseline MRI exam until follow up. In the ARIC study, WML worsening was detected in 70% of African Americans, compared to 54% of European Americans.

5.4 Risk Factors of White Matter Lesions

Vascular risk factors appear to be strongly associated with leukoaraiosis. In the ARIC baseline MRI, hypertension was noted in 65% of participants with severe WML, compared to 43% of participants without WML[111]. Participants with severe WML had a mean systolic blood pressure of 136 mmHg, significantly higher than 124 mmHg, the mean systolic blood pressure observed in participants without WML (p for trend < .0001) [111]. When examined longitudinally over a 10 year interval, a 1-standard deviation increase in mean arterial blood pressure only marginally influenced the odds ratio of increased WML severity grade[113]. However, when stratified by race, midlife systolic blood pressure was a significant predictor of WML progression in African Americans[67] . Cross-sectionally, smoking status has also been implicated as a risk factor for WML severity. At the baseline ARIC MRI exam, the prevalence of moderate or severe WML among current smokers was 43%, compared to 33% in never-smokers (p for trend <.001) [111]. Although it has been examined, diabetes mellitus has not been shown to be associated with WML in the ARIC

study, either cross-sectionally or longitudinally; however a 1-standard deviation increase in fasting glucose was found to significantly increase the odds ratio of WML severity progression[113]. Thus, vascular risk factors, and possibly diabetes, may be related to the etiology of WML.

5.5 Sickle Mutation and White Matter Lesions

In cross-sectional studies, WML prevalence has been associated with both the homozygous and heterozygous sickle mutation, which may be a genetic risk factor for leukoaraiosis. White matter lesions were detected in 15% of patients with sickle cell anemia, compared to 7% of age, sex, and education matched African American controls with normal hemoglobin, in a cross sectional study with a mean study population age of 32 years[102]. Although only one study has been conducted, a greater prevalence of WML has also been observed in SCT carriers. In a small case-control study with 26 participants, children with SCT and sibling controls with normal hemoglobin were prospectively screened by brain MRI. White matter lesions were detected in 10% of participants with SCT, but in none of the sibling controls[51]. The prevalence and severity of WML in adults with SCT is unknown.

5.6 Burden Associated with White Matter Lesions

As with subclinical brain infarctions, WML are generally asymptomatic; however, severe WML and progression of WML severity have been shown to predict adverse neurocognitive outcomes[114-117]. The Leukoaraiosis and Disability Study (LADIS), a European multicenter collaboration including 639 patients aged 65-84 with mild cognitive complaints, found that severe WML at baseline independently predicted disability, dementia, and cognitive impairment[114]. After 3 years of follow up, progression of WML severity

correlated with a decrease in executive function, assessed by the Vascular Dementia Assessment Scale (VADAS) and Stroop test[115]. However, the LADIS study may not represent the general population of elderly adults, as mild cognitive impairment was an inclusion factor for participants. The Cardiovascular Health Study examined older adults using a less restrictive study population. As with the LADIS study, CHS reported adverse outcomes associated WML progression, which was detected by 2 MRI exams performed 5 years apart. A progression in WML severity was associated with decreased cognitive function, assessed by the Mini-Mental State exam and Digit-Symbol Substitution test[116]. WML progression was also found to predict future stroke in the CHS study. After 9 years of follow up from the second MRI exam, WML severity that worsened by at least one grade increased the hazard ratio of stroke by 1.39 (1.02 – 1.88)[117]. The observations from the LADIS and CHS studies suggest that WML, though often subclinical in presentation, are not benign.

5.7 Summary of White Matter Lesions

White matter lesions are very prevalent in elderly populations, and are associated with cognitive decline and stroke. A greater severity, and an accelerated progression of WML severity, has been detected in African Americans. The etiology of WML appears to be related to hypoperfusion, but is not well established. Both sickle cell anemia and SCT have been associated with a greater prevalence of WML, when compared to matched controls with normal hemoglobin. However, it is unknown if older adults with SCT have a greater severity of leukoaraiosis than African Americans with normal hemoglobin.

6. SICKLE CELL TRAIT AND CEREBROVASCULAR DISEASE

6.1 Overview

Sickle cell anemia is a known risk factor for stroke; however, the association between SCT and cerebrovascular disease is less well-established. The published literature entails a total of 17 case reports, 3 retrospective studies, and a single, prospective study examining SCT and stroke.

6.2 Case Reports

Numerous case reports have described stroke in young patients with SCT, with no predisposing cerebrovascular risk factors[118-129]. The majority of these cases were not preceded by known episodes of hypoxia; however, two describe post-operative stroke in children sedated by general anesthesia[121, 122]. Necropsy examinations identified sagittal sinus thrombosis in both, with perioperative hypoxia assumed to trigger erythrocyte sickling and clot formation[118-120].

Several criticisms of the associations between SCT and stroke have been made. Because many of the published case reports antedate modern medical practice, it has been suggested that the evaluations of stroke risk factors were incomplete[130]. A case report describing stroke in a 32 year old with SCT was criticized, due to the presence of lymphocytes in the cerebral spinal fluid, which may be suggestive of meningoencephalitis as the stroke cause, rather than SCT[131]. However, the post-mortem evaluation revealed bilateral, old infarctions in the deep white matter, in addition to new infarcts, which reasonably may have preceded the infection[126]. Another case report, of a 22 year old male with SCT, was criticized due to a history of rheumatic fever and mitral valve disease[131].

However, echocardiography only detected mild mitral valve prolapse and regurgitation, and no mention was made of heart failure or left atrial dilation, more widely accepted cardiac risk factors for stroke[123]. With or without criticisms, case reports are inherently subject to publication bias. They capture attention by describing interesting, medical anomalies, but may not accurately reflect population-based associations.

6.3 Retrospective Studies

Three retrospective studies analyzing sickle cell trait and hospital discharge for stroke have been published. The first, conducted from 1965-1969 at North Carolina Memorial Hospital, included 227 patients with SCT and 16,701 African Americans assumed to have normal hemoglobin. No differences were found in frequency of hospital discharges for stroke; however, this study is deeply flawed. The African American controls were never confirmed to have normal hemoglobin, which may have resulted in substantial misclassification bias. Furthermore, there was no statistical control for age. While the mean age of SCT patients hospitalized with stroke was 38 years, the control group was much older. The mean age of the controls is not published; however, 88% were over the age of 30, and 37% were over the age of 65. As Steen, *et al* point out, if the control group is assumed to have a prevalence of SCT that is 7.8%, then 1,303 members of this group were unidentified SCT carriers[51]. A lifetime stroke incidence among SCT carriers would therefore only need to be 8.2% higher than in African Americans with normal hemoglobin, to explain the 107 excess strokes among the group of African American control group[51].

The second study, conducted more recently in the French Carribean colony of Guadaloupe, analyzed the prevalence of SCT in 295 hospitalizations for stroke. A 10-fold

higher risk for hemorrhagic stroke and a 15-fold lower risk for ischemic stroke were observed in patients with SCT, compared to Guadeloupeans with normal hemoglobin[132]. This study has been criticized for diagnosing stroke type by computed tomography, which may not distinguish between primary hemorrhages, and hemorrhagic bleeding secondary to infarctions[133]. It is also uncertain that the population of Guadeloupe can be generalized to African Americans, due to differences in genetic ethnography. The Guadeloupean population is primarily multiracial creole, an admixture of European, African, Indian, and Amerindian ancestries.

Finally, a recently conducted analysis based on 13,964 African Americans (2,642 with SCT and 139 with SCA) registered with the Kaiser Permanente Northern California health system reported no differences in stroke diagnoses for patients with either SCT, SCA, or normal hemoglobin[134]. While hemoglobin was ascertained by laboratory analysis, the mean age of the study population was 35. In adult populations, only pregnant African American women are routinely tested for sickle hemoglobinopathies, and if positive, the fathers are tested as well[134]. Due to the young age of the study population and low number of ischemic stroke events, the authors concluded the analysis was underpowered to detect statistically significant differences in stroke prevalence among patients with SCT, SCA, or normal hemoglobin[134].

6.4 Prospective Studies

The association between SCT and stroke has been examined in a single, prospective epidemiological study[135]. In this analysis from the ARIC study, 3,497 genotyped African Americans were followed a median of 22 years. The frequency of ischemic stroke was

slightly higher in those with SCT (13%) than participants with homozygous hemoglobin A (10%). In multivariable analyses controlling for the traditional risk factors of stroke, SCT was associated with a 40% higher risk of stroke (OR= 1.4, 95% CI: 1.0 – 2.0), and an incidence rate difference amounting to nearly 2 extra strokes per 1000 person years (IRD = 1.9, 95% CI: 0.3 – 3.8). However, this analysis was based on a single cohort study, and warrants replication in other populations.

7. PUBLIC HEALTH IMPLICATIONS

African Americans are disproportionately burdened by cerebrovascular disease. The disparities in stroke prevalence, incidence, and associated disability and mortality underscore a dire need for effective stroke prevention strategies. Genetic risk factors may present an opportunity to target high risk individuals for prevention. Approximately 41 million individuals living in the United States self-identify as black [136]. With a heterozygous allelic frequency of 8%, sickle cell trait is estimated to affect 3.3 million of these individuals. In 1978, the National Institute of Health recommended mandatory newborn screening for sickle cell hemoglobinopathies[137], and by 2006, all 50 states initiated this practice universally[138]. As a result, SCT is easily identifiable, and documented in the medical records of younger generation African Americans. Although somewhat controversial, student athletes are also screened for SCT, as mandated by the National Collegiate Athletic Association[139]. Three branches of the Department of Defense, the US Air Force, US Marines, and US Navy, screen for SCT as well[140].

Environmental exposures and behaviors are known to influence the phenotypic expression of SCT. Heavy exertion, dehydration, and high altitude are widely accepted to

invoke hemoglobin S polymerization and erythrocyte sickling in SCT carriers. Though the associations between SCT and stroke have been conflicting in retrospective studies, SCT has been associated with stroke in case reports, and a single, prospective study. A higher prevalence of subclinical small vessel disease has also been detected in SCT carriers, in a small, pediatric study. If SCT is indeed associated with stroke and cerebrovascular disease, modification of the phenotypic expression, either by behavior, environment, or pharmacotherapy, may be tenable for the prevention or treatment of stroke in over 3 million African American SCT carriers.

CHAPTER 2: STUDY AIMS AND RATIONALE

INTRODUCTION

Numerous case reports have described stroke in young patients with SCT, with no underlying risk factors for cerebrovascular disease. Additionally, SCT has been associated with stroke in a single, prospective epidemiological study. The pathophysiology of stroke in SCT carriers is unclear; however, and may be related to cerebral small vessel disease. We hypothesize that sickle cell trait is associated with an increased risk of cerebral small vessel disease. The hypercoagulability and vasculopathy observed with SCT lend biological plausibility to this hypothesis, as both conditions contribute to cerebrovascular disease. Previous research examining SCT and small vessel disease is scarce; however, an increased prevalence of WML have been noted in SCT carriers, in a small study of pediatric patients. We intend to conduct the first population-based, epidemiological investigation of sickle cell trait and cerebral small vessel disease. To accomplish this, we will analyze data from the Atherosclerosis Risk in Communities Study (ARIC), which examined a subset of participants by cerebral MRI in 1993-1995 and 2004-2006. Genetic modification by stroke candidate genes, sickle haplotypes, and alpha thalassemia will be considered as well.

PRIMARY AIM: Analyze associations between sickle cell trait and the prevalence, severity, and progression of white matter lesions in African Americans.

We hypothesize that African American participants with SCT will have a greater prevalence and severity of WML at the baseline MRI (1993-1995), and an increased rate in the progression of WML between MRI 1 (1993-1995) and MRI 2 (2004-2006).

Sub-aim: Consider potential modification of the SCT and WML associations by environmental covariates (traditional risk factors for stroke), and genetic covariates (sickle haplotypes, alpha thalassemia, and stroke risk candidate genes).

SECONDARY AIM : Quantify associations between sickle cell trait and the prevalence and incidence of subclinical brain infarctions in African Americans.

We hypothesize that African American participants with SCT will have a greater number of SBI lesions detected at the first MRI (1993-1995), and will develop a greater number of SBI lesions by the second MRI screening (2004-2006).

Sub-aim: Consider potential modification of the SCT and SBI associations by environmental covariates (traditional risk factors for stroke), and genetic covariates (sickle haplotypes, alpha thalassemia, and stroke risk candidate genes).

CHAPTER 3: METHODS

1. INTRODUCTION

We investigated the associations between sickle cell trait and cerebrovascular disease in African Americans participating in the Atherosclerosis Risk in Communities Study. The ARIC study includes cerebral MRI screenings, extensive genomic characterization, and phenotypic data meticulously collected by quality assurance protocols. The following chapter will describe the ARIC study design, sickle cell trait genotyping, imputation of potential modifier genes, stroke and TIA diagnoses, magnetic resonance imaging for subclinical brain infarction and white matter lesion detection, and clinical covariate data collection. Finally, this chapter will conclude with a detailed description of the statistical analyses and power calculations.

2. STUDY DESIGN

The ARIC study is an ongoing, prospective epidemiological cohort study. A biracial, population-based sample (N= 15,792) of African Americans and European Americans aged 45-65 were recruited at the study onset in 1987-1989[141]. Participants were recruited from Jackson, Mississippi; Forsyth County, North Carolina; Washington County, Maryland; and Minneapolis, Minnesota. African Americans (N=4,270) were selected from the Jackson and Forsyth sites, representing urban, suburban, and rural areas. Households in Forsyth County were identified for study inclusion by area sampling, while driver's licenses listings were

used to identify age-eligible residents in Jackson[142]. The cohort participation rate among African Americans was 46%; but participants did not differ from nonresponders in general health status or hospitalizations[141]. The ARIC study includes 5 cohort examinations (visit 1: 1987-1989; visit 2: 1990-1992; visit 3: 1993-1995; visit 4: 1996-1998; and visit 5: 2011-2013), annual telephone surveys during interim years, and ongoing surveillance of hospitalized events. Brain MRI was offered to a random subset of participants in 1993-1995, with a follow up exam offered in 2004-2006. There were 926 African Americans participating in the first MRI exam, with 470 (56%) returning for the second MRI.

3. SICKLE CELL TRAIT GENOTYPING

SCT was genotyped in consenting African Americans using a TaqMan® SNP Genotyping Assay (Applied Biosystems; Foster City, CA). TaqMan® high throughput assays genotype polymorphisms of interest using polymerase chain reactions and dye-labeled probes[143]. First, the DNA is denatured by heating, allowing the double stranded helix to separate. Once denatured, a primer sequence of the DNA binds to catalytic enzymes, or polymerases, which initiate replication. Genotypes are revealed by dye-labeled probes, with DNA sequences complementary to the polymorphism of interest. The 5' ends of the probes contain reporter dyes, while the 3' ends are bound to quenchers. The quenchers absorb the fluorescent energy of the reporter dyes, as long as the probe is intact and bound to the DNA. Probes are released from the DNA by the polymerase *Taq*, as it synthesizes new strands. Fluorescence is then detectable by laser excitation. The fluorescent signals are plotted and measured with sequence detection software, allowing allelic discrimination[143].

Assay reproducibility is validated by genotyping blinded duplicates, or by assaying positive controls, DNA samples with known genotypes determined by a different method[143]. The ARIC central laboratory performs blind replicate redundancy tests in randomly selected samples, representing 5% of the total assays. If discordant genotypes are detected in more than 2%, SNPs are excluded from further analysis. In addition to these quality control measures, we will assess genotyping errors by Hardy-Weinberg equilibrium (HWE) testing [144]. In large, randomly mating populations without selection, genetic drift, genetic mutation, or gene flow, allelic frequencies will remain constant from one generation to the next[145]. While no population strictly meets Hardy Weinberg assumptions, wide deviations from the expected genotype frequencies often indicate genotyping errors. However, genes with a minor allele frequency less than 5% may not follow Hardy Weinberg proportions[144], a potential consideration in our analysis.

4. GENETIC MODIFIERS

Potential genetic modifiers of the association between SCT and cerebrovascular disease were genotyped or imputed, by a variety of methods. Hemoglobin C was genotyped by Taqman, as previously described. The Affymetrix 6.0 GWAS Array, which simultaneously types 906,600 variants in the coding and noncoding regions of the DNA[146], was used to genotype common SNPs. Rare variants were genotyped by the Affymetrix Axiom Exome Chip Array, simultaneously typing 300,000 coding SNPs within the exons. The GWAS and exome chip arrays function similarly, but target different SNPs. Arrays are divided into multiple wells, or features, each containing a specific probe complementary to a particular SNP. Prior to genotyping, the sample DNA is digested by restriction enzymes (Nsp I and Sty I), and DNA fragments are ligated to adaptors[146]. DNA primers recognize the

adapter sequences, and initialize amplification by polymerase chain reactions. Once the DNA is amplified, it is fragmented and labeled with biotin[146]. The biotin-labeled DNA is then washed over the array for 14-16 hours. Available SNPs bind to the complementary probes, and are identified by laser fluorescence.

Genetic markers typed by the GWAS chip were used to impute additional SNPs, based on the 1000 Genomes reference panel. The 1000 Genomes Project sequenced the entire genome of 2500 individuals of Chinese, Japanese, Indian, African, European, and Latin American ancestries. The initial phase included low coverage (~4x) whole genome sequencing, followed by moderate coverage (~20x) whole exome sequencing, and finally high coverage (~40x) sequencing in 500 individuals, as technology advanced and costs decreased[147]. Genotype data in the 1000 Genomes Project is much denser than commercially available GWAS or Exome chips, and can be used as a reference for Monte Carlo Markov models, which impute untyped SNPs, based on known linkage disequilibrium patterns. However, imputed data is highly dependent on the sample from which it is imputed, and ethnic match must be considered[148]. In the ARIC Study, SNP imputations in African Americans have been shown to have better accuracy with the 1000 Genomes reference panels, compared to the HapMap III panels, which were based on a sample with less ethnic diversity[149].

GWAS data were subjected to rigorous quality control. Duplicate samples were identified by PLINK software, and removed from analysis. Individuals with >5% missing genotype calls out of the total number of SNPs typed were also removed, as this may indicate a genotyping problem that affects all SNPs for this individual[148]. As a simple test to rule out sample mix-ups, the concordance between self-reported gender and genotyped sex were also

assessed, using PLINK, and discordant individuals were removed. First degree relatives were also identified by PLINK, defined by alleles that were identical by state and had a distance value of ≥ 0.8. Race-specific imputation quality was ascertained by the squared Pearson correlation (r^2) between the imputed genotype calls and the true genotype[150]. An r^2 value of .90 or greater was considered sufficient for inclusion in the analysis. Quality scores for hemoglobinopathies, stroke modifier SNPs, and sickle haplotype proxies are shown in Tables 2, 3, 4, and 5.

Table 2: Imputation quality scores for hemoglobinopathy modifier genes

Gene	Gene Name	Effect	SNPs	MAF	Affy 6.0 r^2
HBB	Hemoglobin, beta	Beta thalassemia Beta thalassemia Senegal Haplotype	-88 (C-T) -29 (A-G) rs7482144		----- ----- -----
HBA1	Hemoglobin, alpha 2	Alpha thalassemia	rs63751476		-----
HBS1L-MYB	Intergenic region	Increase gamma globin	rs9399137	0.753	0.991
BCL11A	B-cell lymphoma	Increase gamma globin	rs766432	0.074	0.986

The majority of hemoglobinopathy modifiers were not imputable, using the Affymetrix 6.0 microarray and 1000 Genomes reference panel. However, SNPs for the *HBSIL-MYB* and *BCL11A* genes, which are external to the *HBB* gene-like cluster and influence gamma globin production (and fetal hemoglobin concentrations), imputed with high quality.

Table 3. Imputation quality scores for SCA stroke risk candidate genes

Gene	Gene Name	SNPs	MAF	Affy 6.0 r2
ADCY9	Adenylate cyclase 9	rs2072338 rs437115 rs2238432 rs2238426 rs2283497	.167 .475 .261 .149 .457	0.825 0.963 0.928 0.926 0.779
ANXA2	Annexin a2	rs11853426		-----
BMP6	Bone morphogenic protein 6	rs267196 rs408505 rs267201 rs449853	.299 .663 .345 .411	1.0 1.0 0.994 0.987
CCL2	Chemokine ligand 2	rs4586		-----
CSF2	Colony stimulating factor 2	rs25882	.31	0.976
ECE1	Endothelin converting enzyme 1	rs212528 rs212531	.108 .098	1.0 0.991
ERG	V-ets virus e26 oncogene like	rs989554	.722	0.978
MET	Met proto-oncogene	rs38850 rs38859	.053 .09	0.975 0.997
SELP	Selectin P	rs2420378 rs3917733 rs3753306	.114 .075 .53	1.0 0.996 0.944
TEK	Tek tyrosine kinase	rs489347	.68	1.0
TGFBR3	Transforming bone growth factor β-receptor III	rs2148322 rs2765888 rs2007686 rs284875	 .118 .149 .903	----- 0.991 0.993 0.99
TNF-α	Tumor necrosis factor	rs1800629	.135	0.99

With the exception of *ANXA2* and *CCL2*, all SCA stroke risk candidate genes were imputable to the 1000 genomes reference panel, and most were high quality imputations. Stroke risk SNPs with genome-wide significance in large consortia of Europeans were also examined, and shown to impute very well.

Table 4. Imputation quality scores for stroke risk candidate genes with genome-wide significance

Gene	**Gene Name**	**SNPs**	**MAF**	**Affy 6.0 r^2**
HDAC9	Histone deacetylase 9	rs2107595 rs11984041	0.219 0.224	0.978 0.974
PITX2	Paired-like homeodomain 2	rs6843082	0.691	1.0
NINJ2	Ninjurin 2	rs12425791	0.107	1.0

When SNPs could not be imputed by the Affymetrix 6.0 GWAS chip and 1000 Genomes reference panel, the exome chip genotypes were searched. However, none of the missing SNPs were identified by the exome chip. After concluding that SNP data were not available, proxies for SNPs were identified using PLINK. Proxy SNPs were identified by linkage disequilibrium, using publically available data from the 1000 Genomes project. SNP proxies with a correlation of .80 or greater were considered acceptable. Imputation quality scores for Senegal haplotype proxies were assessed, using the GWAS chip and 1000 Genomes reference panel.

Table 5: Imputation quality scores for Senegal haplotype proxies

Gene	Effect	SNP	Proxy SNPs	Proxy r^2	MAF	Affy 6.0 r^2
HBB	Senegal Haplotype	rs7482144	rs10128556	0.890995	0.155	0.971
			rs2071348	0.890995	0.155	0.971
			rs2855039	0.890995		
			rs2855121	0.890995	0.156	0.977
			rs145323148	0.827691	0.164	0.963
			11:5289228	0.827691	0.165	0.965
			rs72872549	0.827691	0.164	0.969

Several proxies for the Senegal haplotype were identified, but unfortunately, the alpha thalassemia deletions were not imputable, or in linkage disequilibrium with any proxy. However, whole genome sequencing data is expected to be available soon.

5. STROKE AND TIA DIAGNOSIS

Stroke prevalence was determined in the ARIC study at the baseline visit (1987-1989)[64]. Participants self-reported neurologic signs and symptoms of stroke using a validated questionnaire. The questionnaires were administered by certified interviewers, and quality assurance was ascertained by audiotape. Participants were asked if they had ever been told by a physician that they had a stroke or TIA, or had experienced sudden episodes of speech dysfunction, vision loss, double vision, weakness or paralysis, numbness or tingling, or dizziness and loss of balance. Based on the questionnaire responses, a computer algorithm diagnosed stroke and TIA, and determined the vascular distribution involved. Stroke symptoms resolving in less than 24 hours were considered TIA. The standardized questionnaire and computer algorithm used by the ARIC study were previously validated in the Asymptomatic Carotid Atherosclerosis Study (ACAS). When compared to a standardized neurological exam by a neurologist, diagnosis agreement with the computer

algorithm was 80.1%. Overall, the computer algorithm classified stroke and TIA with a sensitivity of 87.8% and a specificity of 71.9%[151].

Stroke incidence was determined in the ARIC study by annual telephone contact with the participants, surveillance of hospitalized events, and death records[60]. Discharge lists for hospitalized study participants with ICD-9 diagnosis codes 430-438 were provided to the ARIC study. When keywords indicative of stroke were present in the discharge summaries, the medical records were abstracted[60]. Stroke diagnosis was determined by computer algorithm and physician review of the medical record, which included the discharge summary, imaging reports, neurological consults, and medical history. Discrepancies between diagnoses made by the physician reviewer and computer algorithm were settled by a second physician adjudicator. Agreement rates between the physician reviewer and computer algorithm were 78%[60]. In the majority of discordant diagnoses (65%), the physician adjudicator agreed with the physician reviewer, rather than the computer algorithm[60]. Based on the clinical picture, stroke was classified as either definite or probable, and categorized as thrombotic, cardioembolic, lacunar, subarachnoid hemorrhage, or intracerebral hemorrhage (Table 8). Stroke symptoms resolving in less than 24 hours were considered TIA. Stroke was considered "possible", rather than "definite" or "probable" when there was a clinical presentation of at least one major symptom (hemiparesis, homonymous hemianopia, or aphasia), or at least two minor symptoms (diplopia, vertigo, dysarthria, dysphagia, or dysphonia, unilateral numbness, severe headache, depressed consciousness, meningeal irritation, retinal hemorrhage, or palsy of the iii cranial nerve), but the clinical history, imaging reports, or autopsy were inconclusive. Neurological symptoms attributable to head trauma, neoplasm, coma (due to diabetes, epilepsy, poisoning / drug

overdose, hypovolemia, hypoglycemia, uremia, or liver disease), cerebral vasculitis (due to systemic lupus erythematosus, or radiation), peripheral neuropathy, bleeding disorders (disseminated intravascular coagulopathy or thrombocytopenia), or central nervous system infection involving the brain or meninges were considered ineligible for stroke diagnosis.

Table 6: Definite and Probable Stroke subtype classification

Stroke Subtype	**Definite Diagnosis**	**Probable Diagnosis**
Subarachnoid Hemorrhage	1. Identification by Angiogram + Bloody spinal fluid 2. Identification by CT or MRI 3. Identification by surgery 4. Identification by autopsy	1. Identification by Angiogram * Spinal tap not done 2. Symptoms (1 or more): Severe headache Depressed consciousness Meningeal irritation Retinal hemorrhage + Bloody spinal fluid
Brain Hemorrhage	1. Identification by CT or MRI 2. Identification by surgery 3. Identification by autopsy 4. Major symptom (1) Hemiparesis Homonymous hemianopia Aphasia + Bloody spinal fluid 5. Minor symptoms (2) Diplopia Vertigo / Gait disturbance Dysarthria, dysphagia, or Dysphonia Unilateral numbness + Bloody spinal fluid	1. Major symptoms (1) Hemiparesis Homonymous hemianopia Aphasia 2. Minor symptoms (2) Diplopia Vertigo / Gait disturbance Dysarthria, dysphagia, or Dysphonia Unilateral numbness 3. Coma persisting 24 hours + Bloody spinal fluid * CT / MRI not done or inadequate
Thrombotic Brain Infarction	1. Identification by autopsy 2. Major symptom (1) Hemiparesis Homonymous hemianopia Aphasia + MRI / CT identification 3. Minor symptoms (2) Diplopia Vertigo / Gait disturbance Dysarthria, dysphagia, or Dysphonia Unilateral numbness + MRI / CT identification	1. Major symptom (1) Hemiparesis Homonymous hemianopia Aphasia 2. Minor symptoms (2) Diplopia Vertigo / Gait disturbance Dysarthria, dysphagia, Dysphonia Unilateral numbness * CT / MRI not done or Inadequate * Spinal tap not done

Embolic Brain Infarction (Non-carotid)	1. Identification by autopsy Brain infarction + Emboli in any organ 2. Major symptom (1) Hemiparesis Homonymous hemianopia Aphasia + Valvular heart disease; Atrial fibrillation, MI, Cardiac surgery, myxoma + Identification by MRI / CT 3. Minor symptoms (2) Diplopia Vertigo / Gait disturbance Dysarthria, dysphagia, or Dysphonia Unilateral numbness + Valvular heart disease, Atrial fibrillation, MI, Cardiac surgery, myxoma + Identification by MRI / CT	1. Major symptom (1) Hemiparesis Homonymous hemianopia Aphasia + Valvular heart disease, Atrial fibrillation, MI, Cardiac surgery, myxoma 2. Minor symptoms (2) Diplopia Vertigo / Gait disturbance Dysarthria, dysphagia, or Dysphonia Unilateral numbness + Valvular heart disease, Atrial fibrillation, MI, Cardiac surgery, myxoma * Inconclusive MRI / CT * Spinal tap not done

6. MAGNETIC RESONANCE IMAGING PROTOCOL

Brain MRI was offered twice in the ARIC study: in 1993-1995 and 2004-2006. All exams were performed by certified radiographers following standardized protocols, and read by board-certified radiologists with subspecialty training in neuroradiology. Quality assurance was ascertained by intra- and inter-reader agreement among randomly replicated images. Scanner performance and resolution were monitored by phantom images submitted bi-monthly. Cerebral images were acquired using the anterior commissure / posterior commissure (ACPC) line as an anatomic reference. First, the ACPC line was located in the sagittal view (sectioning the brain into left and right). Axial views (sectioning from top to bottom) were then acquired by aligning images parallel to the ACPC. Contiguous axial images were captured in 5 mm slices, from the vertex of the skull to the base, using General Electric 1.5 Tesla scanners. Proton density (sometimes called spin density) weighting was

achieved by setting the repetition time to 3000 ms and the echo time to 30 ms. T2 weighted images were acquired with a 3000 ms repetition time and a 100 ms echo time, and T1 weighted images used a 500 ms repetition time and a 20 ms echo time. Imaging protocols and equipment remained consistent between MRI 1 and MRI 2 with the exception of axial FLAIR, added to MRI 2 after software upgrades.

7. SUBCLINICAL BRAIN INFARCTIONS

Lacunes were identified in the cortical gray matter and deep nuclear regions by signal hyperintensities using T1 and T2 weighted MRI, and in the white matter and brain stem by hypointensities in T1-weighted images. Focal, non-mass lesions measuring 3 – 20 mm within territories of vascular distributions were considered SBI when asymptomatic, in individuals without prevalent or incident stroke or TIA prior to cerebral MRI. SBI diameters were measured by electronic calipers, using the largest dimension within the contiguous MRI slices. Inter- and intrareader agreement rates among ARIC neuroradiologists were 79% and 82%, respectively[71]. Periventricular spaces, which fill with cerebral spinal fluid, are often indistinguishable from lacunes when imaged by MRI without FLAIR. To avoid false positive identification of SBI, lesions less than 3 mm in diameter were considered perivascular spaces. Although signals from cerebral spinal fluid were filtered by FLAIR in the second MRI exam, lesions less than 3 mm remained classified as perivascular spaces, for consistency of definitions.

8. WHITE MATTER LESIONS

Leukoaraiosis was defined by diffuse areas of white matter hyperintensities, detected by proton density imaging. For the purposes of our analysis, WML were considered

subclinical, when identified in individuals without prevalent or incident stroke or TIA prior to MRI examination. Severity was graded qualitatively by neuroradiologists, by matching lesion patterns to reference images[112]. The reference images ranged from a few "dots" of hyperintensities within subcortical white matter, to continuous periventricular rims of hyperintense regions, to confluent periventricular regions involving the entire white matter[72]. Inter- and intra-reader reliability coefficients in the ARIC study were .68 and .71, respectively[72].

9. COVARIATE ASSESSMENT

The analysis of SCT and cerebrovascular disease accounted for ancestry, as well as confounding by the traditional risk factors of stroke: age, sex, smoking, hypertension, hyperlipidemia, diabetes, coronary heart disease, and atrial fibrillation. Clinical covariates were ascertained during the baseline ARIC exam (1987-1989), by home interviews, health questionnaires, and clinical examinations. Causal pathways between confounders, SCT expression, and cerebrovascular disease are shown in Figures 15 and 16.

Figure 4: Directed acyclic graph for causal pathway between SCT, hypercoagulability, and cerebrovascular disease

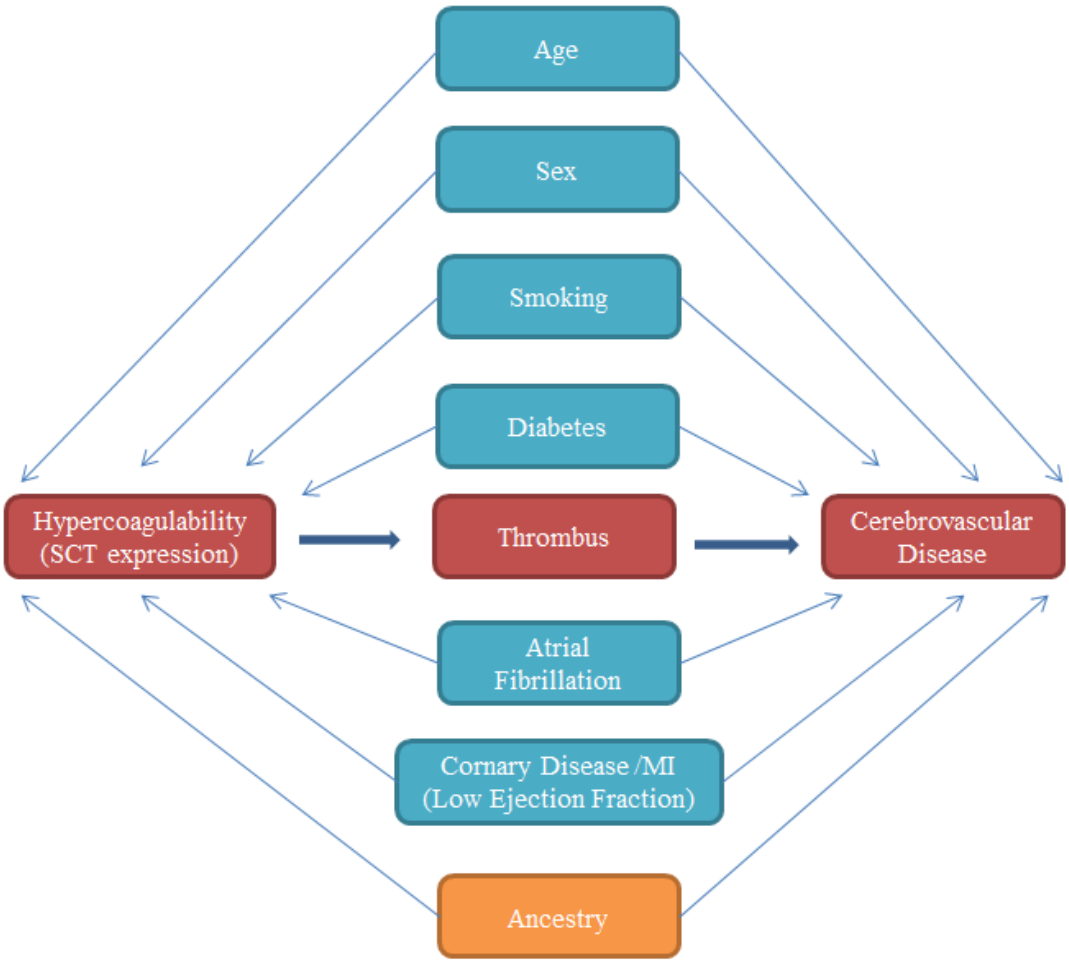
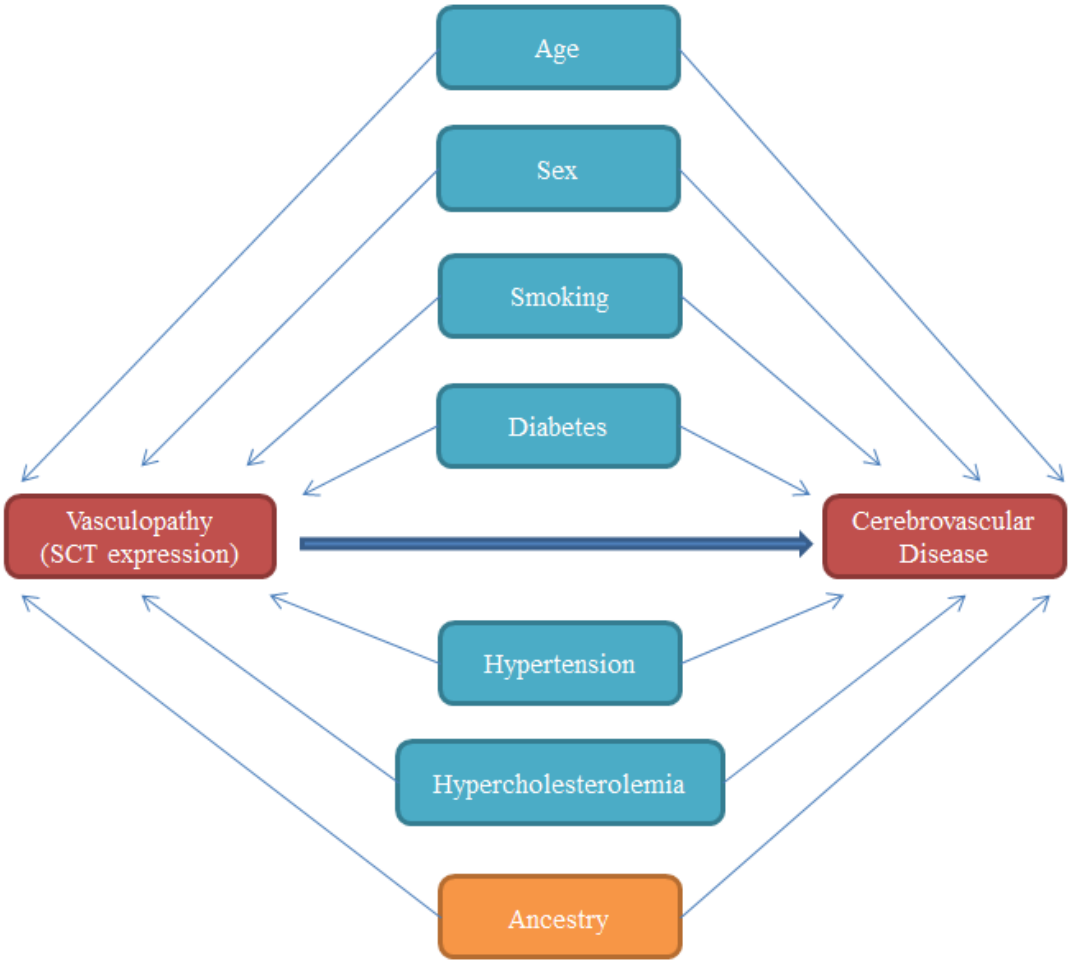


Figure 5: Directed acyclic graph for causal pathway between SCT, vasculopathy, and cerebrovascular disease



9.1 Ancestry

Hypothetically, if SCT is more prevalent among certain subpopulations of African Americans than others, and genetic factors putatively increasing stroke risk are also more prevalent among certain subpopulations, then SCT will appear to be associated with cerebrovascular disease, simply due to confounding by the population structure. To avoid false positive association, population stratification was accounted for with principal

components, previously derived by the ARIC Study using EIGENSTRAT software. Principal components are continuous axes of genetic variation, or eigenvectors of the covariance matrix between samples[152]. The first principal component is the eigenvector explaining most of the genetic variability. The second principal component is the eigenvector that is orthogonal to the first principal component[152]. Frequently, the first and second principal components correspond to geographic axes, such as latitude and longitude, which have influenced human genetic variation[148]. In the ARIC Study, 10 principal components capturing population stratification among African Americans have been identified. Individuals with values less than or greater than 8 standard deviations for any principal component were considered outliers, and removed from the analysis.

9.2 Smoking

Smoking habits were assessed by interview, recording the amount and duration of tobacco use. Smoking will be dichotomized into a yes/no variable, based on current smoking at the baseline visit. Cotinine and carbon monoxide, metabolites of smoking, are often used to validate self-reported smoking. While this was not ascertained in the ARIC study, a meta-analysis of 26 studies validating self-reported smoking (as a yes/no variable) with biomarkers yielded an overall sensitivity of 87.5% and a specificity of 89.9%[153]. Interviewer-administered questionnaires and observational studies had higher estimates of sensitivity and specificity than self-administered questionnaires and interventional studies[153]. Retrospective measures of smoking, such as pack-years, are sometimes validated by test-retest surveys[154]; however, comparisons with gold-standard biomarkers of smoking are not possible, making them potentially less reliable measures.

9.3 Blood Pressure

Seated blood pressures were measured by random-zero mercury manometers, after determining proper cuff sizes based on arm diameters. A total of 3 measurements were made, with only the second and third recorded for analysis. Hypertension was defined by the criteria established by the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood pressure. For our analysis, a systolic blood pressure exceeding 140 mmHg, diastolic blood pressure exceeding 90 mmHg, or use of antihypertensive medications was considered hypertension.

9.4 Cholesterol

Blood was drawn in EDTA tubes, and shipped to the ARIC Central Lipid Laboratory. Cholesterol was measured by an automated process, using an automatic pipetting station and a self-contained centrifugal analyzer with an optical monochromator. Hypercholesterolemia was considered a total cholesterol exceeding 240 mg/dL, based on the established guideline set by the third report of the National Cholesterol Education Program expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (Adult Treatment Panel III), or use of lipid lowering medicines. ARIC participants were instructed to fast 12 hours prior to their study visits, yet this was not always achieved. To avoid misclassification bias, blood labs from participants who were not fasting 12 hours will be excluded from the analysis

Laboratory quality control was ascertained by control samples with known concentrations of analytes, which were measured with each run. Within-day and overall precision were calculated by the standard deviation of the control sample's measured values.

Control sample standard deviations were normalized to the mean, to calculate the coefficient of variation.

$$\text{Coefficient of Variation (CV)} = \sigma / \mu$$

The coefficient of variation allows direct comparison of the precision of one analytical test to another, and is used to assess the reliability of repeated tests. While the ARIC protocols specify quality control and coefficient of variation calculations for lipid labs, these values are nowhere to be found on the website. Further inquiry is underway.

9.5 Diabetes

Blood glucose was measured by the ARIC Central Chemistry Laboratory. Quality control was ascertained by control samples, and the coefficient of variation was calculated to be 2.9%. For the purposes of our analysis, diabetes will be defined by a fasting plasma glucose level exceeding 126 mg/dL, the established guideline set by the American Diabetes Association[157], self-reported diabetes, or use of diabetic medications. While all ARIC participants were instructed to fast 12 hours prior to their study appointments, many were found to be noncompliant on the day of their visits. To avoid measurement bias, lab results from patients who did not fast 12 hours were excluded from the analysis.

9.6 Coronary Heart Disease

Prevalent coronary disease was defined by a history of myocardial infarction, coronary artery bypass graft, percutaneous coronary intervention, or electrocardiogram (ECG) suggestive of myocardial infarction. A standardized, 12-lead ECG and 2-minute rhythm strip were performed at the baseline visit, capturing leads I, II, III, aVR, aVL, aVF,

and the precordial leads V1 through V6. Digital ECG tracings were sent electronically to the ARIC ECG Computing Center for analysis, and assigned a Minnesota code[158]. ECGs with Minnesota codes indicative of myocardial infarction (ST elevation or depression, or inverted T-waves), were visually inspected and adjudicated by the Minneapolis ECG Reading Center. The agreement between myocardial infarction diagnoses made by Minnesota coding and ECG visual inspection is not readily available in ARIC, but has been assessed in a random sample of 768 patients, and reported to be 94.3%[159].

9.7 Atrial Fibrillation

Atrial fibrillation was classified by an ECG with a Minnesota code of 8.3[158]. Agreement between Minnesota coding and atrial fibrillation diagnosis by visual inspection of the ECG was not ascertained in the ARIC study, but has been previously reported. In a test set of 300 ECGs classified by Minnesota coding and visual inspection by an experienced cardiologist, arrhythmia diagnosis agreement was 90%[160]

10. STATISTICAL ANALYSIS

Associations between sickle cell trait and cerebrovascular disease will be assessed, after controlling for ancestry and confounding by the traditional risk factors for stroke (age, gender, hypertension, smoking, diabetes, hypercholesterolemia, coronary heart disease, and atrial fibrillation)[77]. The statistical plans for the WML and SBI aims are shown in Tables 10 and 11. Multiple statistical methods were considered for each analysis. After deliberating the pros and cons, a statistical approach was chosen. The statistical methods are described below, along with alternative approaches and the rationale against them.

Table 7: Statistical plan for white matter lesions aim

Aim 1 Outcomes	Cohort Exams	Approach	Estimate
White Matter Lesion Prevalence	MRI 1	Binomial Regression	Prevalence Ratio
White Matter Lesion Severity	MRI 1	Multiple linear regression	Mean WML severity grade
White Matter Lesion Severity Progression	MRI 1, MRI 2	Analysis of Covariance (ANCOVA)	Mean increase in WML severity grade

Table 8: Statistical plan for subclinical brain infarctions aim

Aim 2 Outcomes	Cohort Exams	Approach	Estimate
Subclinical Brain Infarction Prevalence	MRI 1	Multivariable Poisson regression	Prevalence Ratio
Subclinical Brain Infarction Incidence	MRI 2	Logistic Regression	Odds Ratio

10.1. WML Prevalence

The association between leukoaraiosis and SCT status will be assessed at the baseline MRI, by dichotomizing WML into a yes / no variable. Prevalence ratio comparisons will be estimated by multivariable binomial regression, adjusted for demographics, traditional stroke risk factors, and ancestral principal components:

$$\ln(P(X)) = \alpha + \sum [\beta(SCT) + \beta_i(x_i)]$$

In this model, the probability of prevalent WML is log transformed, which constrains the predicted probabilities to positive numbers. $\beta(SCT)$ is the increment to the log-prevalence for WML in participants with SCT, compared to those without, after controlling for covariates. The prevalence ratio is calculated by exponentiating $\beta(SCT)$. As a derivative of

10.1.a. Alternative Approach

As an alternative to prevalence ratios estimated by binomial regression, prevalence odds ratios may be estimated, using logistic regression. Based on previous analyses from the ARIC Study, WML are expected to be prevalent in approximately 80% of the study population. Odds ratios are useful for approximating risk of rare events (affecting <10% of the population), but overestimates risk when outcomes are common.

10.2. White Matter Lesion Severity

As shown in Table 9, WML severity scores are designed to represent sequentially increasing lesion volumes. The 10-point scale used to score WML severity will be assumed to represent equivalent increments from one grade to the next, and will be modeled as a continuous, ordinal variable.

$$y = \beta_0 + \beta(SCT) + \beta_i(x_i)$$

Where y is the predicted mean score, and the SCT regression coefficient (β) is the difference in the adjusted group mean WML score for participants with SCT, compared to those without. Linearity assumptions will be ascertained, with quadratic terms entered into the model in the event of non-linearity. The beta coefficient for the SCT genotype will be interpreted as the increase in mean WML score associated with SCT.

10.3. White Matter Lesion Progression

The change in WML severity score between the first and second MRI exams, will be assessed by analysis of covariance (ANCOVA):

$$Y_{iMRI2} = \beta_0 + \beta_1 Y_{iMRI1} + \ldots + \varepsilon_i$$

Where Y_{iMRI2} is the WML score for subject *i* at MRI 2 and Y_{iMRI1} is the score at MRI 1. By modeling the WML score at MRI 2 as the dependent variable, and the first WML score as a predictor variable, the baseline MRI score becomes an autoregression coefficient[161]. This overcomes the tendency for regression to the mean, frequently encountered by modeling change with simple linear regression:

$$\Delta y = \beta_0 + \beta(SCT) + \beta_i(x_i) + \ldots$$

With regression to the mean, individuals in the higher end of the distribution for WML scores at MRI 1 would be less likely to be in the higher end of the distribution of WML scores for MRI 2, compared to other study participants. As a result, the change in WML scores (Δy) would be strongly related to the baseline WML score, simply due to chance[161]. With ANCOVA, change is defined relative to the baseline WML score, to correct for regression to the mean[161]. Another approach to correct for regression to the mean is to analyze residual change, by comparing the observed values of WML scores at MRI 2 to the predicted WML scores at MRI 2[162]. However, the analysis of residual change (first described by Blomquist in 1977) is comparable to ANCOVA.

There are some limitations to ANCOVA. The dependent variable is assumed to be normally distributed, and in the event of wide deviation from normality, may require log transformation. ANCOVA also assumes linearity between continuous covariates and the dependent variable, and is difficult to interpret when there is interaction between the categorical exposure variable and the covariates, making the SCT-stratified slopes nonparallel. Linearity assumptions will be tested by scatter plots, and interaction tested by significant cross-product terms. The time interval between each MRI exam is also assumed

to be constant, at 10 years. However, it is possible that a 10 year interval will not adequately describe the time between examinations. In this case, the actual time variable will be entered into the model as a covariate.

10.3.a. Alternative Approach

As an alternative to ANCOVA, the change in WML severity may be modeled by logistic regression or Cox proportional hazards. This approach has been used on numerous occasions in the literature[113, 163]. WML progression is dichotomized as either an increase of 1 or more grade, or the absence of progression. This dichotomous outcome is then modeled as the dependent variable. It is not clear if this approach accounts for regression to the mean, however.

10.4. Subclinical Brain Infarction Prevalence

The number of subclinical brain infarctions in an individual will be treated as a count variable. Based on previous analyses from the baseline MRI examination in the ARIC study, the number of SBI per individual is expected to follow a Poisson distribution[68].

$$\Pr(y_i \mid x_i) = \exp(-\mu_i) * \mu_i^{y_i} / y_i!$$

Where y is the observed count, μ is the expected, or mean count, and x terms are the exposure and covariate variables[164]. $\Pr(y_i \mid x_i)$ is the conditional probability of the observed count, based on the covariate pattern. Baseline prevalence ratios for participants with SCT, compared to those with normal hemoglobin will be analyzed, using Poisson regression:

$$\log(Y_k) = \alpha + \beta(\text{SCT}) + \beta_i(x_i)$$

Where Y_k is the mean count of a particular outcome, in the exposed or unexposed strata. Poisson regression will be used to predict the logarithmic mean count of subclinical brain infarctions in participants with and without SCT, controlling for clinical covariates. The prevalence ratio is calculated by exponentiating the regression coefficient β(SCT), and is interpreted as the relative increase in number of SBI associated with SCT, compared to study participants with normal hemoglobin. The Poisson distribution will be confirmed with goodness of fit statistics, by calculating the deviance to degrees of freedom ratio[165]. While a deviance/d.f. ratio close to one suggests good fit, a ratio exceeding 1.5 indicates overdispersion, leading to estimates with spurious standard errors.

10.4.a Alternative Approach

As an alternative, prevalence of SBI may be dichotomized as a yes / no variable, and the baseline prevalence difference or prevalence ratios may be calculated, binomial log-risk models. While this is a simpler approach than considering the actual count of SBI, these analyses result in loss of information.

10.5. Subclinical Brain Infarction Incidence

Incident subclinical brain infarction will be defined by new infarctions detected at MRI 2 (2004-2006), which were not present at MRI 1 (1993-1995). Previous analyses of incident SBI in the literature have relied on logistic regression, restricting the study population to study subjects participating in both MRI 1 and MRI 2[116].

$$\ln [P(X)/(1-P(X)] = \alpha + \sum [\beta(SCT) + \beta_i(x_i) + \ldots]$$

The dependent variable for the logistic model, or logit, is a log-transformed odds ratio. The SCT beta coefficient describes the increment to the log-odds of incident SBI in participants with SCT, taking other covariates into account. The odds ratio comparing participants with and without SCT is estimated by exponentiating the SCT regression coefficient. Assuming the incidence of SBI is low (<10%), the odds ratio can be considered an approximation of risk. Because the time interval between each MRI exam is likely not constant for all study participants, the actual time between each exam will be entered as a covariate. Logistic regression assumes continuous variables will be linear on the log scale, and that all covariate pairs have multiplicative joint effects. Linearity assumptions will be verified by plotting continuous variable on the log scale, and multiplicativity assumptions tested by entering cross-product interaction terms into the model.

10.5.a. Alternative Approach

As an alternative to logistic regression, incident SBI may be analyzed by Cox regression. The presence of incident subclinical brain infarctions between MRI 1 (1993-1995) and MRI 2 (2004-2006) would be dichotomized, and hazard ratios will be estimated. The hazard ratio is the antilog of the average of differences in log hazards for a particular outcome[166], and is calculated by Cox regression:

$$h(t) = h_0 * \exp[\beta(SCT) + \beta_i(x_i)]$$

Where h(t) is the hazard function, and h_0 is the baseline hazard in unexposed individuals with a value of zero for each of the covariates[167]. Baseline hazards cancel out in the hazard ratio between two individuals, $h(t_1) / h(t_2)$, written logarithmically as $[\log h(t_1) – \log h(t_2)]$. This difference in log hazards is averaged across all individuals in the risk set with a particular

covariate pattern, and smoothed over time at risk to give the hazard ratio. Cox regression makes no assumptions concerning the survival function, and can be used to model continuous, stepwise, parametric, and nonparametric distributions[167]. However, as with all derivatives of linear models, Cox regression assumes a linear relationship between the predictor variables and the log hazard. Functional forms of predictor variables and linearity assumptions will be tested by plotting Martingale residuals, which can be used to reveal departures from linearity[168]. Cox regression also assumes the hazard ratio for the exposed to unexposed individuals remains proportional over time. Nonproportionality will be assessed by plotting Shoenfeld residuals, or by visually inspecting the plots of the negative log of the cumulative hazards[167]. When the slopes of the negative log cumulative hazards for the exposed and unexposed groups cross, proportionality is considered violated. In this event, interaction terms between covariates and time can be entered into the model, or the hazard ratio can simply be interpreted as an average effect over the range of times observed[167]. However, Cox regression may not be appropriate for our study design, due to the interval censored data and potential for large number of ties, assuming follow-up exams were performed over nearly uniform time intervals.

11. GENETIC MODIFICATION

Modification of the associations between SCT and cerebrovascular endpoints by traditional stroke risk factors, stroke candidate genes, sickle haplotypes, and alpha thalassemia will assessed by exploratory multivariable regression modeling. Each potential genetic modifier will be entered into the regression models separately. Cross-product interaction terms between SCT and potential modifiers will assess interaction on the additive scale for ordinary linear regression models, and interaction on the multiplicative scale for

Poisson models and Cox regression models. Effect measure modification will be considered changes to model variance or the main effect estimate, but will be considered exploratory.

12. POWER CALCULATIONS

A total of 960 African Americans were examined in ARIC by cerebral MRI. After removing subjects with prevalent stroke / TIA, and accounting for missing covariates or principal components, we estimate an effective sample size of 600. The association between SCT and prevalence of WML is not well known. In a single pediatric study, WML were detected in 10% of children with SCT, but in none of their sibling controls with homozygous hemoglobin A[51]. In the ARIC study, WML were detected in 80.7% of African Americans overall[111]. Using this as a reference, we expect 80% power to detect a prevalence difference of 0.15 in participants with SCT, compared to those without (Figure 17). While the pediatric study may not be directly comparable to older African Americans; the reported effect size (prevalence difference of 0.10) is close to 0.15, the effect that our study is powered to estimate.

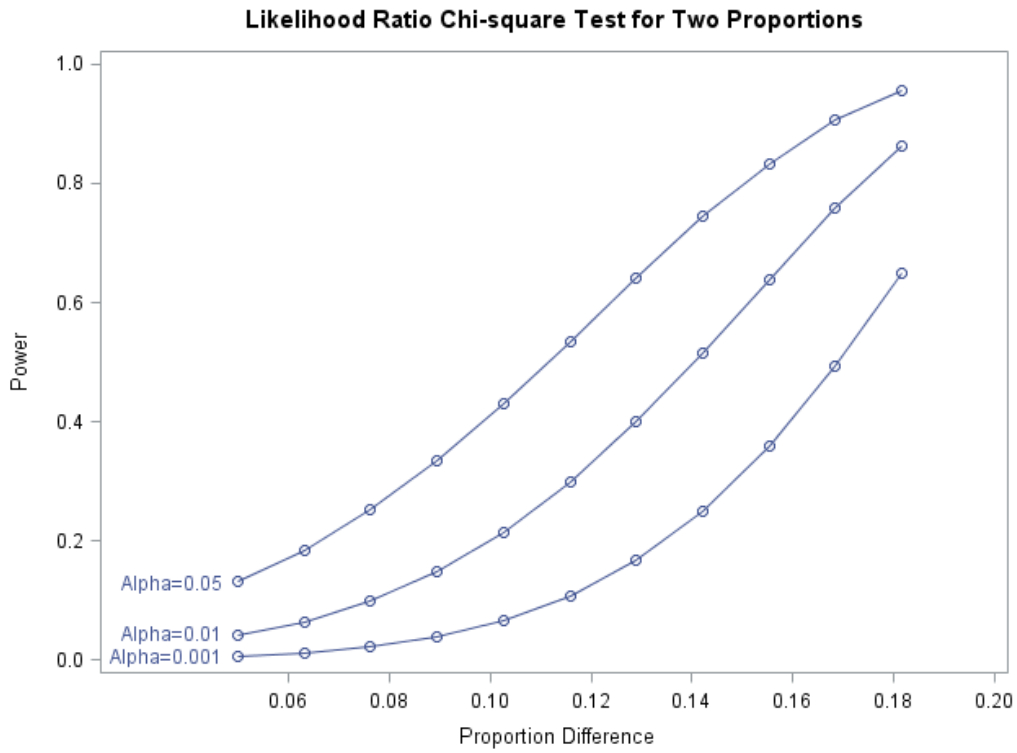


Figure 6. Power calculations for baseline WML prevalence difference associated with SCT status

To date, the association between SCT and WML severity has been unexplored. With WML severity scored by a continuous, ordinal (0 to 9) scale, with an estimated standard deviation of 1[111], we expect 80% power to detect a 0.5 difference in mean WML severity score (Figure 19).

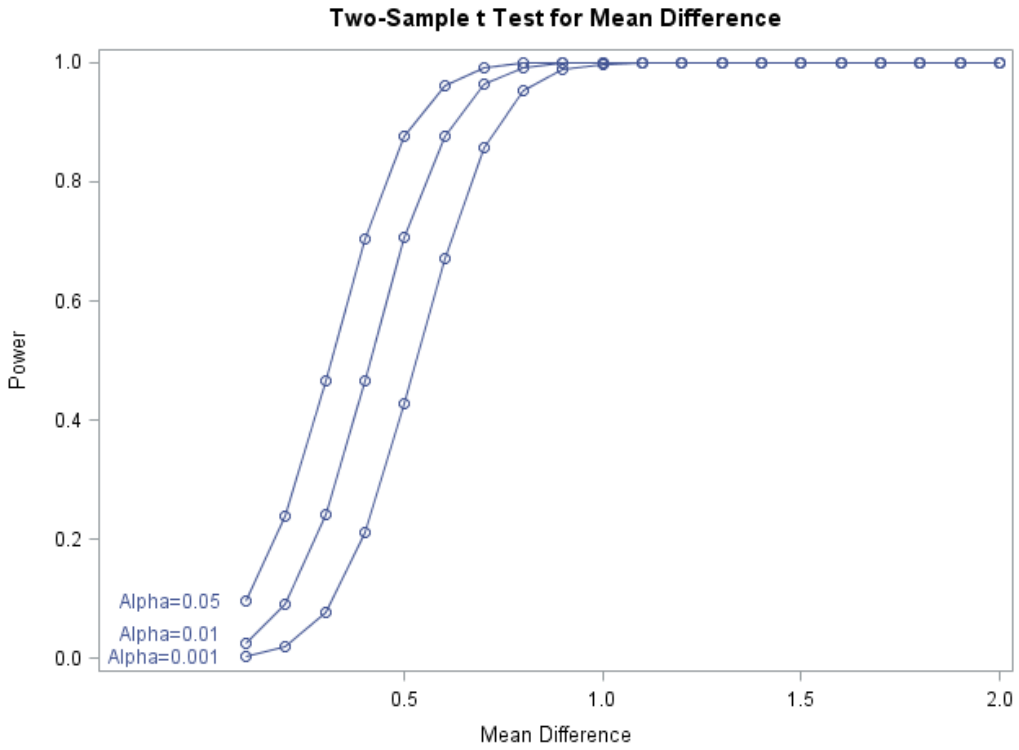


Figure 7: Power calculation for mean WML severity scores, associated with SCT status

MANUSCRIPT 1

Sickle Cell Trait and White Matter Lesions in the Atherosclerosis Risk in Communities (ARIC) Study

Melissa C. Caughey, MPH[1,2]; Laura R. Loehr, MD, PhD[2]; Nigel S. Key, MB, ChB[1];

Vimal K. Derebail, MD, MPH[1]; Rebecca F. Gottesman, MD, PhD[3];

Abhijit V. Kshirsagar, MD, MPH[1]; Megan L. Grove, MS[4]; Gerardo Heiss, MD, PhD[2];

Thomas H. Mosley, PhD[5]; Kari E. North, PhD[2]

[1]Department of Medicine, University of North Carolina at Chapel Hill; Chapel Hill, NC

[2]Department of Epidemiology, University of North Carolina at Chapel Hill; Chapel Hill, NC

[3]Department of Neurology, Johns Hopkins University; Baltimore, MD

[4] School of Public Health, Human Genetics Center, The University of Texas Health Science Center; Houston, TX

[5]University of Mississippi Medical Center; Jackson, MS

Short Title: Sickle Cell Trait and White Matter Lesions: ARIC study

Corresponding author:

Melissa C. Caughey, MPH

University of North Carolina at Chapel Hill

Email: caughey@med.unc.edu

Abstract

Background and Purpose: Cerebral white matter lesions (WML) are associated with adverse neurologic outcomes, and have been detected in individuals with sickle cell trait (SCT). To date, no large, epidemiological studies have investigated this association.

Methods: A population-based sample of African Americans (N=844, mean age=62, female=64%) was prospectively imaged by cerebral MRI in 1993-1995 in the Atherosclerosis Risk in Communities (ARIC) Study, with 470 (56%) returning for a follow up MRI in 2004-2006. Hemoglobin S (rs334) and hemoglobin C (rs33930165) were directly genotyped, and participants identified with sickle cell anemia or hemoglobin C excluded. WML on MRI were evaluated and scored using a qualitative, 10-point scale. Associations between SCT and WML prevalence and severity were analyzed using ordinal logistic and linear regression. All models were adjusted for age, sex, smoking, diabetes, hypertension, and 10 ancestral principal components.

Results: SCT was identified in 56 (6.6%) participants at the baseline MRI. Individuals with SCT had more prevalent (86% vs. 79%), and more severe (mean score 1.5 vs. 1.3) WML than individuals without SCT. SCT was associated with a 20% increased odds of WML prevalence (POR 1.2, 95% CI: 0.7 – 2.0), and an adjusted mean severity score that was 0.2 (-0.1 – 0.5) points higher.

Conclusion: There was a subtle but insignificant trend for greater WML prevalence and severity associated with SCT.

Introduction

White matter lesions (WML), the result of cerebral small vessel disease, are detectable by hyperintensities in the cerebral white matter by MRI. Though often unaccompanied by overt neurological symptoms, WML have been shown to predict cognitive impairment, dementia, and stroke[114-117]. In elderly populations, WML are particularly prevalent[112], with a greater severity observed in African Americans than whites[111]. When examined longitudinally, African Americans appear to have an accelerated progression of WML severity as well[113].

The etiology of WML is believed to be vascular in origin, and possibly related to chronic hypoperfusion[107]. Consistent with this hypothesis, a greater prevalence of WML has been observed in patients with sickle cell anemia, a Mendelian hemoglobinopathy complicated by microvascular occlusion, hypoperfusion, and heightened stroke risk[102]. Sickle cell trait (SCT), the heterozygous carrier state of sickle cell anemia, is prevalent in 7–9% of African Americans and 0.2% of non-Hispanic whites[19], affecting over 3 million Americans. Though hypoperfusion is not a well-established phenotype of SCT, a greater prevalence of WML has been observed in pediatric carriers in a single study[51]. Individuals with SCT also appear to have increased risk of ischemic strok[169, 170]. However, the pathophysiological mechanisms of stroke associated with SCT are poorly understood, and may be related to cerebral small vessel disease. To date, no large, epidemiological studies have examined the prevalence or the progression of WML in African American adults with SCT. We hypothesized that SCT would be associated with a greater prevalence, severity, and progression of WML in older African Americans imaged prospectively by cerebral MRI in the Atherosclerosis Risk in Communities (ARIC) Study.

Methods

The ARIC study is a biracial, prospective epidemiological cohort based on 4 U.S. areas. Included in this study is a population-based sample (N= 4,270) of African Americans aged 45-65 was recruited with informed consent in 1987-1989 from urban, suburban, and rural areas of Jackson, Mississippi and Forsyth County, North Carolina. The ARIC study includes 5 cohort examinations spaced approximately 3 years apart (with a larger gap between the 4th and 5th visits), with annual telephone surveys during interim years, and ongoing surveillance of hospitalized events, and an additional ancillary study visit (the Brain MRI visit) in a subset of participants. All study protocols were approved by the University of Mississippi and Wake Forest University Institutional Review Boards.

Brain MRI was offered to a random subset of participants in 1993-1995, with a follow up ancillary exam offered in 2004-2006. Standard imaging exclusions included metallic implants, aneurysm clips, hearing aids, cardiac pacemakers, spinal cord stimulators or other internal electrical devices. Examinations were performed by certified radiographers, using standardized protocols and consistent imaging equipment (GE Signa 1.5 T) for both the baseline and follow-up MRI. Axial images were aligned parallel to the anterior commissure / posterior commissure line, and contiguous images were captured in 5 mm slices, from the vertex of the skull to the base. The cerebral white matter was imaged by proton density weighting, acquired with a repetition time of 3000 ms and echo time of 30 ms. All images were read by board-certified ARIC radiologists with subspecialty training in neuroradiology. Leukoaraiosis was defined by diffuse areas of white matter hyperintensities, and severity was graded using a qualitative 10-point scale, by matching lesion patterns to reference images[112].

The inter- and intra-reader reliability coefficients in the ARIC study were .68 and .71, respectively[72].

For the purposes of our analysis, prevalent stroke was considered a cerebrovascular accident occurring prior to the baseline MRI in 1993-1995. Stroke history was initially determined at the ARIC study onset in 1987-1989, by a validated computer algorithm[64], and thereafter by hospital surveillance with verification by physician review. The diagnostic criteria and adjudication have previously been described[60].

Hemoglobin S and hemoglobin C were directly genotyped, using TaqMan® high throughput assays, as previously described. For quality assurance, blind duplicate genotyping was performed in a random sample representing 5% of the total assays (kappa coefficients 0.83 for hemoglobin S, and 0.93 for hemoglobin C). In addition to direct genotyping, high coverage DNA sequencing was performed. Any discrepancies were identified and adjudicated by review of quality control data and re-genotyping, to yield the final hemoglobin S and hemoglobin C classifications.

Ancestry was quantified using EIGENSTRAT 5.0.1 (David Reich, open source), based on genomic variation characterized by the HumanExome BeadChip v1.0 (Affymetrix, Santa Clara, CA), as previously described[171]. First degree relatives were identified by PLINK (Shaun Purcell, http://pngu.mgh.harvard.edu/purcell/plink)[172].

Clinical covariates were ascertained at time of the first MRI examination. Seated blood pressures were measured by random-zero mercury manometers. A total of 3 measurements were made, with the average of the second and third recorded for analysis. Hypertension was considered a systolic blood pressure $\geq$ 140 mmHg, a diastolic blood

pressure ≥ 90 mmHg, or antihypertensive medication use. Fasting cholesterol and glucose were assessed by ARIC central laboratories. Hypercholesterolemia was considered a total fasting cholesterol ≥ 6.2 mmol/L. Diabetes was defined by either a blood glucose level ≥ 7 mmol/L, self-reported diabetes, or use of diabetic medications. Current smoking status was determined by self-report.

Statistical Analysis

All analyses were performed using SAS 9.3 (SAS Institute, Cary, NC). Categorical variables were compared by a χ^2 test, and continuous variables compared by analysis of variance. Categorical variables with expected cell counts < 5 were analyzed using Fisher's Exact test. Multivariable analyses contrasting SCT with homozygous hemoglobin A (HbAA) were adjusted for age, sex, white matter lesion risk factors (smoking, hypertension, and diabetes)[107], and 10 ancestral principal components. The association between SCT and WML severity score was assessed by linear regression, modeling WML score as a continuous, ordinal variable. After verifying proportional odds, associations between SCT and WML prevalence were analyzed by ordinal logistic regression, contrasting a WML score of at least 3 to scores of 2, 1, or 0; a score of at least 2 to scores of 1 or 0; and a score of at least 1 to a score of 0. In a subset of study participants returning for the second MRI, progression of WML severity score was assessed by linear regression, modeling the change in score associated with SCT, with adjustments for age, sex, smoking, hypertension, diabetes, and elapsed time between each MRI examination. Progression of WML was also dichotomized and analyzed by logistic regression. WML progression was considered an increase of at least 2 severity points, to minimize the likelihood of spurious progression due simply to inaccuracies of the qualitative severity assessment.

Power calculations for difference in WML prevalence and severity score were conducted *a priori*, based on an assumed study population of 800, SCT prevalence of 7%, WML prevalence of 80%, and mean WML score standard deviation of 1.0. We expected 80% power to detect a WML prevalence difference of 14%, and mean WML score difference of 0.4, with significance at the $\alpha = 0.05$ level (2-sided).

Results

A total of 937 African Americans were genotyped for hemoglobin S and hemoglobin C, and examined by MRI in 1993-1995. First degree relatives (n=18), those with missing or inadequate genotype calls (n=3), and participants identified with hemoglobin C trait (N=24), hemoglobin SC disease (n=1), or sickle cell anemia (n=1) were excluded. To minimize measurement error of WML due to previous stroke, we additionally excluded 26 with history of cerebrovascular accident. After removing subjects with missing clinical covariates (n=18) a final study population of 844 remained.

Out of 844 African Americans, 56 (6.6%) were identified with SCT. The mean age at the baseline MRI was 62 years, and 64% were women. Study participants with SCT were more often men (50% vs. 35%; p=0.02) and less obese (28 kg/m^2 vs. 30 kg/m^2 BMI; p=0.007), with a slightly greater prevalence of current smoking. (**Table 1**).

White matter lesions were detected in the majority of study participants (80%). A greater prevalence was observed in those with SCT (86% vs. 79%; p=0.2) than without SCT, yielding an unadjusted prevalence difference of 7% (-0.3 – 16%; p=0.2). Severity scores of WML ranged from 0 to 8, with an overall mean score of 1.3 ± 1.2. The mean WML score among those with SCT was 1.5 (1.2 – 1.9); slightly higher than the mean score among those

with HbAA, 1.3 (1.2 – 1.4). When severity scores were categorized into scores of 1, 2, or 3+, a consistent but insignificant pattern of higher WML scores was observed in the group with SCT, compared to those with HbAA (**Table 2**).

In multivariable regression analysis controlling for age, sex, ancestry and WML risk factors (smoking, hypertension, and diabetes), SCT was associated with a WML mean score that was 0.2 (-0.1 – 0.5) points higher than HbAA study participants, but was not statistically significant. As shown in **Figure 1**, the odds of WML associated with SCT were similar among differing cut points (contrasting a WML score of at least 3 to scores of 2, 1, or 0; a score of at least 2 to scores of 1 or 0; and a score of at least 1 to a score of 0). As a result, we present the cumulative POR of 1.2 (0.7 – 2.0) as a summary measure, which takes the ordinality of WML scores into account.

Of the 844 study participants examined by brain MRI in 1993-1995, 470 (56%) returned for the follow-up MRI in 2004-2006. Within this subset, 36 (7.5%) were identified with SCT. The mean elapsed time between the baseline and follow up MRI examinations was 10.8 ± 0.9 years, which did not differ by SCT status (p=0.8). By the second MRI, WML prevalence (defined by any score greater than zero) was nearly ubiquitous (100% in participants with SCT, and 98% in those with HbAA). The overall mean WML severity score was 2.3 ± 1.5, and did not differ by SCT status (p= 0.9). The overall mean increase in WML scores between the baseline and follow up scans was 1.1 ± 1.1 points; however, progression was marginally attenuated in those with SCT, compared to HbAA participants (0.9 ± 0.9 points, compared to 1.2 ± 1.1 points, respectively; p=0.1). With WML score change modeled as a function of SCT status, adjusted for age, sex, smoking, diabetes, hypertension, elapsed time between MRI examinations, and ancestry, SCT was associated with a progression that

was 0.3 points lower (95% CI: 0.6 points lower to 0.3 points higher) than those with HbAA. When WML progression was analyzed dichotomously, SCT was associated with a 30% lower adjusted risk of WML progression (OR = 0.7, 95% CI: 0.3 – 1.5). Additional analyses are shown in **Appendix I**.

Discussion

At the baseline MRI examination, we observed a subtle, yet imprecise, elevation in WML prevalence and severity in study participants with SCT, compared to HbAA participants. The WML prevalence difference, 7%, was lower than our projected power to detect a difference of 14%. Similarly, the adjusted increase in WML score associated with SCT, 0.2 points, was less than 0.4 points, the minimum score difference our study was powered to detect. Though not statistically significant, there was a consistent pattern of increased WML prevalence and severity among those with SCT. Among the subset returning for the follow up MRI, SCT did not appear to be associated with increased progression in WML severity; however, estimates were limited by imprecision and small sample sizes.

Although generally asymptomatic, WML have been shown to predict adverse neurocognitive outcomes[114-117]. In the Leukoaraiosis and Disability Study, a European multicenter collaboration including 639 patients aged 65-84, WML severity independently predicted disability, dementia, and cognitive impairment[114]. Consistent with this finding, the Cardiovascular Health Study, which longitudinally examined WML progression over a 5 year interval in 1,921 elderly subjects, reported decline in cognitive function and increased stroke risk associated with WML progression[117].

In cross-sectional studies, both sickle cell anemia and SCT have been associated with increased WML prevalence. In an observational study of 141 young adults with sickle cell anemia and 44 age-, sex-, and education-matched African American controls with normal hemoglobin, WML were detected in 15% with sickle cell anemia, compared to 7% of the controls[102]. Though suggestive of an increased risk of WML associated with sickle cell anemia, the prevalence difference, 8% (95% CI: -1.4 – 17.6%), was imprecise. Although to our knowledge only one study has been conducted, a greater prevalence of WML has also been observed in carriers. In a small, pediatric case-control study with prospective MRI screenings, WML was detected in 10% of children with SCT, but in none of the sibling controls[51].

The etiology of WML remains uncertain, but is thought to be vascular in origin. The deep white matter is particularly vulnerable to ischemia and hypoperfusion. Its network of small penetrating arteries and arterioles has a poor anastomatic system of collateral vessels, supplied solely by the cortical and leptomeningeal arteries[107, 108]. In animal experiments, WML have been induced by eliciting cerebral ischemia in gerbils[109] and rats[110]. While animal studies are not always generalizable to humans, human brain autopsies have confirmed a 20% decreased concentration of arterioles and afferent capillaries within white matter lesions, compared to healthy white matter tissue[106]. The decrease in capillary density may be the result of vascular occlusion or chronic ischemia, inducing downstream endothelial cell apoptosis by blocking the flow of nutrients[173].

Microvascular hypoperfusion is a known complication of sickle cell anemia, as sickled erythrocytes impede flow through the capillary beds, causing vaso-occlusive crises, target organ damage, and pain. Though less established, there is some evidence that carrier

status is associated with microvascular hypoperfusion as well. In the resting state, only 1% of erythrocytes in the venous circulation of individuals with SCT are sickled, in contrast to 30-60% in those with sickle cell anemia[10]. However, even in the absence of sickling, blood assays of SCT carriers indicate abnormal red cell rigidity[174], which may impede the microcirculation. While a capillary lumen diameter typically measures 4-5 μm, red blood cells span 8-9 μm[175]. In order to pass through the capillaries, erythrocytes must be flexible. As a possible adaption to diminished red cell flexibility, capillaries in the skeletal muscle of individuals with SCT are larger in diameter, but sparser, with an overall reduced surface area for gas exchange[175]. Along with muscle biopsies, hypoperfusion has also been assessed indirectly, using the hematocrit to blood viscosity ratio. This index has been observed to be reduced in SCT carriers as well[174]. However, it is unknown if these indices of hypoperfusion observed in individuals with SCT indicate hypoperfusion of the cerebral microcirculation.

We observed an increased frequency and severity of WML in African American study participants with SCT, compared to those with HbAA genotypes. However, our estimates were imprecise, which may reflect either no association between SCT and WML, or deficiencies of the study design. The WML scores were subject to measurement error, with an intra-reader reliability coefficient of 0.71. The sample size of participants with SCT (N=56) at the baseline MRI was also quite small for an epidemiological study, with an attrition rate of 44% by the follow up exam. To overcome the variability inherent with small sample sizes, a partnership with other cohort studies examining African Americans by cerebral MRI may be worthwhile.

Additionally, white matter abnormalities associated with SCT may be better detected by newer MRI scanners with enhanced resolution. Volumetric analyses of WML have greater

reliability than qualitative scoring, improving the precision of the outcome. Diffusion tensor imaging with fractional anisotropy has also been shown to detect subtle white matter abnormalities, which were underestimated by conventional imaging of patients with sickle cell anemia[176, 177]. Fractional anisotropy and volumetric analyses, which will be available in the 3rd MRI examination (2011 – 2013) of the ARIC study, may enhance detection of possible white matter abnormalities associated with SCT.

In conclusion, the relation between SCT and WML, if any, appears to be subtle. We did not detect a statistically significant association within the subset of ARIC study participants examined by cerebral MRI. A larger sample of individuals with SCT, and greater precision of WML measurements, may improve the ability to detect an association.

Table 1: Baseline (1993-1995) demographics and clinical characteristics of African American participants in the ARIC brain MRI examinations

Characteristic	SCT (N=56) *Mean ± SD or N (%)*	HbAA (N=788) *Mean ± SD or N (%)*	*P-value*
Demographics			
Age (years)	61 ± 5	62 ± 5	0.6
Female	28 (50%)	512 (65%)	0.02
Medical History			
Current smoker	15 (27%)	142 (18%)	0.1
Hypertension	32 (57%)	494 (63%)	0.4
Hypercholesterolemia	9 (16%)	162 (21%)	0.4
Diabetes	15 (27%)	196 (25%)	0.7
Coronary heart disease	1 (2%)	33 (4%)	0.7
Body mass index (kg/m^2)	28 ± 4	30 ± 5	0.007

Table 2: Frequencies of WML scores at baseline MRI (1993-1995), stratified by SCT status

Genotype	Score = 0	Score = 1	Score = 2	Score = 3+
SCT	8 (14%)	30 (54%)	10 (18%)	8 (14%)
HbAA	163 (21%)	403 (51%)	125 (16%)	97 (12%)

*Cochrane-Maental-Haenzsel p = 0.3

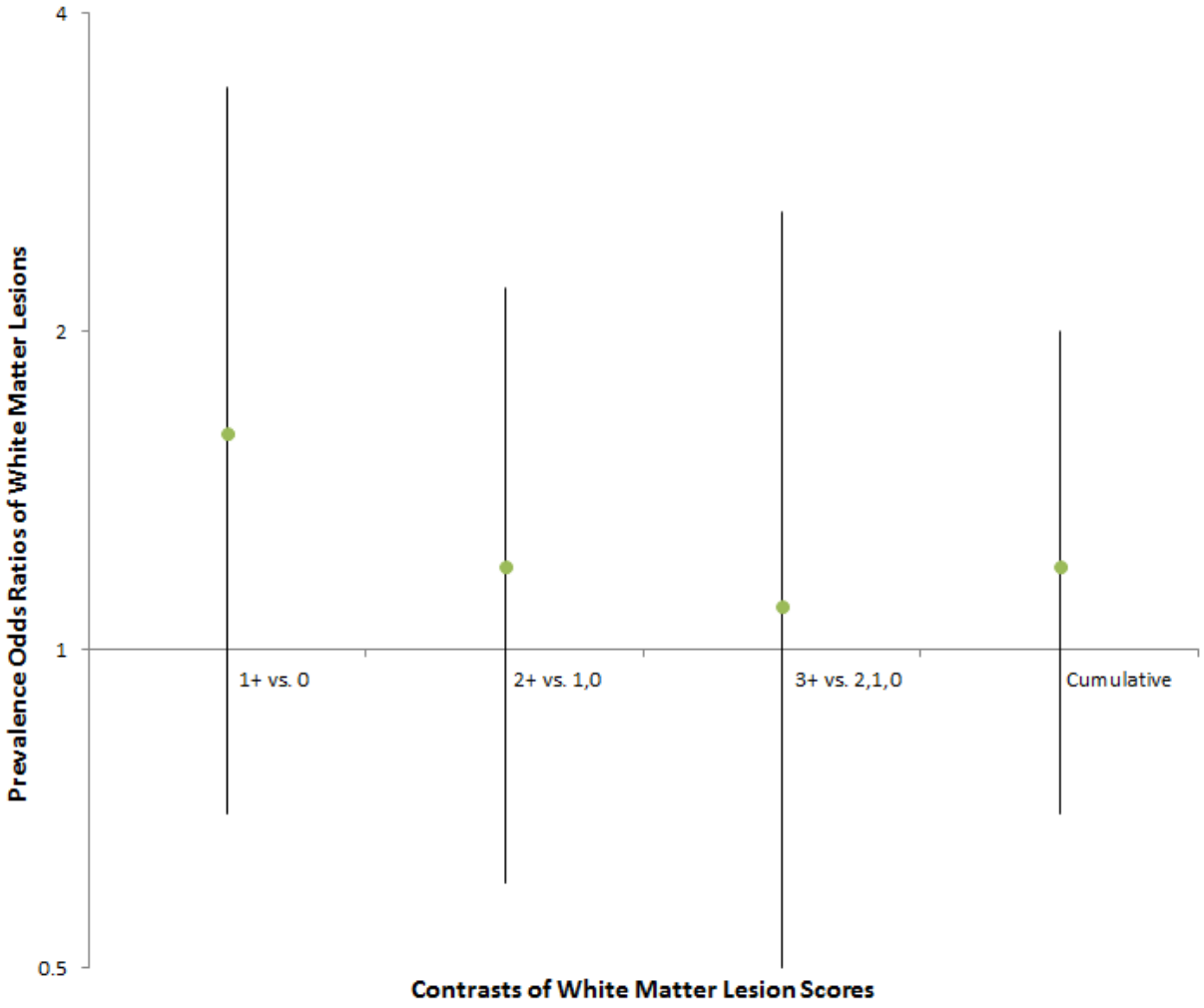


Figure 1: Individual and cumulative prevalence odds ratios of WML associated with SCT. Contrasts = score of 1 or greater vs. score of 0; score of 2 or greater vs. score of 1 or 0; score of 3 or greater vs. score of 2, 1, or 0. All models adjusted for age, sex, hypertension, smoking, diabetes, and ancestry. The cumulative odds ratio is presented as a summary measure.

Acknowledgments

The authors thank the staff and participants of the ARIC study for their important contributions.

Funding Sources

The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HSN268201100006C, HHSN268201100007C, HHSN268201100008C, HSN268201100009C, HHSN268201100010C, HHSN268201100011C, and SN268201100012C), with the ARIC Brain MRI examination funded by R01-HL70825. Genotyping of hemoglobin S and hemoglobin C was funded by grant 5K12HL087097. Support for ARIC exome chip genotyping was provided by Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium through the National Institutes of Health (NIH) American Recovery and Reinvestment Act of 2009 (ARRA) (5RC2HL102419). NSK was supported by 1UO1HL117659.

Disclosures

None.

MANUSCRIPT 2

Sickle Cell Trait and Subclinical Brain Infarction in the Atherosclerosis Risk in Communities (ARIC) Study

Melissa C. Caughey, MPH[1,2]; Laura R. Loehr, MD, PhD[2]; Nigel S. Key, MB, ChB[1];

Rebecca F. Gottesman, MD, PhD[3]; Vimal K. Derebail, MD, MPH[1]; Gerardo Heiss, MD, PhD[2];

Abhijit V. Kshirsagar, MD, MPH[1]; Megan L. Grove, MS[4]; Thomas H. Mosley, PhD[5];

Kari E. North, PhD[2]

[1]Department of Medicine, University of North Carolina at Chapel Hill; Chapel Hill, NC

[2]Department of Epidemiology, University of North Carolina at Chapel Hill; Chapel Hill, NC

[3]Department of Neurology, Johns Hopkins University School of Medicine; Baltimore, MD

[4] School of Public Health, Human Genetics Center, The University of Texas Health Science Center; Houston, TX

[5]University of Mississippi Medical Center; Jackson, MS

Short Title: Sickle Cell Trait and Lacunes: ARIC study

Corresponding author:

Melissa C. Caughey, MPH

University of North Carolina at Chapel Hill

Email: caughey@med.unc.edu

Abstract

Background and Purpose: Though frequently asymptomatic, infarctions of the small cerebral arteries predict adverse neurologic outcomes. The etiology of subclinical brain infarctions (SBI) is heterogeneous, arising from lipohyalinosis, microatheroma, and thrombosis. It is unknown if sickle cell trait (SCT), a genotype associated with hypercoagulability, is related to increased SBI risk.

Methods: A population-based sample of African Americans (N=828, mean age=62, female=64%) was prospectively imaged by cerebral MRI in 1993-1995 in the Atherosclerosis Risk in Communities Study (ARIC), with 477 (58%) returning for a follow up MRI in 2004-2006. Hemoglobin S (rs334) and hemoglobin C (rs33930165) were directly genotyped, and participants identified with sickle cell anemia or hemoglobin C excluded. SBI was detected in the cortical and subcortical regions by T1 and T2 weighted MRI, and in the white matter by T1 weighted imaging. Associations between SCT and the prevalence and progression of SBI were analyzed using logistic regression, adjusted for age, sex, smoking, diabetes, hypertension, total cholesterol, prevalent coronary disease, and 10 ancestral principal components.

Results: At the baseline MRI, SCT was prevalent in 55 (6.6%). SBI was detected in 14%, with a lower prevalence in participants with SCT (11% vs. 15%). SCT was not associated with prevalent SBI (POR = 0.7; 95% CI: 0.3 – 1.8), or incidence of new infarctions by the follow up exam (OR = 1.4; 95% CI: 0.6 – 3.1).

Conclusion: The associations between SCT and SBI remain elusive. Our estimates of effect are insignificant and demonstrate wide variability, possibly due to small sample sizes.

Introduction

Though often asymptomatic, cerebral infarctions of the small, deep, penetrating arteries have been shown to predict adverse neurological outcomes and future stroke[104, 105]. The etiology of subclinical brain infarction (SBI) is mixed, resulting from vascular occlusion by lipohyalinosis, microatheroma, and thrombosis[98]. Evidence of SBI is inferred from magnetic resonance imaging (MRI), by the presence of lacunes, small fluid filled cavities of necrosed cerebral tissue resulting from infarction. In observational studies, the prevalence of SBI appears to be correlated with age, increasing from 8% in 55-59 year olds to 23% in 65-72 year olds[71], to 43% for those over the age of 85 years[100]. A greater number of infarctions are also reported in African Americans. In a previous analysis from the ARIC baseline MRI examination, SBI prevalence was detected in 21% of African Americans, compared to 10% of white study participants[71].

Even in pediatric populations, a high prevalence of SBI has been observed in patients with sickle cell anemia[101, 178, 179], a Mendelian hemoglobinopathy complicated by hypercoagulability, microvascular occlusion, and elevated stroke risk. The carrier state, sickle cell trait (SCT) is prevalent in over 3 million Americans, with a heterozygous allelic frequency of 7-9% in African Americans and 0.2% of non-Hispanic whites[180]. Though generally considered benign, the SCT genotype has been associated with thrombosis and elevated factors of coagulation[181]. The prevalence of SBI in individuals with SCT has never been reported, and it is unknown if the hypercoagulability associated with SCT increases SBI risk. We conducted an epidemiological investigation of the prevalence and progression of SBI associated with SCT, using prospective MRI data from the Atherosclerosis Risk in Communities (ARIC) Study.

Methods

The ARIC study is a biracial, prospective, epidemiological cohort representing 4 U.S. areas. A population-based sample (N= 4,270) of African Americans aged 45-65 was recruited with informed consent in 1987-1989 from urban, suburban, and rural areas of Jackson, Mississippi and Forsyth County, North Carolina. The ARIC study includes 5 cohort examinations, with annual telephone surveys during interim years and ongoing surveillance of hospitalized events. All study protocols were approved by the University of Mississippi and Forsyth County Institutional Review Boards.

Brain MRI was offered to a random subset of participants in 1993-1995, with a follow up exam offered in 2004-2006. Standard imaging exclusions included metallic implants, aneurysm clips, hearing aids, cardiac pacemakers, spinal cord stimulators or other internal electrical devices. Examinations were performed by certified radiographers, using standardized protocols and consistent imaging equipment (GE Signa 1.5 T) for both the baseline and follow up MRI. Axial images were aligned parallel to the anterior commissure / posterior commissure line, and contiguous images were captured in 5 mm slices, from the vertex of the skull to the base. T1 weighted images were acquired with a repetition time of 500 ms and echo time of 20 ms, and T2 weighted images were captured using a 3000 ms repetition time and echo time of 100 ms. All images were read by board-certified ARIC radiologists with subspecialty training in neuroradiology.

Lacunes were identified in the cortical gray matter and deep nuclear regions by signal hyperintensities using T1 and T2 weighted MRI, and in the white matter and brain stem by hypointensities in T1-weighted images. SBI diameters were measured by electronic calipers, using the largest dimension within the contiguous MRI slices. Because small periventricular

spaces are often indistinguishable from lacunes, a minimum 3 mm cut point was used. Focal, non-mass lesions detected within territories of vascular distributions were considered SBI when measuring 3-20 mm, in asymptomatic individuals with no prior stroke. Inter- and intrareader agreement rates among ARIC neuroradiologists were 79% and 82%, respectively[71].

For the purposes of our analysis, prevalent stroke was considered a cerebrovascular accident occurring prior to the baseline MRI in 1993-1995. Stroke history was determined at the ARIC study onset in 1987-1989, by self-reported signs and symptoms and a validated computer algorithm[64]. Incident stroke over the course of follow up was captured by hospital surveillance. Hospital discharge lists for study participants with ICD-9 codes 430-438 were provided to the ARIC study, and stroke diagnoses were verified by physician review of the discharge summary, imaging reports, neurological consults, and medical history, as previously described[60]. For quality assurance, stroke diagnosis was also determined by a computer algorithm. Any disagreements between the physician diagnosis and computer algorithm were adjudicated by a second physician reviewer. Agreement rates between the physician reviewer and computer algorithm were 78%[60]. In the majority of discordant diagnoses (65%), the physician adjudicator agreed with the physician reviewer, rather than the computer algorithm[60].

Hemoglobin S and hemoglobin C were directly genotyped, using TaqMan® high throughput assays, as previously described. For quality assurance, blind duplicate genotyping was performed in a random sample representing 5% of the total assays (kappa coefficients 0.83 for hemoglobin S, and 0.93 for hemoglobin C). In addition to direct genotyping, high coverage DNA sequencing was performed. Any discrepancies were identified and

adjudicated by review of quality control data and re-genotyping, to yield the final hemoglobin S and hemoglobin C classifications.

Ancestry was quantified using EIGENSTRAT 5.0.1 (David Reich, open source), based on genomic variation characterized by the HumanExome BeadChip v1.0 (Affymetrix, Santa Clara, CA), as previously described[171]. First degree relatives were identified by PLINK (Shaun Purcell, http://pngu.mgh.harvard.edu/purcell/plink)[172].

Clinical covariates were ascertained at time of the first MRI examination. Seated blood pressures were measured by random-zero mercury manometers. A total of 3 measurements were made, with the average of the second and third recorded for analysis. Hypertension was considered a systolic blood pressure ≥ 140 mmHg, a diastolic blood pressure ≥ 90 mmHg, or antihypertensive medication use. Fasting cholesterol and glucose were assessed by ARIC central laboratories. Hypercholesterolemia was considered a total fasting cholesterol ≥ 6.2 mmol/L. Diabetes was defined by either a blood glucose level ≥ 7 mmol/L, self-reported diabetes, or use of diabetic medications. Standardized, 12-lead electrocardiograms were performed, and assigned a Minnesota code[158] by the ARIC ECG Reading Center. Prevalent coronary heart disease was defined by self report, history of myocardial infarction, coronary artery bypass graft, or percutaneous coronary intervention; or electrocardiogram (ECG) suggestive of prior myocardial infarction. ECGs with Minnesota codes indicative of myocardial infarction (ST elevation or depression, or inverted T-waves), were visually inspected and adjudicated by the Minneapolis ECG Reading Center. Current smoking was determined by self report.

A total of 937 African Americans were genotyped for hemoglobin S and hemoglobin C, and examined by MRI in 1993-1995. First degree relatives (N=18), those with missing or inadequate genotype calls (N=3), and participants identified with hemoglobin C trait (N=24), hemoglobin SC disease (N=1), or sickle cell anemia (N=1) were excluded. Additionally, 26 participants with history of stroke were excluded, along with 36 with missing clinical covariates, resulting in a final study population of 828.

Statistical Analysis

All analyses were performed using SAS 9.3 (SAS Institute, Cary, NC). Categorical variables were compared using a χ^2 test, and continuous variables compared by ANOVA. Categorical variables with expected cell counts < 5 were analyzed using Fisher's Exact test. Multivariable analyses contrasting SCT with homozygous hemoglobin A (HbAA) were adjusted for traditional risk factors of cerebrovascular disease (age, sex, smoking, hypertension, diabetes, total cholesterol, prevalent coronary disease), along with 10 ancestral principal components. Prevalence odds ratios of SBI were assessed by logistic regression, after dichotomizing presence of infarctions at the baseline MRI. In the subset of study participants returning for the second MRI, progression of SBI was dichotomously defined by an increase in infarction number since the first MRI. The odds ratio of SBI progression associated with SCT was assessed by logistic regression, with adjustments for traditional risk factors, ancestry, and elapsed time between each MRI examination.

Power calculations for the prevalence odds ratio of SBI associated with SCT was conducted *a priori*, based on an assumed study population of 800, SCT prevalence of 7%, and SBI prevalence of 20% among the reference (HbAA) group. We hypothesized that SCT

would be associated with greater prevalence of SBI, and expected 80% power to detect a prevalence odds ratio of 2.3, with significance at the $\alpha = 0.05$ level (2-sided).

Results

Out of 828 African Americans, SCT was prevalent in 55 (6.6%) participating in the baseline MRI examination. The mean age at the study onset was 62 years, and 64% were women. Study participants with SCT were more often men (51% vs. 35%; p=0.02) and less obese (28 kg/m^2 vs. 30 kg/m^2 BMI; p=0.006), with a slightly greater prevalence of current smoking. (**Table 1**).

At the baseline MRI, SBI was detected in 14%, with a greater prevalence in participants with the HbAA genotype (15% vs. 11%; p=0.4). The number of detected infarctions ranged from 0 to 5, however, the majority identified with SBI (77%) had a single infarction. When categorized into groups of zero, 1, and more than 1 infarction, individuals with SCT had a lower prevalence of single infarctions, but similar prevalence of multiple infarctions (**Table 2**).

In multivariable regression analysis controlling for age, sex, smoking, hypertension, diabetes, total cholesterol, prevalent coronary disease, and ancestry, SCT was not associated with prevalent SBI at the baseline MRI (POR = 0.7; 95% CI: 0.3 – 1.8), as shown in **Table 3**.

Of the 828 study participants examined by brain MRI in 1993-1995, 477 (58%) returned for the follow up MRI in 2004-2006. Within this subset, 36 (7.5%) were identified with SCT. The mean elapsed time between the baseline and follow up MRI examinations was 10.8 ± 0.9 years, which did not differ by SCT status (p=0.8). An increase in SBI number was detected in 108 participants (23%), with a higher frequency of progression in those with

SCT (28% vs. 22%, p=0.4). Of those classified with SBI progression, the mean increase in infarction number was 1.4 ± 0.7, which was similar among HbAA and SCT genotypes (p=0.6). In multivariable regression analysis controlling for age, sex, smoking, hypertension, diabetes, total cholesterol, prevalent coronary disease, ancestry, and elapsed time between MRI examinations, SCT was not significantly associated new infarctions (OR = 1.4; 95% CI: 0.6 – 3.1). Additional analyses are shown in **Appendix II**.

Discussion

This is the first epidemiological study to examine SCT and SBI prevalence and progression in the general population of African Americans. Contrary to our expectation, SCT was not associated with a higher prevalence of SBI. However, the estimate was widely variable, making it inconclusive. Our initial power calculation assumed a higher risk associated with SCT, and we expected 80% power to detect an SBI odds ratio of 2.3. When examined as a protective exposure, we would have had 80% power to detect an SBI odds ratio of 0.3. Our analysis was therefore underpowered to detect the observed odds ratio of 0.7 (0.3 – 1.8). In the subset of study participants returning for the follow up MRI, we observed a 40% higher risk of SBI progression in those with SCT; however, the estimate was insignificant and limited by the small sample size.

SBI is commonly observed in elderly populations, and though asymptomatic, has been associated with subtle neurological deficits and future cognitive decline. In the Rotterdam Scan Study, which prospectively imaged 1015 stroke-free, elderly subjects by MRI, those with SBI were found to have a steeper decline in cognitive function, and more than twice the risk of developing dementia or Alzheimer's disease[104]. SBI has also been

associated with future stroke. In the Cardiovascular Health Study, examining 3,324 stroke-free elderly subjects by cerebral MRI, SBI was correlated with a 50% higher stroke risk over a 4 year follow up period[105]. Progression of SBI has also been associated with adverse neurological outcomes, as well. In a subset of the Cardiovascular Health Study undergoing cerebral MRI with follow up imaging 5 years later, new infarctions were correlated with decreased scores on both the Modified Mini-Mental State Examination and the Digit-Symbol Substitution test[103].

Brain autopsies identifying the source of SBI are infrequently done, as this requires uninterrupted serial sections of the basal ganglia, tracing the infarcted area proximally until an occlusion is found[98]. However, from the few published autopsy reports, the etiology of SBI appears to be heterogeneous, arising from obstructions of the small penetrating arteries by embolic particles of atheroma, or blockages due to lipohyalinosis, a hypertrophic adaption to hypertension that eventually occludes the arterial lumen[61]. In the absence of identifiable obstructions by post-mortem examination, thrombus is suspected, as blood clots lyse and disappear within days of an embolic event[98]. The thrombotic etiology of SBI is supported by a case-control analysis from the ARIC study, comparing coagulation factors in 196 SBI cases and 214 age- race- and sex matched controls without SBI[99]. The odds of SBI were nearly doubled when contrasting the highest with lowest tertiles of d-dimer and von Willebrand factor, suggesting an association with coagulation.

Hypercoagulability, thrombosis, and microvascular occlusion are known complications of sickle cell anemia, a Mendelian hemoglobinopathy characterized by misshapen red blood cells, vaso-occlusive crisis, acute chest syndrome, and pain. Even in childhood, SBI is commonly observed in patients with sickle cell anemia, prevalent in 27%

of patients under the age of 6[178], and 37% under the age of 14[179]. Though unaccompanied by neurological deficits, SBI has been associated with poor academic achievement in children with sickle cell anemia, and with future stroke[101]. The prevalence of SBI in adults with sickle cell anemia is less well established; however, in a cross sectional study of 149 participants with a mean age of 32 years, lacunes were detected in 13% of patients with sickle cell anemia, compared to 2% of age, sex, and education matched African American controls[102]. To date, no publications have examined heterozygous SCT as a genetic risk factor for SBI.

Even in heterozygous carriers, hemoglobin S is associated with hypercoagulability. Under conditions of exertion, dehydration, and high altitude, SCT erythrocytes are known to sickle and polymerize[11, 27]. The sickling deformation exposes phosphatidylserine on the cell membrane surface, facilitating the assembly of coagulation enzymatic complexes[29]. Laboratory assays of healthy individuals with SCT show elevated markers of coagulation (prothrombin fragment 1+2, thrombin-antithrombin complex, and d-dimer)[26], and epidemiological studies report twice the risk of pulmonary embolism and venous thrombosis[33, 34]. Though seemingly plausible, it is unknown if the hypercoagulability and greater prevalence of thromboembolism observed in individuals with SCT conveys a heightened risk of SBI.

In conclusion, additional research is required to draw conclusions relating SCT to SBI. Our estimates were too variable to be definitive, and the analysis was underpowered to detect even moderate effects.

Table 1: Baseline (1993-1995) demographics and clinical characteristics of African American participants in the ARIC brain MRI examinations

Characteristic	SCT (N=55) *Mean ± SD or N (%)*	HbAA (N=773) *Mean ± SD or N (%)*	*P-value*
Demographics			
Age (years)	61 ± 5	62 ± 5	0.6
Female	27 (49%)	512 (65%)	0.02
Medical History			
Current smoker	15 (27%)	139 (18%)	0.09
Hypertension	31 (56%)	484 (63%)	0.4
Hypercholesterolemia	8 (15%)	160 (21%)	0.3
Diabetes	15 (27%)	192 (25%)	0.7
Coronary heart disease	1 (2%)	33 (4%)	0.7
Body mass index (kg/m^2)	28 ± 4	30 ± 5	0.006

Table 2: Frequencies of SBI counts at baseline MRI (1993-1995), stratified by SCT status

Genotype	Normal Count = 0	Mild Count = 1	Moderate Count > 1
SCT	49 (89%)	3 (5%)	3 (5%)
HbAA	659 (85%)	74 (10%)	40 (5%)

*Cochrane-Maental-Haenzsel p = 0.3

Table 3: Multivariable regression models contrasting SCT with HbAA

Model Outcome and Adjustments	Odds Ratio (95% CI)
Prevalent SBI at baseline MRI	
Demographics	0.72 (0.30 – 1.76)
Demographics, clinical covariates	0.69 (0.28 – 1.70)
Demographics, clinical covariates, ancestry	0.71 (0.28 – 1.78)
Incident SBI at follow up MRI (≥1 new infarct)	
Demographics, elapsed time	1.35 (0.62 – 2.95)
Demographics, clinical covariates, elapsed time	1.42 (0.64 – 3.13)
Demographics, clinical covariates, ancestry, elapsed time	1.40 (0.63 – 3.12)

Acknowledgments

The authors thank the staff and participants of the ARIC study for their important contributions.

Funding Sources

The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HSN268201100006C, HHSN268201100007C, HHSN268201100008C, HSN268201100009C, HHSN268201100010C, HHSN268201100011C, and SN268201100012C). Genotyping of hemoglobin S and hemoglobin C was funded by grant 5K12HL087097. Support for ARIC exome chip genotyping was provided by Building on GWAS for NHLBI-diseases: the U.S. CHARGE consortium through the National Institutes of Health (NIH) American Recovery and Reinvestment Act of 2009 (ARRA) (5RC2HL102419). NSK was supported by 1UO1HL117659.

Disclosures

None.

DISCUSSION

1. Introduction

This is the first epidemiological study to examine associations between SCT and cerebral small vessel disease. In a previous analysis from the ARIC study, an increased risk of stroke was observed in participants with SCT, compared to those with homozygous hemoglobin A. The pathophysiology of stroke in SCT carriers is unclear, however. Though not considered a hematological disorder, the SCT phenotype is nonetheless associated with hypercoagulability, vasculopathy, and possibly hypoperfusion. In the general population, future stroke has been predicted by cerebral small vessel disease. Subclinical brain infarctions, blockages of the small cerebral vessels by atheroma, lipohyalinosis, and thrombus, are associated with hypercoagulability. Thus, it seems biologically plausible that individuals with SCT would have heightened risk of SBI, a potential etiologic pathway to stroke. In addition, white matter lesions, often indicative of axon demyelination and hypoperfusion, have been observed with greater frequency in pediatric SCT carriers. However, these findings have yet to be replicated in a large, epidemiological study.

2. Key Findings

In the present study, we observed no statistically significant associations between SCT and cerebral small vessel disease. There was a trend for greater WML prevalence and

severity among those with SCT, as well as a higher 11-year incidence of SBI; however, the estimates were imprecise and inconclusive.

Associations between SCT and WML prevalence were examined by ordinal regression. First, the odds of having a WML score of 1, 2, or ≥ 3 were contrasted with having a score of 0. Next, the odds of a WML score of ≥ 2 were contrasted to a score of 0 or 1. Finally, the odds of a WML ≥ 3 were contrasted with a score of 0, 1, or 2. The advantage of this analysis is that it minimizes cut-point bias, which may result from arbitrarily assigned WML prevalence definitions. Because none of these contrasts significantly differed from each other and the assumptions of proportional odds were upheld, the cumulative logit was exponentiated to estimate the summary odds ratio of WML associated with SCT. We observed a 20% greater odds of WML associated with SCT, which was non-significant.

The association between SCT and WML severity was assessed by linear regression. A multiple regression model was used, to examine the mean difference in severity scores between individuals with SCT and HbAA. Among those with SCT, the mean WML score was 0.2 points higher than those with HbAA; however, our analysis was only powered to detect a mean difference of 0.4 or greater. The progression of WML by the follow-up MRI was examined by 2 methods. First, a linear model was used, analyzing the mean increase in WML score associated with SCT. We also analyzed progression using a logistic model, by defining progression as a longitudinal increase in severity score of at least 2 points. Though more conservative, we chose at least 2 points over 1 point to avoid spurious progression, which may simply reflect inaccuracies of the qualitative scoring method. However, the analyses of WML progression were non-significant, regardless of methodological technique.

Baseline prevalence and longitudinal progression of SBI were both analyzed by logistic regression. SBI prevalence was considered one or more infarction detected by MRI, and progression of SBI was defined by an increase in number of infarctions between the first and second MRI examinations. At the baseline MRI, there were only 6 cases of SBI among the 56 SCT carriers, which severely limited the analysis. By the follow up exam, the SBI prevalence among those with SCT had increased from 11% to 31%; however, only 36 individuals with SCT returned for the second MRI. Compared to HbAA individuals, those with the SCT genotype had 30% lower odds of SBI at the baseline MRI. The confidence intervals were implausibly broad though, and the estimate was non-significant. By the follow up exam, participants with SCT had a 40% higher odds of newly detected SBI lesions. Again, the confidence intervals were broad and this analysis was non-significant and inconclusive.

With a sample size of only 56 SCT carriers imaged by cerebral MRI, analyses of modification by traditional stroke risk factors, sickle haplotype, co-inherited hemoglobinopathies, or stroke candidate genes did not appear feasible. However, we discovered that most potential genetic modifiers were highly imputable using the Affymetrix 6.0 microarray and 1000 Genomes reference panel. SNPs from candidate genes associated with stroke in children with sickle cell anemia, *ADCY9, BMP6, CSF2, ECE1, ERG, MET, SELP, TEK, TGFBR3,* and *TNF-α,* imputed well, each with a quality score greater than .90. Stroke candidate genes with genome-wide significance, *HDAC9, PITX2,* and *NINJ2*, also imputed with high quality. SNPs from two genes influencing gamma globin and fetal hemoglobin concentrations, *HBSIL-MYB* and *BCL11A*, imputed with high quality. The SNP for the most common alpha-thalassemia deletion in African Americans was not imputable to

the 1000 genomes panel, using either the Affymetrix array or Exome chip. Likewise, the SNP for the Senegal sickle haplotype was not imputable; however, several proxy SNPs identified by PLINK imputed with very high quality. This preliminary analysis provides proof of concept that potential genetic modification of association between SCT and cerebrovascular disease could be assessed, provided an adequately large sample size.

3. Public Health Implications

Since 2006, all 50 of the United States have mandated neonatal screenings for sickle hemoglobinopathies. Consequently, SCT status is now documented in the medical record of younger African Americans. With a heterogeneous allelic frequency of 7-9% in African Americans and 0.2% in non-Hispanic whites, SCT is estimated to affect over 3 million Americans. Increasingly, research suggests SCT may not be entirely benign, raising the possibility that the SCT genotype may be used to target high risk individuals for potential therapeutic interventions. However, at the current state of science, potential interventions – which may be behavioral, environmental, pharmacological, or genetic; and the ideal time to initiate them, are unknown.

4. Strengths and Limitations

The ARIC study includes a large sample of older African Americans with SCT identified by direct genotyping. A subset of study participants was randomly selected for cerebral MRI examination, with images read by trained neuroradiologists using standardized criteria. Despite these strengths, we were unable to conclusively examine the associations between SCT and cerebral small vessel disease. Of the 844 African Americans undergoing cerebral MRI, only 56 were SCT carriers. As a result, we did not have the statistical power

to detect even moderately higher risks of cerebral small vessel disease in those with SCT. The assessment of WML was additionally hindered by the qualitative severity scores, which were subject to measurement error and only fair inter-reader and intra-reader reliability.

5. Future Directions

Our analysis of SCT and cerebral small vessel disease is inconclusive. To gain a more definitive understanding of the cerebrovascular pathophysiology associated with SCT, an analysis with adequate statistical power is required. The statistical power would be improved with a larger sample of SCT carriers, or better precision of the image measurements. In order to augment the sample size, collaborations with other cohort studies examining African Americans by cerebral MRI could be forged. However, data harmonization is a potential concern, if SCT genotypes are identified using differing methods, or brain MRI is performed by scanners with different image resolution.

MRI scanner resolution has improved dramatically over the past 20 years, and by the time the ARIC cohort was examined at visit 5 (2011 – 2013), scanners had been upgraded from 1.5 to 3.0 Tesla magnets. The newer scanners not only have better resolution, but the capability to measure 3D volumes, obviating the need to score WML severity with a qualitative scale. Diffusion tensor imaging, another feature available in the visit 5 MRI, has been shown to enhance detection of small vessel disease and white matter abnormalities as well, which may improve precision of the outcome. Though currently unmeasured in the ARIC study, cerebral artery tortuosity and dolichoectasia could be assessed, using the visit 5 MRI images. This has previously been measured in other cohort studies, using custom-built

image analysis software. The association between SCT and brain volumes, or micro-hemorrhage could also be analyzed from the visit 5 MRI.

In the future, the effect of genetic modification on the association between SCT and cerebrovascular disease could be assessed as well. This would obviously require a much larger sample size than currently available in the ARIC study, but could be potentially interesting. Alpha thalassemia deletions are highly prevalent (30%) in African Americans, and known to modify the amount of hemoglobin S produced when co-inherited with the sickle mutation. It would be reasonable to hypothesize that alpha thalassemia would influence the associations between SCT and cerebrovascular disease. Modification by stroke candidate genes could also be an interesting area of further research.

Influences from the environment, lifestyle, and behavior affecting SCT and cerebrovascular disease risk could also be examined. This, too, would require a much larger sample size than currently available, but could be enormously beneficial if shown to influence associations between SCT and cerebrovascular disease. Unlike genetic modifiers, the environmental, lifestyle, and behavioral factors are largely modifiable. If an exposure such as smoking or hormone replacement therapy were shown to interact with SCT to cause heightened coagulation and cerebrovascular disease, strategies could be developed to mitigate risk. It is already known that high altitude, dehydration, and heavy exertion cause erythrocyte sickling in SCT carriers, and an assessment of how these exposures may influence an individual with SCT's likelihood of developing cerebrovascular disease could also be assessed.

APPENDIX 1

Figure 1aI. Determination of final study population

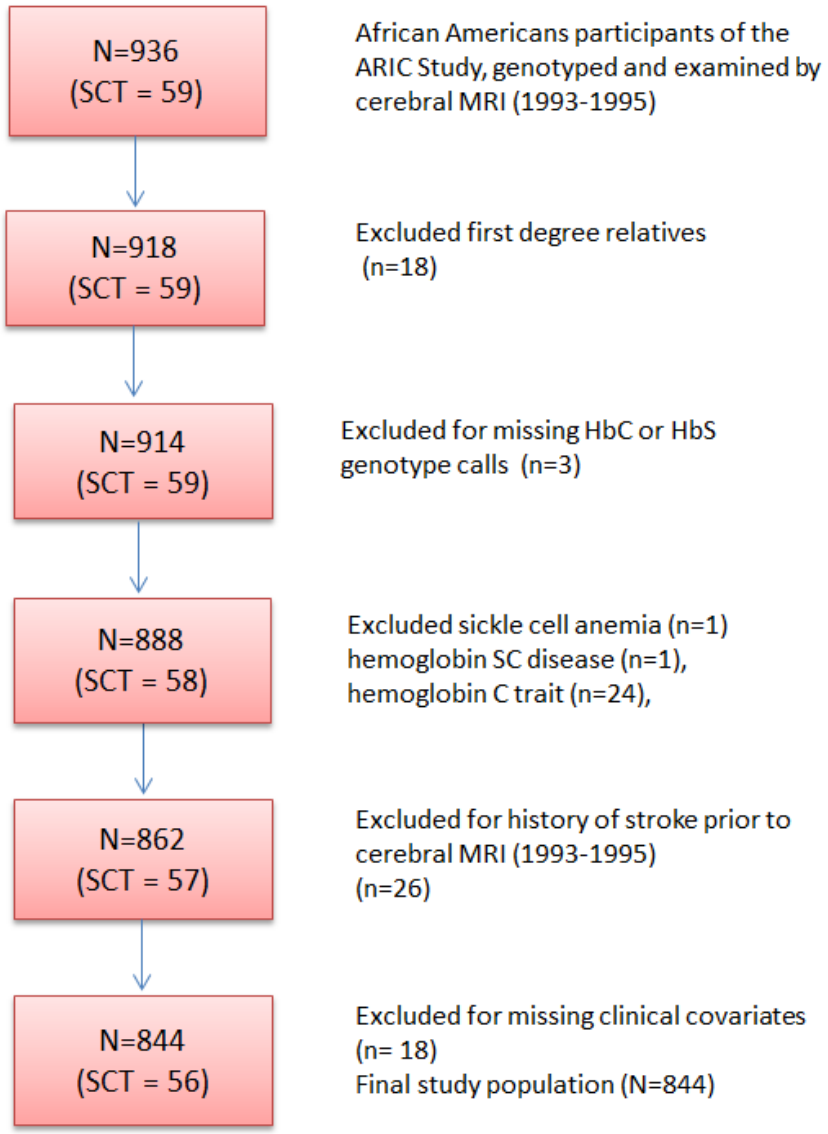
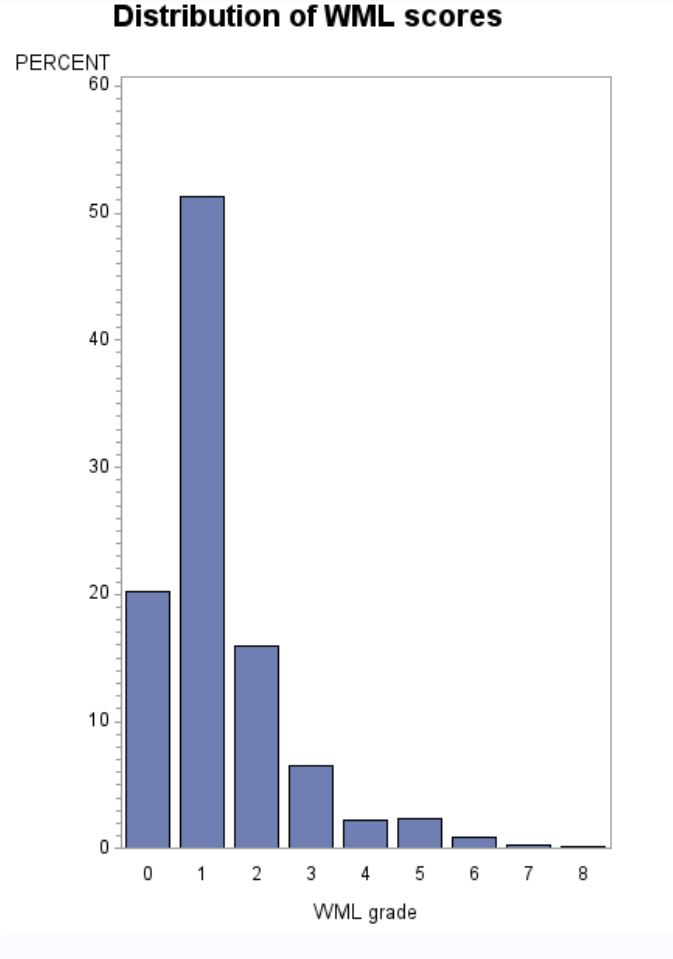


Figure 2aI. Distribution of white matter scores at baseline MRI (1993-1995)



The majority of study participants (80%) were identified with WML at the baseline MRI. Most (52%) had a mild presentation, classified as a WML score of 1. The distribution of scores was unimodal and skewed right (skewness = 1.8), with a kurtosis value of 4.3 and significant departure from normality (Kolmogov-Smirnov test $p < 0.01$).

Figure 3aI. Baseline MRI (1993-1995) white matter score distributions, stratified by SCT status.

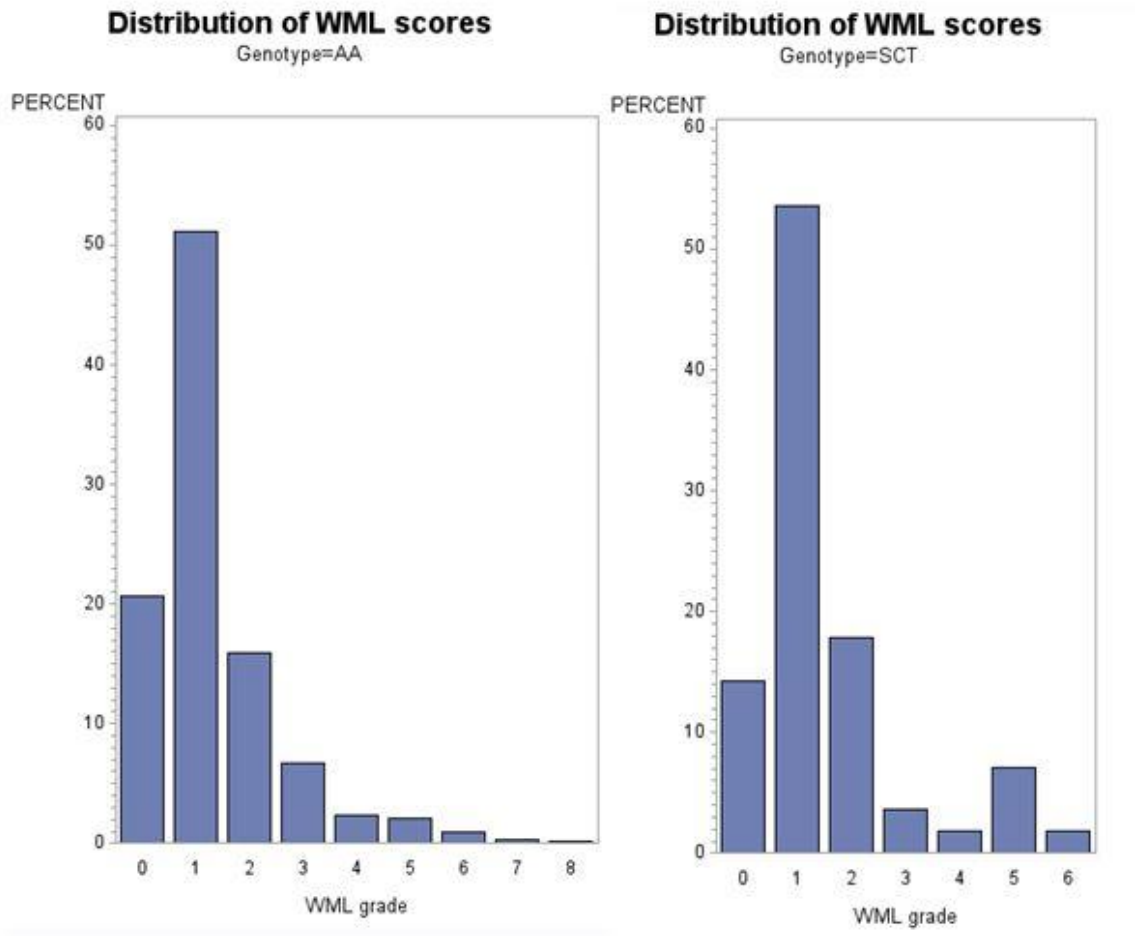
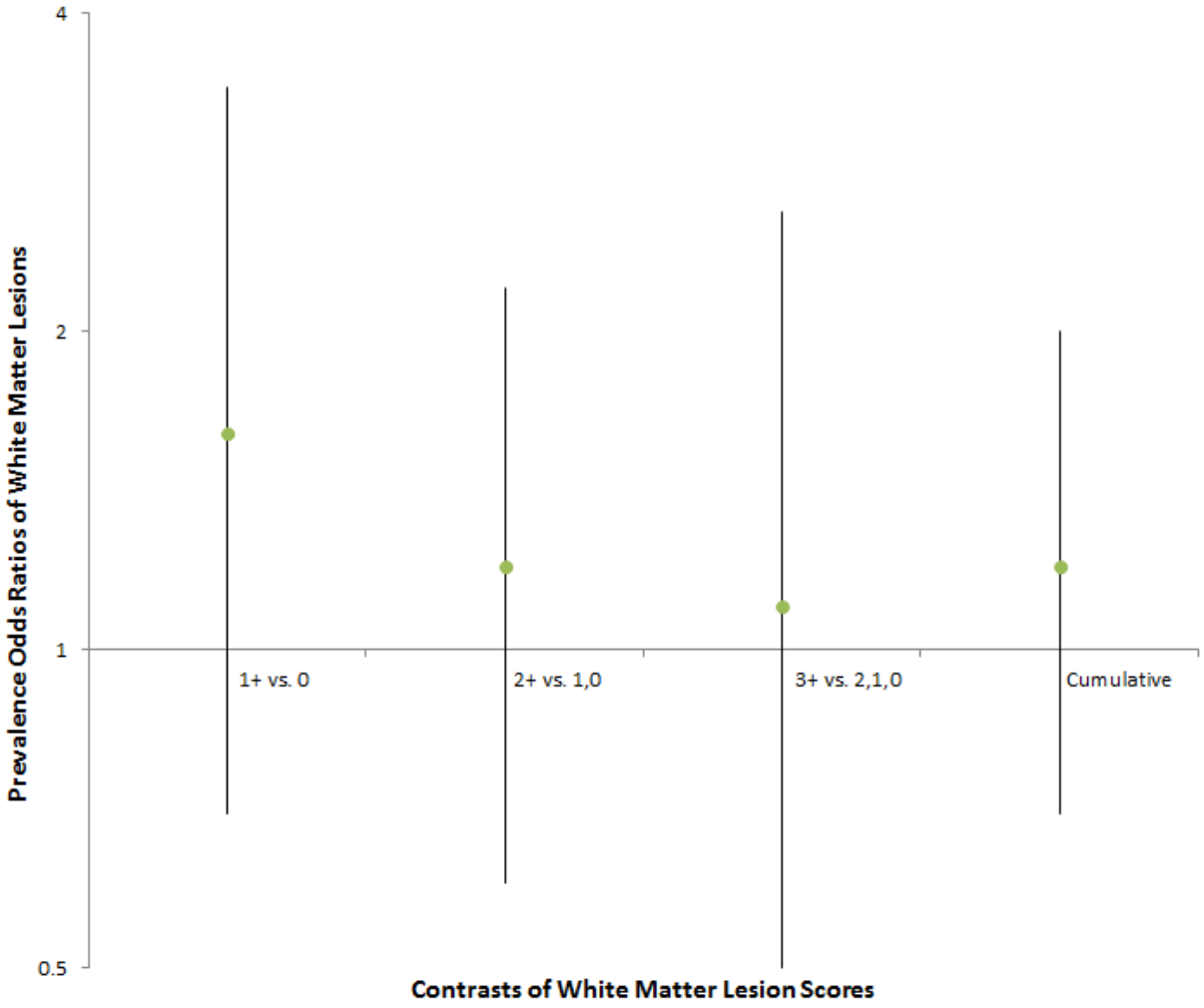


Table 1aI. Distribution of baseline WML scores, stratified by SCT status

	WML Scores (0-8)								
	0	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***
SCT	8 (14%)	30 (54%)	10 (18%)	2 (4%)	1 (2%)	4 (7%)	1 (2%)	0	0
HbAA	163 (21%)	403 (51%)	125 (16%)	53 (7%)	18 (2%)	16 (2%)	7 (1%)	2 (0.3%)	1 (0.1%)

Figure 4aI. Individual and cumulative prevalence odds ratios of WML associated with SCT. Contrasts = score of 1 or greater vs. score of 0; score of 2 or greater vs. score of 1 or 0; score of 3 or greater vs. score of 2, 1, or 0. All models adjusted for age, sex, hypertension, smoking, diabetes, and ancestry.



Proportional odds test: $\chi^2 = 39.9$ (32 degrees of freedom); p=0.2. There is insufficient evidence to reject the null hypothesis, which assumes the increment to the log odds associated with SCT is equivalent for each contrast. The cumulative prevalence odds ratio is therefore reported as a summary estimate.

Table 2a. Linear regression models analyzing absolute and relative increment to baseline WML score associated with SCT.

Model	Beta	p-value	Interpretation
WML mean score			SCT is insignificantly associated with a mean WML score that is 0.2 points higher than those with HbAA
Crude	0.2	0.2	
Demographics	0.2	0.1	
Demographics, Covariates, Ancestry	0.2	0.2	
Log of WML mean score			SCT is insignificantly associated with a mean WML score that is 9% higher, attenuating to a score 7% higher after full adjustments.
Crude	0.09	0.2	
Demographics	0.09	0.1	
Demographics, Covariates, Ancestry	0.07	0.2	

*Demographics = age, sex
Covariates = hypertension, diabetes, smoking

Justification for the absolute rather than relative model:

Linear models assume homoscedasticity and a linear relationship between the dependent variable and any continuous predictor variables. WML is a non-normally distributed variable; however, its distribution of residuals approximates a normal distribution reasonably well, demonstrating near-homoscedasticity (Figure 5a). The continuous variables (age and the 10 principal components) demonstrate linear relationships with WML score, shown by the nearly equal distribution of residuals both above and below the reference line (Figure 6a). Because risk is best measured on the additive scale, and the linear model demonstrates reasonable fit, the linear model without log transformation of the WML variable is preferred.

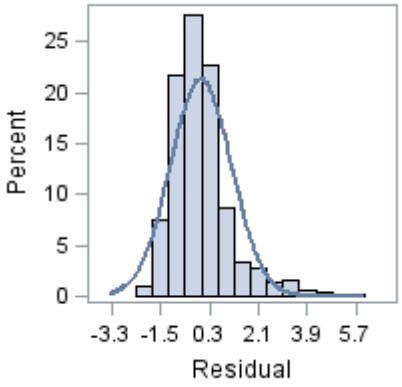
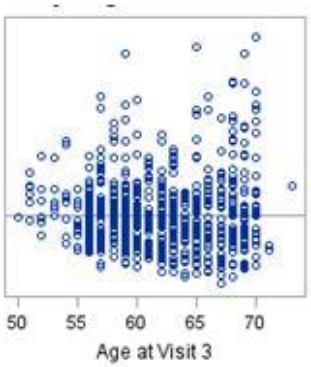
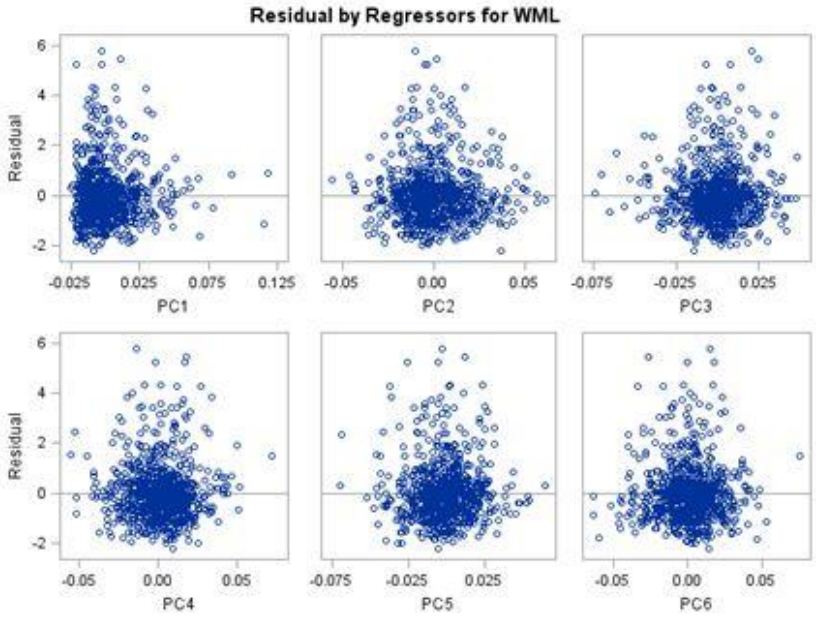


Figure 5aI. Near-homoscedastic distribution of WML residuals

Figure 6aI. Residuals for continuous variables (age and 10 principal components)





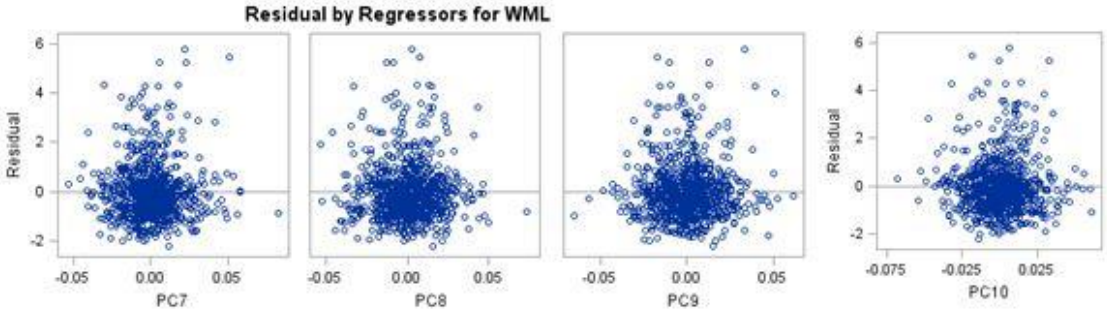
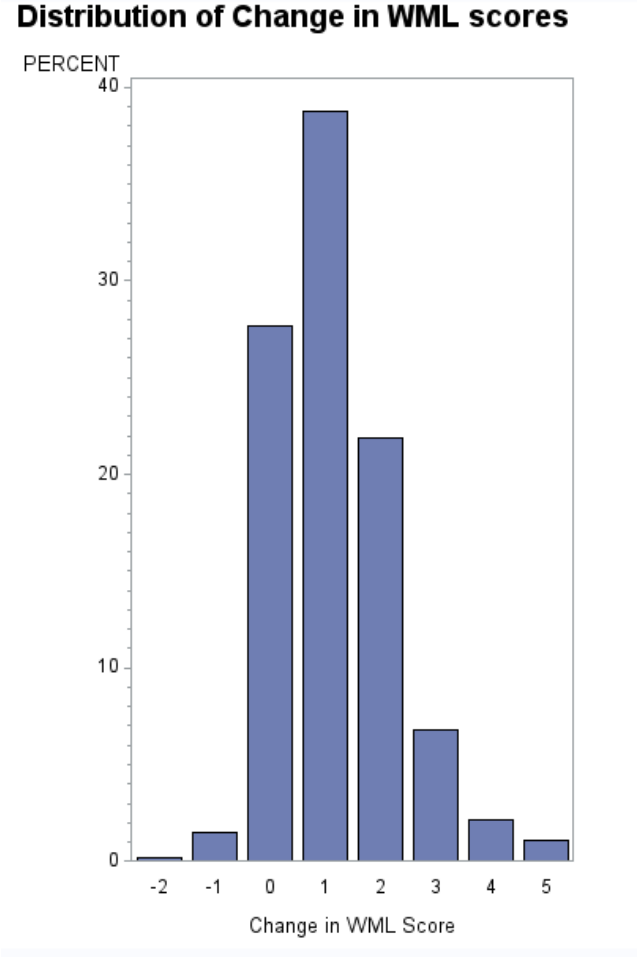


Table 3aI. Demographics and baseline clinical characteristics of study participants examined by cerebral MRI in 1993-1995, who either did or did not return for follow up imaging in 2004-2006.

Baseline Characteristics	Follow-up MRI (N=470) *Mean ± SD or N (%)*	No Follow-up MRI (N=377) *Mean ± SD or N (%)*	*P-value*
Demographics			
Age (years)	61 ± 4	62 ± 5	0.0001
Female	297 (63%)	245 (65%)	0.6
Medical History			
Sickle cell trait	36 (7.7%)	20 (5.3%)	0.2
Current smoker	74 (16%)	85 (23%)	0.01
Hypertension	271 (58%)	257 (68%)	0.002
Systolic blood pressure	131 ± 20	136 ± 22	0.0007
Diastolic blood pressure	75 ± 10	75 ± 11	0.9
Total cholesterol	208 ± 41	211 ± 39	0.3
Diabetes	98 (21%)	113 (30%)	0.002
Coronary heart disease	8 (2%)	26 (7%)	0.0001
Body mass index (kg/m^2)	29 ± 5	30 ± 6	0.05
Magnetic Resonance Imaging			
White matter lesions	360 (77%)	315 (84%)	0.01
Mean WML score	1.1 ± 1.0	1.6 ± 1.4	<0.0001

Study participants returning for the follow up MRI in 2004-2006 had fewer comorbid conditions at the baseline MRI in 1993-1995 than those not returning for the follow-up exam. In addition, white matter lesions detected at the baseline MRI were both less prevalent and less severe in those returning for follow-up imaging. The remaining population examined by cerebral MRI in 2004-2006 was likely influenced by selective survival.

Figure 7aI. Distribution of change in WML score between baseline (1993-1995) and follow-up (2004-2006) cerebral MRI examinations.



An increase in the WML severity score was observed in the majority (71%) of participants; however no change was observed in 27%, and a decrease in severity score was detected in 2%.

Figure 8aI. Distributions of change in WML score between baseline (1993-1995) and follow-up (2004-2006) cerebral MRI examinations, stratified by SCT status.

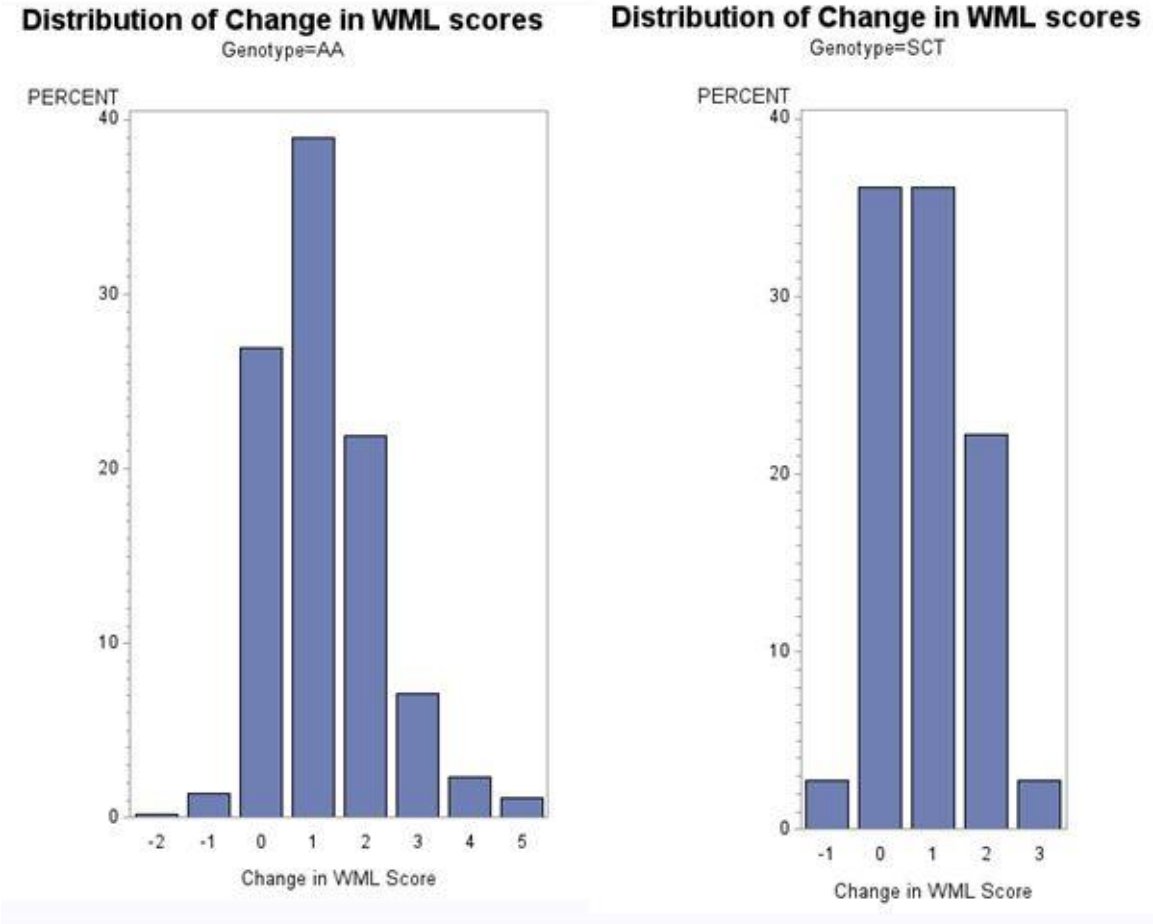


Table 4aI. Distribution of WML score change, stratified by SCT status

	-2	***-1***	***0***	***1***	***2***	***3***	***4***	***5***
SCT	0	1 (3%)	12 (35%)	12 (35%)	8 (25%)	1 (3%)	0	0
HbAA	1 (0.2%)	6 (1%)	115 (27%)	165 (39%)	88 (21%)	32 (8%)	10 (2%)	5 (1%)

APPENDIX II

Due to the small number of subclinical infarctions (n=6) observed in the 56 SCT cases imaged by cerebral MRI at the baseline examination, it was suggested that prevalence estimates should be examined at the follow up MRI exam. Because SBI is strongly associated with age, it was assumed that the greater number of lesions (in both individuals with SCT and HbAA) by the follow up examination would enhance the statistical precision and power to detect an association. Indeed, by the follow up examination, the SBI prevalence had increased in both groups (Table 1aII).

Table 1aII: Frequencies of SBI counts at baseline (1993-1995) and follow up (2004-2006) MRI examinations, stratified by SCT status

Genotype	Normal SBI Count = 0	Mild SBI Count = 1	Moderate SBI Count > 1	Total SBI Count
*MRI Exam 1**				
SCT	49 (89%)	3 (5%)	3 (5%)	6 (11%)
HbAA	659 (85%)	74 (10%)	40 (5%)	114 (15%)
MRI Exam 2†				
SCT	24 (69%)	8 (23%)	3 (9%)	11 (31%)
HbAA	315 (74%)	78 (18%)	30 (7%)	108 (26%)

*Cochrane-Maental-Haenzsel p = 0.6
†Cochrane-Maental-Haenzsel p = 0.5

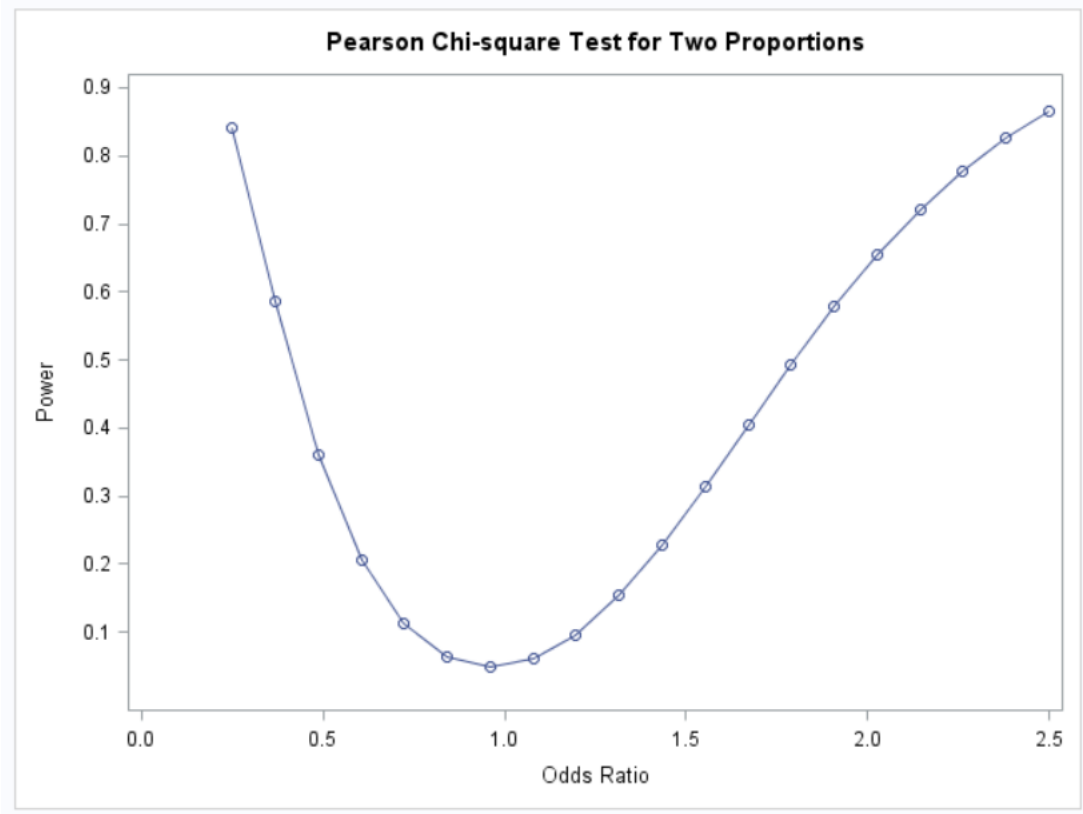
The prevalence of SBI in SCT carriers increased from 11% to 31% by the follow up examination, and from 15% to 26% in those with the HbAA genotype. However, the study population attending the follow up MRI was strikingly different from those either choosing not to, or unable, to participate in the follow up MRI (Table 2aII).

Table 2aII. Demographics and baseline clinical characteristics of study participants examined by cerebral MRI in 1993-1995, who either did or did not return for follow up imaging in 2004-2006.

Baseline Characteristics	Follow-up MRI (N=470) *Mean ± SD or N (%)*	No Follow-up MRI (N=377) *Mean ± SD or N (%)*	*P-value*
Demographics			
Age (years)	61 ± 4	62 ± 5	0.0001
Female	297 (63%)	245 (65%)	0.6
Medical History			
Sickle cell trait	36 (7.7%)	20 (5.3%)	0.2
Current smoker	74 (16%)	85 (23%)	0.01
Hypertension	271 (58%)	257 (68%)	0.002
Systolic blood pressure	131 ± 20	136 ± 22	0.0007
Diastolic blood pressure	75 ± 10	75 ± 11	0.9
Total cholesterol	208 ± 41	211 ± 39	0.3
Diabetes	98 (21%)	113 (30%)	0.002
Coronary heart disease	8 (2%)	26 (7%)	0.0001

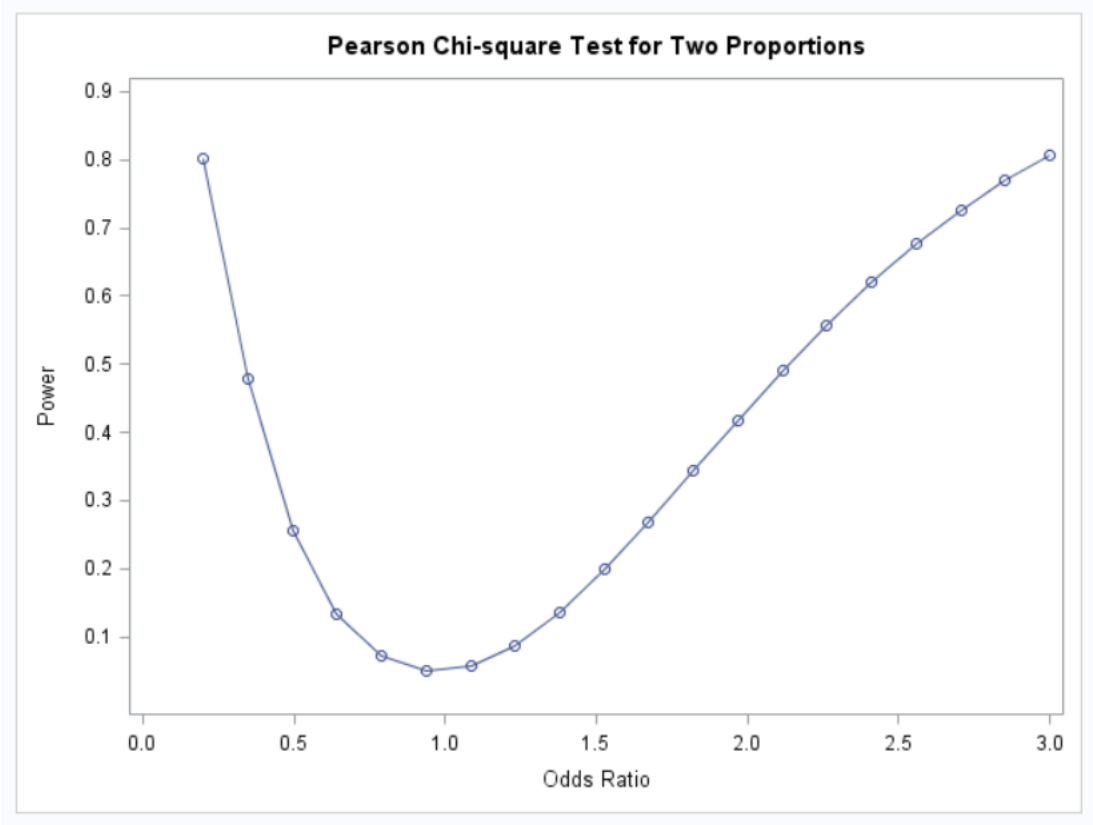
Moreover, the projected power to detect an association between SBI and SCT was not improved by the follow up, due to the overall decreased sample of study participants imaged at the second MRI (Figures 1aII and 2aII).

Figure 1aII: Power curves for odds ratios of SBI prevalence among those with SCT, compared to study participants with HbAA, at the first MRI: Assumed sample size = 800, assumed prevalence of SCT = 7%, assumed reference prevalence of SBI = 20%.



We expected 80% power to detect an odds ratio of 2.3 (if SCT is associated with increased risk of SBI), or an odds ratio of 0.3 (if SCT is associated with decreased risk of SBI).

Figure 2aII: Power curves for odds ratios of SBI prevalence among those with SCT, compared to study participants with HbAA, at the follow-up MRI: Assumed sample size = 450, assumed prevalence of SCT = 7%, assumed reference prevalence of SBI = 30%.



We expected 80% power to detect an odds ratio of 3.0 (if SCT is associated with increased risk of SBI), or an odds ratio of 0.2 (if SCT is associated with decreased risk of SBI).

Based on the power calculations, the insignificant prevalence odds ratios estimates for the baseline and follow-up MRI examinations were unsurprising (Table 3aII).

Table 3aII: Multivariable regression models of prevalent SBI, contrasting SCT with HbAA at the baseline (1993-1995) and follow-up (2004-2006) MRI examinations

Model Outcome and Adjustments	Odds Ratio (95% CI)
Prevalent SBI at baseline MRI	
Demographics	0.72 (0.30 – 1.76)
Demographics, clinical covariates	0.69 (0.28 – 1.70)
Demographics, clinical covariates, ancestry	0.71 (0.28 – 1.78)
Prevalent SBI at follow-up MRI	
Demographics	1.35 (0.63 – 2.87)
Demographics, clinical covariates	1.36 (0.63 – 2.93)
Demographics, clinical covariates, ancestry	1.40 (0.64 – 3.05)

REFERENCES

1. Steinberg M, Forget B, Higgs D, Nagel R, eds. *Disorders of Hemoglobin : Genetics, Pathophysiology, and Clinical Management.* New York: Cambridge University Press; 2009.

2. Tortora G, Grabowski S. *Principles of Anatomy and Physiology*. New York, NY: John Wiley & Sons, Inc.; 2000.

3. Liebhaber SA, Cash FE, Ballas SK. Human alpha-globin gene expression. The dominant role of the alpha 2-locus in mRNA and protein synthesis. *J Biol Chem*. 1986;261:15327-15333.

4. Tomkins J. The human beta-globin pseudogene is non-variable and functional. *Answers in Research Journal*. 2013;6:293.

5. Hawks J. *Reading a Genetic Map: The Beta-Globin Gene Cluster.* http://johnhawks.net/taxonomy/term/928: ; 2011.

6. Harris S, Barrie PA, Weiss ML, Jeffreys AJ. The primate psi beta 1 gene. An ancient beta-globin pseudogene. *J Mol Biol*. 1984;180:785-801.

7. Marotta CA, Wilson JT, Forget BG, Weissman SM. Human beta-globin messenger RNA. III. Nucleotide sequences derived from complementary DNA. *J Biol Chem*. 1977;252:5040-5053.

8. NEEL JV, WELLS IC, ITANO HA. Familial differences in the proportion of abnormal hemoglobin present in the sickle cell trait. *J Clin Invest*. 1951;30:1120-1124.

9. Chikezie PC. Sodium metabisulfite-induced polymerization of sickle cell hemoglobin incubated in the extracts of three medicinal plants (Anacardium occidentale, Psidium guajava, and Terminalia catappa). *Pharmacogn Mag*. 2011;7:126-132.

10. Pauling L, Itano H, Singer S, Wells I. Sickle cell anemia, a molecular disease. *Science*. 1949;110:543.

11. Martin TW, Weisman IM, Zeballos RJ, Stephenson SR. Exercise and hypoxia increase sickling in venous blood from an exercising limb in individuals with sickle cell trait. *Am J Med*. 1989;87:48-56.

12. Bookchin RM, Balazs T, Landau LC. Determinants of red cell sickling. Effects of varying pH and of increasing intracellular hemoglobin concentration by osmotic shrinkage. *J Lab Clin Med*. 1976;87:597-616.

13. LANGE RD, MINNICH V, MOORE CV. Effect of oxygen tension and of pH on the sickling and mechanical fragility of erythrocytes from patients with sickle cell anemia and the sickle cell trait. *J Lab Clin Med*. 1951;37:789-802.

14. Serjeant GR, Petch MC, Serjeant BE. The in vivo sickle phenomenon: a reappraisal. *J Lab Clin Med*. 1973;81:850-856.

15. Cholera R, Brittain NJ, Gillrie MR, et al. Impaired cytoadherence of Plasmodium falciparum-infected erythrocytes containing sickle hemoglobin. *Proc Natl Acad Sci U S A*. 2008;105:991-996.

16. Aidoo M, Terlouw DJ, Kolczak MS, et al. Protective effects of the sickle cell gene against malaria morbidity and mortality. *Lancet*. 2002;359:1311-1312.

17. Platt OS, Brambilla DJ, Rosse WF, et al. Mortality in sickle cell disease. Life expectancy and risk factors for early death. *N Engl J Med*. 1994;330:1639-1644.

18. Piel FB, Patil AP, Howes RE, et al. Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. *Lancet*. 2013;381:142-151.

19. Lorey FW, Arnopp J, Cunningham GC. Distribution of hemoglobinopathy variants by ethnicity in a multiethnic state. *Genet Epidemiol*. 1996;13:501-512.

20. Gabriel A, Przybylski J. Sickle-Cell Anemia: A Look at Global Haplotype Distribution. *Nature Education*. 2010;3:2.

21. Powars DR, Chan L, Schroeder WA. Beta S-gene-cluster haplotypes in sickle cell anemia: clinical implications. *Am J Pediatr Hematol Oncol*. 1990;12:367-374.

22. Hattori Y, Kutlar F, Kutlar A, McKie VC, Huisman TH. Haplotypes of beta S chromosomes among patients with sickle cell anemia from Georgia. *Hemoglobin*. 1986;10:623-642.

23. Schroeder WA, Powars DR, Kay LM, et al. Beta-cluster haplotypes, alpha-gene status, and hematological data from SS, SC, and S-beta-thalassemia patients in southern California. *Hemoglobin*. 1989;13:325-353.

24. Akinsheye I, Alsultan A, Solovieff N, et al. Fetal hemoglobin in sickle cell anemia. *Blood*. 2011;118:19-27.

25. Helley D, Girot R, Guillin MC, Bezeaud A. Sickle cell disease: relation between procoagulant activity of red blood cells from different phenotypes and in vivo blood coagulation activation. *Br J Haematol*. 1997;99:268-272.

26. Westerman MP, Green D, Gilman-Sachs A, et al. Coagulation changes in individuals with sickle cell trait. *Am J Hematol*. 2002;69:89-94.

27. Kark JA, Posey DM, Schumacher HR, Ruehle CJ. Sickle-cell trait as a risk factor for sudden death in physical training. *N Engl J Med*. 1987;317:781-787.

28. Thogmartin JR, Wilson CI, Palma NA, Ignacio SS, Shuman MJ, Flannagan LM. Sickle cell trait-associated deaths: a case series with a review of the literature. *J Forensic Sci*. 2011;56:1352-1360.

29. Singer ST, Ataga KI. Hypercoagulability in sickle cell disease and beta-thalassemia. *Curr Mol Med*. 2008;8:639-645.

30. Zwaal RF, Schroit AJ. Pathophysiologic implications of membrane phospholipid asymmetry in blood cells. *Blood*. 1997;89:1121-1132.

31. Lawrie AS, Pizzey A, Trompeter S, et al. Procoagulant activity in patients with sickle cell trait. *Blood Coagul Fibrinolysis*. 2012;23:268-270.

32. McCormick W. Abnormal Hemoglobins: II. the Pathology of Sickle Cell Trait. *American Journal of Medical Sciences*. 1961;241:329.

33. Heller P, Best WR, Nelson RB, Becktel J. Clinical implications of sickle-cell trait and glucose-6-phosphate dehydrogenase deficiency in hospitalized black male patients. *N Engl J Med*. 1979;300:1001-1005.

34. Austin H, Key NS, Benson JM, et al. Sickle cell trait and the risk of venous thromboembolism among blacks. *Blood*. 2007;110:908-912.

35. Austin H, Lally C, Benson JM, Whitsett C, Hooper WC, Key NS. Hormonal contraception, sickle cell trait, and risk for venous thromboembolism among African American women. *Am J Obstet Gynecol*. 2009;200:620.e1-620.e3.

36. Aird W. *Endothelial Biomedicine.* Cambridge, UK: Cambridge University Press; 2007.

37. Blum A, Yeganeh S, Peleg A, et al. Endothelial function in patients with sickle cell anemia during and after sickle cell crises. *J Thromb Thrombolysis*. 2005;19:83-86.

38. Zawar SD, Vyawahare MA, Nerkar M, Jawahirani AR. Non-invasive detection of endothelial dysfunction in sickle cell disease by Doppler ultrasonography. *J Assoc Physicians India*. 2005;53:677-680.

39. Ataga KI, Moore CG, Hillery CA, et al. Coagulation activation and inflammation in sickle cell disease-associated pulmonary hypertension. *Haematologica*. 2008;93:20-26.

40. De Caterina R, Libby P, Peng HB, et al. Nitric oxide decreases cytokine-induced endothelial activation. Nitric oxide selectively reduces endothelial expression of adhesion molecules and proinflammatory cytokines. *J Clin Invest*. 1995;96:60-68.

41. Presley TD, Perlegas AS, Bain LE, et al. Effects of a single sickling event on the mechanical fragility of sickle cell trait erythrocytes. *Hemoglobin*. 2010;34:24-36.

42. Kato GJ, Hebbel RP, Steinberg MH, Gladwin MT. Vasculopathy in sickle cell disease: Biology, pathophysiology, genetics, translational medicine, and new research directions. *Am J Hematol*. 2009;84:618-625.

43. Ahanchi SS, Tsihlis ND, Kibbe MR. The role of nitric oxide in the pathophysiology of intimal hyperplasia. *J Vasc Surg*. 2007;45 Suppl A:A64-73.

44. Majumdar S, Webb S, Norcross E, et al. Stroke with intracranial stenosis is associated with increased platelet activation in sickle cell anemia. *Pediatr Blood Cancer*. 2013;60:1192-1197.

45. Rothman SM, Fulling KH, Nelson JS. Sickle cell anemia and central nervous system infarction: a neuropathological study. *Ann Neurol*. 1986;20:684-690.

46. Haque AK, Gokhale S, Rampy BA, Adegboyega P, Duarte A, Saldana MJ. Pulmonary hypertension in sickle cell hemoglobinopathy: a clinicopathologic study of 20 cases. *Hum Pathol*. 2002;33:1037-1043.

47. Gutierrez J, Sacco RL, Wright CB. Dolichoectasia-an evolving arterial disease. *Nat Rev Neurol*. 2011;7:41-50.

48. Sho E, Nanjo H, Sho M, et al. Arterial enlargement, tortuosity, and intimal thickening in response to sequential exposure to high and low wall shear stress. *J Vasc Surg*. 2004;39:601-612.

49. Gibbons GH, Dzau VJ. The emerging concept of vascular remodeling. *N Engl J Med*. 1994;330:1431-1438.

50. Steen RG, Langston JW, Ogg RJ, Manci E, Mulhern RK, Wang W. Ectasia of the basilar artery in children with sickle cell disease: relationship to hematocrit and psychometric measures. *J Stroke Cerebrovasc Dis*. 1998;7:32-43.

51. Steen RG, Hankins GM, Xiong X, et al. Prospective brain imaging evaluation of children with sickle cell trait: initial observations. *Radiology*. 2003;228:208-215.

52. Steinberg MH, Embury SH. Alpha-thalassemia in blacks: genetic and clinical aspects and interactions with the sickle hemoglobin gene. *Blood*. 1986;68:985-990.

53. Sylvester-Jackson DA, Page SL, White JM, et al. Unbiased analysis of the frequency of beta-thalassemia point mutations in a population of African-American newborns. *Arch Pathol Lab Med*. 1993;117:1110-1114.

54. Gill FM, Sleeper LA, Weiner SJ, et al. Clinical events in the first decade in a cohort of infants with sickle cell disease. Cooperative Study of Sickle Cell Disease. *Blood*. 1995;86:776-783.

55. Sheehan VA, Luo Z, Flanagan JM, et al. Genetic modifiers of sickle cell anemia in the BABY HUG cohort: influence on laboratory and clinical phenotypes. *Am J Hematol*. 2013.

56. Bhardwaj U, Zhang YH, Lorey F, McCabe LL, McCabe ER. Molecular genetic confirmatory testing from newborn screening samples for the common African-American, Asian Indian, Southeast Asian, and Chinese beta-thalassemia mutations. *Am J Hematol*. 2005;78:249-255.

57. Lionnet F, Hammoudi N, Stojanovic KS, et al. Hemoglobin sickle cell disease complications: a clinical study of 179 cases. *Haematologica*. 2012;97:1136-1141.

58. Fauci A, Braunwald E, Kasper D, et al. *Harrison's Principles of Internal Medicine, 17th Edition.* New York, NY: McGraw-Hill Professional; 2008.

59. Go AS, Mozaffarian D, Roger VL, et al. Executive summary: heart disease and stroke statistics--2013 update: a report from the American Heart Association. *Circulation*. 2013;127:143-152.

60. Rosamond WD, Folsom AR, Chambless LE, et al. Stroke incidence and survival among middle-aged adults: 9-year follow-up of the Atherosclerosis Risk in Communities (ARIC) cohort. *Stroke*. 1999;30:736-743.

61. Fisher CM. Lacunar strokes and infarcts: a review. *Neurology*. 1982;32:871-876.

62. Sudlow CL, Warlow CP. Comparing stroke incidence worldwide: what makes studies comparable? *Stroke*. 1996;27:550-558.

63. Tirschwell DL, Longstreth WT,Jr. Validating administrative data in stroke research. *Stroke*. 2002;33:2465-2470.

64. Toole JF, Lefkowitz DS, Chambless LE, Wijnberg L, Paton CC, Heiss G. Self-reported transient ischemic attack and stroke symptoms: methods and baseline prevalence. The ARIC Study, 1987-1989. *Am J Epidemiol*. 1996;144:849-856.

65. Elmaoğlu M, Çelik A. *MRI Handbook: MR Physics, Patient Positioning, and Protocols.* New York: Springer; 2012.

66. Vermeer SE, Longstreth WT,Jr, Koudstaal PJ. Silent brain infarcts: a systematic review. *Lancet Neurol*. 2007;6:611-619.

67. Gottesman RF, Coresh J, Catellier DJ, et al. Blood pressure and white-matter disease progression in a biethnic cohort: Atherosclerosis Risk in Communities (ARIC) study. *Stroke*. 2010;41:3-8.

68. Bezerra DC, Sharrett AR, Matsushita K, et al. Risk factors for lacune subtypes in the Atherosclerosis Risk in Communities (ARIC) Study. *Neurology*. 2012;78:102-108.

69. Brilliant M, Hughes L, Anderson D, Ghobrial M, Elble R. Rarefied white matter in patients with Alzheimer disease. *Alzheimer Dis Assoc Disord*. 1995;9:39-46.

70. Revesz T, Hawkins CP, du Boulay EP, Barnard RO, McDonald WI. Pathological findings correlated with magnetic resonance imaging in subcortical arteriosclerotic encephalopathy (Binswanger's disease). *J Neurol Neurosurg Psychiatry*. 1989;52:1337-1344.

71. Bryan RN, Cai J, Burke G, et al. Prevalence and anatomic characteristics of infarct-like lesions on MR images of middle-aged adults: the atherosclerosis risk in communities study. *AJNR Am J Neuroradiol*. 1999;20:1273-1280.

72. Liao D, Cooper L, Cai J, et al. Presence and severity of cerebral white matter lesions and hypertension, its treatment, and its control. The ARIC Study. Atherosclerosis Risk in Communities Study. *Stroke*. 1996;27:2262-2270.

73. National Heart Lung and Blood Institute, The. *Stroke Belt Initiative: Projects Accomplishments and Lessons Learned.* http://www.nhlbi.nih.gov/health/prof/heart/other/sb_spec.pdf: .

74. Reeves MJ, Bushnell CD, Howard G, et al. Sex differences in stroke: epidemiology, clinical presentation, medical care, and outcomes. *Lancet Neurol*. 2008;7:915-926.

75. Brown DL, Boden-Albala B, Langa KM, et al. Projected costs of ischemic stroke in the United States. *Neurology*. 2006;67:1390-1395.

76. Centers for Disease Control and Prevention (CDC). Prevalence of stroke--United States, 2005. *MMWR Morb Mortal Wkly Rep*. 2007;56:469-474.

77. Roger VL, Go AS, Lloyd-Jones DM, et al. Heart disease and stroke statistics--2012 update: a report from the American Heart Association. *Circulation*. 2012;125:e2-e220.

78. Kim JS, Lee HS, Park HY, et al. Endothelial function in lacunar infarction: a comparison of lacunar infarction, cerebral atherosclerosis and control group. *Cerebrovasc Dis*. 2009;28:166-170.

79. Hassan A, Hunt BJ, O'Sullivan M, et al. Markers of endothelial dysfunction in lacunar infarction and ischaemic leukoaraiosis. *Brain*. 2003;126:424-432.

80. Ince B, Petty GW, Brown RD,Jr, Chu CP, Sicks JD, Whisnant JP. Dolichoectasia of the intracranial arteries in patients with first ischemic stroke: a population-based study. *Neurology*. 1998;50:1694-1698.

81. International Stroke Genetics Consortium (ISGC), Wellcome Trust Case Control Consortium 2 (WTCCC2), Bellenguez C, et al. Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. *Nat Genet*. 2012;44:328-333.

82. Traylor M, Farrall M, Holliday EG, et al. Genetic risk factors for ischaemic stroke and its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide association studies. *Lancet Neurol*. 2012;11:951-962.

83. Ikram MA, Seshadri S, Bis JC, et al. Genomewide association studies of stroke. *N Engl J Med*. 2009;360:1718-1728.

84. Ohene-Frempong K, Weiner SJ, Sleeper LA, et al. Cerebrovascular accidents in sickle cell disease: rates and risk factors. *Blood*. 1998;91:288-294.

85. Strouse JJ, Jordan LC, Lanzkron S, Casella JF. The excess burden of stroke in hospitalized adults with sickle cell disease. *Am J Hematol*. 2009;84:548-552.

86. Kissela B, Schneider A, Kleindorfer D, et al. Stroke in a biracial population: the excess burden of stroke among blacks. *Stroke*. 2004;35:426-431.

87. Adams R, McKie V, Nichols F, et al. The use of transcranial ultrasonography to predict stroke in sickle cell disease. *N Engl J Med*. 1992;326:605-610.

88. Driscoll MC, Hurlet A, Styles L, et al. Stroke risk in siblings with sickle cell anemia. *Blood*. 2003;101:2401-2404.

89. Flanagan JM, Frohlich DM, Howard TA, et al. Genetic predictors for stroke in children with sickle cell anemia. *Blood*. 2011;117:6681-6684.

90. Adams RJ, Kutlar A, McKie V, et al. Alpha thalassemia and stroke risk in sickle cell anemia. *Am J Hematol*. 1994;45:279-282.

91. Kennedy AP, Walsh DA, Nicholson R, Adams JG,3rd, Steinberg MH. Influence of HbS levels upon the hematological and clinical characteristics of sickle cell trait. *Am J Hematol*. 1986;22:51-54.

92. Sebastiani P, Ramoni MF, Nolan V, Baldwin CT, Steinberg MH. Genetic dissection and prognostic modeling of overt stroke in sickle cell anemia. *Nat Genet*. 2005;37:435-440.

93. Hoppe C, Klitz W, D'Harlingue K, et al. Confirmation of an association between the TNF(-308) promoter polymorphism and stroke risk in children with sickle cell anemia. *Stroke*. 2007;38:2241-2246.

94. Hoppe C, Klitz W, Cheng S, et al. Gene interactions and stroke risk in children with sickle cell anemia. *Blood*. 2004;103:2391-2396.

95. Kelly-Hayes M, Beiser A, Kase CS, Scaramucci A, D'Agostino RB, Wolf PA. The influence of gender and age on disability following ischemic stroke: the Framingham study. *J Stroke Cerebrovasc Dis*. 2003;12:119-126.

96. Centers for Disease Control and Prevention, The. *Differences in Disability among Black and White Stroke Survivors --- United States, 2000--2001 MMWR 54(01); 3-6.* http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5401a2.htm: ; 2005.

97. Kochanek K, Xu J, Murphy S, Minino A, Kung H. Deaths: Final data for 2009. *Natl Vital Stat Rep*. 2011;60.

98. Fisher CM. Capsular infarcts: the underlying vascular lesions. *Arch Neurol*. 1979;36:65-73.

99. Gottesman RF, Cummiskey C, Chambless L, et al. Hemostatic factors and subclinical brain infarction in a community-based sample: the ARIC study. *Cerebrovasc Dis*. 2009;28:589-594.

100. Bryan RN, Wells SW, Miller TJ, et al. Infarctlike lesions in the brain: prevalence and anatomic characteristics at MR imaging of the elderly--data from the Cardiovascular Health Study. *Radiology*. 1997;202:47-54.

101. DeBaun MR, Armstrong FD, McKinstry RC, Ware RE, Vichinsky E, Kirkham FJ. Silent cerebral infarcts: a review on a prevalent and progressive cause of neurologic injury in sickle cell anemia. *Blood*. 2012;119:4587-4596.

102. Vichinsky EP, Neumayr LD, Gold JI, et al. Neuropsychological dysfunction and neuroimaging abnormalities in neurologically intact adults with sickle cell anemia. *JAMA*. 2010;303:1823-1831.

103. Longstreth WT,Jr, Dulberg C, Manolio TA, et al. Incidence, manifestations, and predictors of brain infarcts defined by serial cranial magnetic resonance imaging in the elderly: the Cardiovascular Health Study. *Stroke*. 2002;33:2376-2382.

104. Vermeer SE, Prins ND, den Heijer T, Hofman A, Koudstaal PJ, Breteler MM. Silent brain infarcts and the risk of dementia and cognitive decline. *N Engl J Med*. 2003;348:1215-1222.

105. Bernick C, Kuller L, Dulberg C, et al. Silent MRI infarcts and the risk of future stroke: the cardiovascular health study. *Neurology*. 2001;57:1222-1229.

106. Moody DM, Thore CR, Anstrom JA, Challa VR, Langefeld CD, Brown WR. Quantification of afferent vessels shows reduced brain vascular density in subjects with leukoaraiosis. *Radiology*. 2004;233:883-890.

107. Pantoni L, Garcia JH. Pathogenesis of leukoaraiosis: a review. *Stroke*. 1997;28:652-659.

108. Ovbiagele B, Saver JL. Cerebral white matter hyperintensities on MRI: Current concepts and therapeutic implications. *Cerebrovasc Dis*. 2006;22:83-90.

109. Kurumatani T, Kudo T, Ikura Y, Takeda M. White matter changes in the gerbil brain under chronic cerebral hypoperfusion. *Stroke*. 1998;29:1058-1062.

110. Pantoni L, Garcia JH, Gutierrez JA. Cerebral white matter is highly vulnerable to ischemia. *Stroke*. 1996;27:1641-6; discussion 1647.

111. Liao D, Cooper L, Cai J, et al. The prevalence and severity of white matter lesions, their relationship with age, ethnicity, gender, and cardiovascular disease risk factors: the ARIC Study. *Neuroepidemiology*. 1997;16:149-162.

112. Manolio TA, Kronmal RA, Burke GL, et al. Magnetic resonance abnormalities and cardiovascular disease in older adults. The Cardiovascular Health Study. *Stroke*. 1994;25:318-327.

113. Knopman DS, Penman AD, Catellier DJ, et al. Vascular risk factors and longitudinal changes on brain MRI: the ARIC study. *Neurology*. 2011;76:1879-1885.

114. LADIS Study Group. 2001-2011: a decade of the LADIS (Leukoaraiosis And DISability) Study: what have we learned about white matter changes and small-vessel disease? *Cerebrovasc Dis*. 2011;32:577-588.

115. Jokinen H, Gouw AA, Madureira S, et al. Incident lacunes influence cognitive decline: the LADIS study. *Neurology*. 2011;76:1872-1878.

116. Longstreth WT,Jr, Arnold AM, Beauchamp NJ,Jr, et al. Incidence, manifestations, and predictors of worsening white matter on serial cranial magnetic resonance imaging in the elderly: the Cardiovascular Health Study. *Stroke*. 2005;36:56-61.

117. Longstreth WT,Jr, Arnold AM, Kuller LH, et al. Progression of magnetic resonance imaging-defined brain vascular disease predicts vascular events in elderly: the Cardiovascular Health Study. *Stroke*. 2011;42:2970-2972.

118. THOMPSON RK, WAGNER JA, MacLEOD CM. Sickle cell disease; report of a case with cerebral manifestations in the absence of anemia. *Ann Intern Med*. 1948;29:921-928.

119. DIGGS LW, JONES RS. Clinicopathological conference: sickle cell disease; atypical case. *Am J Clin Pathol*. 1952;22:1194-1200.

120. ENDE N, PIZZOLATO P, ZISKIND J. Sicklemia. *Ann Intern Med*. 1955;42:1065-1075.

121. SCHENK EA. Sickle Cell Trait and Superior Longitudinal Sinus Thrombosis. *Ann Intern Med*. 1964;60:465-470.

122. Dalal FY, Schmidt GB, Bennett EJ, Ramamurthy S. Sickle-cell trait. A report of a postoperative neurological complication. *Br J Anaesth*. 1974;46:387-388.

123. Handler CE, Perkin GD. Sickle cell trait and multiple cerebral infarctions. *J R Soc Med*. 1982;75:550-553.

124. Greenberg J, Massey EW. Cerebral infarction in sickle cell trait. *Ann Neurol*. 1985;18:354-355.

125. Feldenzer JA, Bueche MJ, Venes JL, Gebarski SS. Superior sagittal sinus thrombosis with infarction in sickle cell trait. *Stroke*. 1987;18:656-660.

126. Reyes MG. Subcortical cerebral infarctions in sickle cell trait. *J Neurol Neurosurg Psychiatry*. 1989;52:516-518.

127. Radhakrishnan K, Thacker AK, Maloo JC, el-Mangoush MA. Sickle cell trait and stroke in the young adult. *Postgrad Med J*. 1990;66:1078-1080.

128. Riggs JE, Ketonen LM, Wang DD, Valanne LK. Cerebral infarction in a child with sickle cell trait. *J Child Neurol*. 1995;10:253-254.

129. Gines R, Truemper E. Stroke in a 9-year-old with sickle cell trait. *Am J Emerg Med*. 1998;16:104-105.

130. Roach ES. Sickle cell trait: innocent until proven guilty. *Arch Neurol*. 2005;62:1781-1782.

131. Dowling MM. Sickle cell trait is not a risk factor for stroke. *Arch Neurol*. 2005;62:1780-1781.

132. Lannuzel A, Salmon V, Mevel G, Malpote E, Rabier R, Caparros-Lefebvre D. Epidemiology of stroke in Guadeloupe and role of sickle cell trait. *Rev Neurol (Paris)*. 1999;155:351-356.

133. Golomb MR. Sickle cell trait is a risk factor for early stroke. *Arch Neurol*. 2005;62:1778-1779.

134. Bucknor MD, Goo JS, Coppolino ML. The Risk of Potential Thromboembolic, Renal and Cardiac Complications of Sickle Cell Trait. *Hemoglobin*. 2013.

135. Caughey MC, Loehr LR, Key NS, et al. Sickle Cell Trait and Incident Ischemic Stroke in the Atherosclerosis Risk in Communities Study. *Stroke*. 2014.

136. United States Census Bureau, The. *State & County Quick Facts.* http://quickfacts.census.gov/qfd/states/00000.html: U.S. Department of Commerce; 2012.

137. Consensus conference. Newborn screening for sickle cell disease and other hemoglobinopathies. *JAMA*. 1987;258:1205-1209.

138. Benson JM, Therrell BL,Jr. History and current status of newborn screening for hemoglobinopathies. *Semin Perinatol*. 2010;34:134-144.

139. Jordan LB, Smith-Whitley K, Treadwell MJ, Telfair J, Grant AM, Ohene-Frempong K. Screening U.S. college athletes for their sickle cell disease carrier status. *Am J Prev Med*. 2011;41:S406-12.

140. O'Connor FG, Bergeron MF, Cantrell J, et al. ACSM and CHAMP summit on sickle cell trait: mitigating risks for warfighters and athletes. *Med Sci Sports Exerc*. 2012;44:2045-2056.

141. Jackson R, Chambless LE, Yang K, et al. Differences between respondents and nonrespondents in a multicenter community-based study vary by gender ethnicity. The Atherosclerosis Risk in Communities (ARIC) Study Investigators. *J Clin Epidemiol*. 1996;49:1441-1446.

142. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol*. 1989;129:687-702.

143. Schleinitz D, DiStefano J, Kovacs P. Targeted SNP genotyping using the TaqMan assay. In: DiStefano J, ed. *Disease Gene Identification: Methods and Protocols.* New York, NY: Humana Press; 2011:77.

144. Hosking L, Lumsden S, Lewis K, et al. Detection of genotyping errors by Hardy-Weinberg equilibrium testing. *Eur J Hum Genet*. 2004;12:395-399.

145. Hardy GH. Mendelian Proportions in a Mixed Population. *Science*. 1908;28:49-50.

146. *How Affymetrix GeneChip DNA Microarrays Work.* http://public.tgen.org/tgen.org/downloads/autism/Genotypingessentials.pdf: Affymetrix.

147. Clarke L, Zheng-Bradley X, Smith R, et al. The 1000 Genomes Project: data management and community access. *Nat Methods*. 2012;9:459-462.

148. Iles M. Genome-wide association studies. In: Teare D, ed. *Genetic Epidemiology.* New York, NY: Humana Press; 2011:89.

149. Chanda P, Yuhki N, Li M, et al. Comprehensive evaluation of imputation performance in African Americans. *J Hum Genet*. 2012;57:411-421.

150. Liu EY, Buyske S, Aragaki AK, et al. Genotype imputation of Metabochip SNPs using a study-specific reference panel of ~4,000 haplotypes in African Americans from the Women's Health Initiative. *Genet Epidemiol*. 2012;36:107-117.

151. Karanjia PN, Nelson JJ, Lefkowitz DS, et al. Validation of the ACAS TIA/stroke algorithm. *Neurology*. 1997;48:346-351.

152. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38:904-909.

153. Patrick DL, Cheadle A, Thompson DC, Diehr P, Koepsell T, Kinne S. The validity of self-reported smoking: a review and meta-analysis. *Am J Public Health*. 1994;84:1086-1093.

154. Brigham J, Lessov-Schlaggar CN, Javitz HS, Krasnow RE, McElroy M, Swan GE. Test-retest reliability of web-based retrospective self-report of tobacco exposure and risk. *J Med Internet Res*. 2009;11:e35.

155. Chobanian AV, Bakris GL, Black HR, et al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *JAMA*. 2003;289:2560-2572.

156. National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III) final report. *Circulation*. 2002;106:3143-3421.

157. American Diabetes Association. Standards of medical care in diabetes--2012. *Diabetes Care*. 2012;35 Suppl 1:S11-63.

158. Blackburn H. Classification of the electrocardiogram for population studies: Minnesota Code. *J Electrocardiol*. 1969;2:305-310.

159. Kottke TE, Daida H, Bailey KR, Hammill SC, Crow RS. Agreement and coding reliability of the Minnesota and Mayo electrocardiographic coding systems. *J Electrocardiol*. 1998;31:303-312.

160. Kors JA, van Herpen G, Wu J, Zhang Z, Prineas RJ, van Bemmel JH. Validation of a new computer program for Minnesota coding. *J Electrocardiol*. 1996;29 Suppl:83-88.

161. Twisk J. *Applied Longitudinal Data Analysis for Epidemiology.* New York: Cambridge University Press; 2003.

162. Blomqvist N. On the relation between change and initial value. *Journal of the American Statistical Association*. 1977;72:746.

163. Inzitari M, Studenski S, Rosano C, et al. Anemia is associated with the progression of white matter disease in older adults with high blood pressure: the cardiovascular health study. *J Am Geriatr Soc*. 2008;56:1867-1872.

164. Long S. *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks, CA: SAGE Publications, Inc; 1997.

165. Stokes M, Davis C, Koch G. *Categorical Data Analysis using SAS: Third Edition.* Cary, NC: SAS Institute; 2012.

166. Cole S. Lecture 2: Comparing groups in time: Cox. *EPID 722.* 2013.

167. Allison P. *Survival Analysis using SAS: A Practical Guide.* Second Edition ed. Cary, NC: SAS Institute; 2010.

168. Harrell FE,Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med*. 1996;15:361-387.

169. Greenberg J, Massey EW. Cerebral infarction in sickle cell trait. *Ann Neurol*. 1985;18:354-355.

170. Radhakrishnan K, Thacker AK, Maloo JC, el-Mangoush MA. Sickle cell trait and stroke in the young adult. *Postgrad Med J*. 1990;66:1078-1080.

171. Grove ML, Yu B, Cochran BJ, et al. Best practices and joint calling of the HumanExome BeadChip: the CHARGE Consortium. *PLoS One*. 2013;8:e68095.

172. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81:559-575.

173. Carmeliet P. Mechanisms of angiogenesis and arteriogenesis. *Nat Med*. 2000;6:389-395.

174. Connes P, Sara F, Hardy-Dessources MD, et al. Effects of short supramaximal exercise on hemorheology in sickle cell trait carriers. *Eur J Appl Physiol*. 2006;97:143-150.

175. Vincent L, Feasson L, Oyono-Enguelle S, et al. Remodeling of skeletal muscle microvasculature in sickle cell trait and alpha-thalassemia. *Am J Physiol Heart Circ Physiol*. 2010;298:H375-84.

176. Scantlebury N, Mabbott D, Janzen L, et al. White matter integrity and core cognitive function in children diagnosed with sickle cell disease. *J Pediatr Hematol Oncol*. 2011;33:163-171.

177. Sun B, Brown RC, Hayes L, et al. White matter damage in asymptomatic patients with sickle cell anemia: screening with diffusion tensor imaging. *AJNR Am J Neuroradiol*. 2012;33:2043-2049.

178. Kwiatkowski JL, Zimmerman RA, Pollock AN, et al. Silent infarcts in young children with sickle cell disease. *Br J Haematol*. 2009;146:300-305.

179. Bernaudin F, Verlhac S, Arnaud C, et al. Impact of early transcranial Doppler screening and intensive therapy on cerebral vasculopathy outcome in a newborn sickle cell anemia cohort. *Blood*. 2011;117:1130-40; quiz 1436.

180. Lorey FW, Arnopp J, Cunningham GC. Distribution of hemoglobinopathy variants by ethnicity in a multiethnic state. *Genet Epidemiol*. 1996;13:501-512.

181. Westerman MP, Green D, Gilman-Sachs A, et al. Coagulation changes in individuals with sickle cell trait. *Am J Hematol*. 2002;69:89-94.

182. Portnoy BA, Herion JC. Neurological manifestations in sickle-cell disease, with a review of the literature and emphasis on the prevalence of hemiplegia. *Ann Intern Med*. 1972;76:643-652.

183. Strouse JJ, Lanzkron S, Urrutia V. The epidemiology, evaluation and treatment of stroke in adults with sickle cell disease. *Expert Rev Hematol*. 2011;4:597-606.

184. Mosley TH,Jr, Knopman DS, Catellier DJ, et al. Cerebral MRI findings and cognitive functioning: the Atherosclerosis Risk in Communities study. *Neurology*. 2005;64:2056-2062.