CROSS-SECTIONAL ANALYSIS OF CANDIDATE GENES FOR COGNITIVE IMPAIRMENT IN THE OSTEOPOROTIC FRACTURES IN MEN STUDY

By

Jenna M. Gribbin

Presented to the Department of Public Health & Preventative Medicine
and the Oregon Health & Science University
School of Medicine
in partial fulfillment of
the requirements for the degree of

Master of Public Health

February 2013

Department of Public Health and Preventative Medicine

School of Medicine

Oregon Health & Science University

CERTIFICATE OF APPROVAL

This is to certify that the Master's thesis of

Jenna M. Gribbin

has been approved

Mentor/Advisor

Member

Member

TABLE OF CONTENTS

Figures and Tables

Acknowledgements

I am very grateful to my thesis committee for their support and guidance. My chair, Carrie Neilson, PhD, MPH helped me through each step of this process and was critical in pulling the whole project together. Jodi Lapidus, PhD helped me understand relevant statistical concepts in great depth through her ability to clearly explain even the most complex topics. Gene Bowman, ND, MPH worked with me to develop the idea behind the thesis, and also to learn all of the dense background material required to work with this topic. I also want to thank my friends and family for their support, especially my fiancé, Jeff Donovan.

Abstract

Background: Cognitive deficits are an early feature of Alzheimer's disease (AD) and vascular dementia (VD). Brain white matter hyperintensities (WMH), identified with MRI, are thought to represent cerebrovascular pathology and are a risk factor for cognitive decline. Six genes in the 17q25 region were recently identified as being associated with WMH burden. I aimed to test if these genes are associated with cognitive dysfunction in elderly men.

Methods: This study was conducted in Osteoporotic Fractures in Men Study (MrOS), a cohort of community-dwelling me age 65 and older. This was a cross-section study comparing 26 SNPs in six genes, *TRIM65, TRIM47, FBF1, MRPL38, ACOX1* and *WBP2,* with outcomes of two cognitive tests, Trails Making Test Part B (Trails B) and the Modified Mini Mental State Examination (3MS).

Results: Mean Trails B and 3MS scores were determined to be 129.5±56.3 seconds and 94.1±5.0 points respectively. SNPs in *ACOX1, FBF1,* and *MRPL38* were associated with Trails B and 3MS scores at $p<0.05$ levels. However, after correction for multiple comparisons, only one SNP (rs11651351) in *ACOX1* was found to be significantly associated with 3MS in a recessive effects model ($\beta=6.55$, $p<0.0001$).

Conclusions: White matter burden has been linked to decreased cognitive function as well as to the six genes studied here. So, the relationships observed in this study are plausible, but warrant further study in a larger population to verify the relationships.

INTRODUCTION

Cognitive decline is characterized by impaired memory, loss of language skills, impaired judgment, and personality changes among other symptoms. Cognitive deficits are an early feature of Alzheimer's disease (AD) and vascular dementia (VD). This study examines the genetic basis of cognitive function in a large cohort of older men. The clinical characteristics, proposed mechanisms and epidemiology of AD and VD will be introduced, then followed by a description of risk factors already shown to be associated with brain aging, as well as the rational for the candidate genes included in this study.

Research Question

Brain white matter hyperintensities (WMH), identified with MRI, are thought to represent cerebrovascular pathology and are a risk factor for cognitive decline.[1] Recently, six genes were found to be associated with WMH burden.[2] WMH accumulation is associated with disturbances in executive skills, which likely occur earlier than its impact on more global cognitive functions.[3,4] Trail Making Test Part B (Trails B) and the modified mini-mental (3MS) reflect executive skills and global cognition, respectively.[5,6] Thus, we tested the hypothesis that single nucleotide polymorphisms (SNPs) in these genes previously linked to WMH are also associated with the cognitive phenotype driven predominately by the WMH burden.

Cognitive Impairment Overview

Cognitive decline is a substantial burden, with 5.4 million Americans afflicted in 2012, 5.2 million of whom are over the age of 65.[7] Age is a primary risk factor, so as people live to older ages, disease burden increases. The two most common forms of age-related cognitive impairment are AD and VD. AD was characterized by Alois Alzheimer

in 1906 as he observed the hallmark plaques and tangles in a woman at autopsy whom had developed a dementia. VD is caused by strokes, or other vascular risk factors (i.e., hypertension, diabetes), which ultimately disturb blood flow to the brain. Both, AD and VD, however, can appear phenotypically similar, with symptoms including memory loss, language loss, personality changes and impaired problem-solving skills.[8]

Burden of Disease

By 2050 an estimated 21 million Americans will be over the age of 85 – nearly quadruple the number in 2012.[7] Age has been clearly indicated as a risk factor for dementia.[9] One in eight people age 65 and older has AD and it is estimated that 13.9% of people over the age of 71 have some form of dementia.[7] In men alone, there are currently 1.8 million individuals over the age of 65 who are living with AD, meaning 11% of men in this age group have AD or some form of dementia.[7]

There is a high financial burden associated with cognitive impairment in the elderly. For example, in 2008, Medicaid payments for Medicare beneficiaries over age 65 with AD and other dementias were nineteen times those for cognitively intact patients.[7] These costs are associated with the long-term care needs of these patients, particularly their utilization of nursing homes and assisted living facilities. Additionally, Medicare recipients with Alzheimer's and other dementia patients paid an average of $9,368 each in out-of-pocket expenses to cover long term care needs in 2008.[7]

In 2008, 82,435 people died of AD, although this is likely an underestimate as death certificates may list an acute condition as the primary cause of death instead of the underlying cognitive impairment.[7] Deaths from other major causes are decreasing, while deaths from AD and other dementias are increasing. Between 2000 and 2008, death

attributed to AD increased 66% while those attributed to heart disease decreased 13% over the same time period. AD is the fifth leading cause of death in people over 65 years.

Symptoms, Diagnosis and Treatment

Clinically, cognitive decline, induced by AD or VD, includes a variety cognitive symptoms including memory loss, confusion with time or place, challenges in solving problems, trouble understanding visual images and spatial relationships, poor judgment, and changes in mood and/or personality.[7] Progression through the disease is unique for each individual, and no effective treatment exists. Because cognitive tests are not a routine part of regular examinations, elderly patients are usually diagnosed with clinical dementia long after the initial onset of symptoms.[7] Subtle differences in the initial symptoms seem to be related to different forms of dementia, with memory loss being more prevalent in AD and impaired judgment being more common in VD, although disease presentation can vary greatly between individuals.[7] Patients eventually reach a stage in which they can no longer care for themselves and have impaired mobility, which is when a diagnosis of dementia is given.[7] This frequently leads to vulnerability to infections, Alzheimer's-related pneumonia, and eventually death (Figure 1).

VD is diagnosed when an MRI image displays vascular injuries in the brain that accompany low cognitive function. AD can be diagnosed when other causes of dementia have been ruled out such as VD, tumors, or stroke.[1] AD is not generally confirmed until after death, when brain material can be closely examined.[8] For these reasons there is much overlap in the diagnosis of Alzheimer's and VD.[10]

There is currently no treatment for AD or VD. The primary drugs used to maintain mental function are Cholinesterase inhibitors, thought to slow mental decline by

increasing brain levels of the neurotransmitter acetylcholine. In clinical trials they are relatively effective showing and odds ratio of 1.56 (95% CI 1.32 – 1.85) when comparing those who improved to those who stayed the same or deteriorated over six months of use.[11] Efficacy does seem to vary across populations and even at its best only temporarily reduces symptoms.[7] Interventions to manage the disease such as management of coexisting conditions, participation in activities, medications, and support groups, can increase the quality of life for those afflicted with cognitive impairment, but nothing has been found that can stop or reverse the progression of the disease.[12]

Role of Executive Functioning in AD and VD

Executive function is considered to be the higher-order cognitive capacities that are necessary to support independent, purposive, goal-directed behavior.[13] The National Center for Learning Disabilities defines it as the set of mental processes used to perform activities such as planning, organizing, strategizing, paying attention, remembering details and managing time and space.[14] Its presence is critical for elderly individuals to maintain independent lives, multi-task, and plan. Decline in executive functioning is frequently seen in conjunction with decline in memory or global cognitive function, but some studies suggest that preclinical deficits in executive functioning precede cognitive impairment.[15] Executive function may be impaired two to three years prior to AD diagnosis.[16] Furthermore, executive functioning is one of the primary cognitive domains that have been shown to be independently associated with certain forms of cerebrovascular pathologies on MRI scans of brain tissue.

Mechanisms Underlying Cognitive Impairment

We are interested in identifying genes associated with cognitive decline, but the mechanisms by which genetic difference cause the disease processes are important to consider. The mechanisms by which AD affects the brain are not entirely clear. There is ample evidence to support the disease progression listed below, which is initiated by disturbance of neuronal protein homeostasis, characterized by the following processes (Figure 2) [17]:

1. Faulty processing of the amyloid precursor protein by beta secretase enzymes predominately lending toward longer cleavages of the amino acid chains of beta-amyloid protein that begins to form "sticky" oligomers and then deposit in the extracellular space as plaques. In addition, the microtubule matrices inside the neurons begin to break down as tau proteins begin disentangling, become phosphorylated and eventually become neurofibrillary tangles. Disturbed metabolism leads to degeneration of neuron, atrophy of the axon and eventual loss of synapses necessary for cell connectivity and transmission.
2. Oxidative stress damages the mitochondria and together these phenomena are the underlying features responsible for the neurodegeneration seen in AD.
3. As these molecular and cellular processes progress, structural changes are observed in the brain using MRI: the entorhinal cortex, hippocampus, and cerebral cortex begin to atrophy while the ventricles begin to expand; all signs of neurodegeneration.
4. This sequence is followed by cognitive decline and eventual loss of the ability to maintain independent living.

While the “plaques” and “tangles” appear to have an important role in AD pathology, there is evidence that cerebrovascular disease also plays a role in many cases of AD.

VD is also associated with symptoms of cognitive deficit and motor dysfunctions in the case of stroke induced dementia. MRI methods are readily available to characterize early features of a vascular “mediated” cognitive decline and the cerebrovascular pathology, including small vessel atherosclerosis.[3] These are seen on MRI images as WMHs and are associated with accelerated cognitive decline, VD and AD. One distinct difference between the two, from a clinical perspective, is that AD has an insidious presentation taking years if not decades to evolve, while VD generates abrupt changes. AD is the most commonly diagnosed form of cognitive impairment, but the importance of other forms of cognitive decline, such as VD are becoming increasing recognized as significant contributors to this public health burden.[1] However, the presence of VD does not exclude Alzheimer’s. In fact, it is now believed that both are frequently present together.[7] For these reasons there is much overlap in the diagnosis of Alzheimer’s and VD, and the term vascular cognitive impairment has recently been proposed to encompass this phenomenon.[10]

Most epidemiological literature is focused on AD, but there is a frequent co-occurrence of AD pathology and other pathologic phenomena important to the risk for cognitive decline. This thesis highlights “vascular driven pathways” to cognitive decline, where VD is the end of the clinical spectrum. Because Alzheimer’s and VD have similar risk factors (e.g., age, education, vascular pathologies such as high blood pressure, high cholesterol and diabetes), public health efforts to reduce dementia incidence from each may be similar. Additionally, current cognitive tests alone cannot distinguish between

VD and AD (or other, less common, causes of cognitive impairment) From a genetic standpoint, the genotypes linked to these phenotypes may also overlap. This overlap introduces some ambiguity when evaluating genetic associations, but it is currently very difficult to clearly separate these cognitive function phenotypes. Because of the homogenous nature of the cognitive phenotypes, this thesis will consider literature pertaining to both cognitive phenotypes.

Risk Factors for Cognitive Decline

Identification of risk factors for dementia or cognitive decline might allow us to take the measures necessary to prevent the onset or to delay the progression. Understanding disease pathology is important as it can provide early warning signs and aid in secondary prevention of the onset of symptoms. It is also necessary to further understand the disease itself and develop effective treatments and prevention methods. Identifying genetic risk factors, however, might allow primary prevention, where those with increased risk for cognitive decline could take preventive measures before any disease pathologies begin, and far before symptoms are experienced. Because there is no known way to prevent the onset of the disease, research must take a two-pronged approach, identifying risk factors and identifying potential prevention strategies.

AD, like many chronic diseases that plague our society, is caused by a combination of genetic and environmental factors and its underlying pathology does not always present clinically as AD dementia. The greatest risk factor for development of AD is aging, with the majority cases being diagnosed after age 65, and is present in over half of individuals over age 85.[7] A higher prevalence in women is confounded by the fact that women live

longer than men. There is no strong evidence to believe that gender modifies risk for AD dementia.[7]

AD does show familial patterns and is considered to be at least partially heritable. Table 1 summarizes the results found in various studies examining the heritability of cognition. The first study showing increased risk of AD in family members of people with AD was published in 1981 and showed that subjects with relatives who had autopsy-confirmed AD had increased risk of experiencing cognitive decline with the probability of 15.1% ± 2.6% of developing AD by age 84 compared to 5.5% ± 3.0% in the control group, suggesting the possibility for genetic transmission.[18] In 1993, a common genetic variant in the gene coding for apolipoprotein E (*APOE*) was found to be associated with increased risk for AD. Individuals carrying one or more copy of *APOE* e4 allele have 3.68 times the odds (95% CI 3.30 – 4.11) of developing AD than non-carriers.[19,20] This association has been confirmed numerous times and across various populations, but it is neither *necessary* nor *sufficient* for the development of cognitive impairment and its causal pathway.[21] Since the discovery of the association with *APOE,* many more genetic associations have been identified, particularly since the advent of genome-wide association studies (GWAS), which allows large numbers of markers to be simultaneously assessed. None of these genes however, has been successful in explaining large amounts of variation, with *APOE* estimated to explain only 7%-9% of total variation,[22] and all other genes explaining even less of AD pathologies, leaving much of the genetic associations poorly understood.[23]

In addition to genetic risk factors, vascular risk factors are important in cognition pathology. Development of AD has been linked to various vascular risk factors, including

high blood pressure in mid-life, type 2 diabetes, high body weight, and high cholesterol levels (Table 2).[24] Additionally cerebrovascular disease is known to be a risk factor for VD, but is also considered as a risk factor for AD.[24] Forms of cerebrovascular disease, such as ischemic infarcts or other changes in white matter in MRI imaging, have been clearly linked to increased risk of dementia, and co-occurrence with AD is common.[7] It is now thought that age-related cognitive impairment can be attributed to AD pathology and/or cerebrovascular disease.[1] This suggests that cardiovascular risk factors might be used to both predict and provide prevention targets for these related diseases, and many already have.[25] Additionally, because cerebrovascular diseases can be visualized as infarcts or white matter changes using MRI, this imaging could be used as predictive and diagnostic tools for AD.[1]

White Matter Changes/Hyperintensities and Cognitive Decline

MRI of the brain can permit the study of structural and functional changes in the brain that are pertinent to cognitive behavior. Brain MRI can detect structural changes that occur with aging. For example, changes in the gray matter, specifically the hippocampal change in volume can predict decline in memory recall.[26] Changes in the white matter also occur but have been understudied in relation to AD risk and cognitive decline outside of stroke and VD. These white matter signals, called lesions, are usually identified as leukoariosis or white matter hyperintensities (WMH) on T2 weighted or FLAIR MRI scans, and are prevalent in up 60-92% of the elderly population.[27,28] WMH have been shown to represent cerebrovascular disease[25,29], atherosclerosis[30] and demyelination of neurons in elders at autopsy.[31] They have been associated with vascular risk factors such as diabetes, fasting glucose, hypertension, systolic blood pressure, and

endothelial dysfunction.[32,33] White matter burden and changes are associated with vascular cognitive impairment, a syndrome that includes both cognitive impairment in one or more cognitive domains and evidence of vascular brain injury (stroke or subcortical WMH).[1] WMHs, a form of subcortical vascular brain injury, have recently been shown to independently predict dementia in the general population with the associations being found with both VD and AD.[34]

Various forms of white matter damage have been observed to be related to vascular cognitive impairment. White matter hyperintensities, which include areas of demyelination as well as silent infarcts,[33] have been clearly linked to all-type dementia.[35] Additionally, larger volumes, and increased numbers of macroscopic infarcts have been associated with an increased likelihood of dementia.[1] In general, it appears that greater total and periventricular white matter hyperintensity burden is associated with cognitive impairment.[36] While the mechanisms by which white matter changes affect cognition are not well understood, the association has been consistently observed.[31]

Not only have white matter hyperintensities been linked to all form dementia and AD,[31] but also more specifically, white matter lesions may impact particular cognitive domains more than others. For example, literature is mounting that show WMH associated with lower executive function. In a population-based study, Prins *et al*. found that periventricular white matter lesions, brain infarcts and generalized brain atrophy were associated with the rate of decline in executive function.[3] More recently WMH volume was also linked to increased relative risk for lower executive function in elderly community-dwelling subjects even after adjustments for sex, age, education and cardiovascular risk factors (Relative Risk = 1.55, 95% CI 1.06-2.26).[4]

Genetic Basis for White Matter Hyperintensities

In 2011, Fornage *et al.* conducted a meta-analysis of genome wide association studies in which they examined white matter hyperintensity (WMH) burden in 9,361 stroke-free individuals from 7 cohorts.[2] Verhaaren *et al.* replicated these results in 2011.[37] Using a significance threshold of $p<5 \times 10^{-8}$, the study identified SNPs in or around six genes having genome-wide statistically significant associations with WMH burden: *WBP2, TRIM65, TRIM47, MRPL38, FBF1* and *ACOX1*. These genes are in the 17q23 region meaning they are on chromosome 17, on the long arm of the chromosome as represented by "q", and are in the 23rd band that you can see using a microscope and counting away from the centromere. The current known function of these six genes provides limited insight into their role in cognitive decline mediated by silent vascular brain injury, in fact little is known about most of them (Table 3). *TRIM65, TRIM47* and *FBF1* have been linked to apoptosis which may occur in white matter lesions[38,39], but no specific biologic mechanisms have been established. Increased *WBP2* expression has been associated with decreased noradrenaline transporters in mice, and decreased noradrenaline has been associated with AD in humans.[40,41,42] *MRPL38* and *ACOX1* are both associated with oxidative stress, but again specific mechanisms are not clear.[42-45]

Fornage *et al.* further examined two of these SNPs for associations with stroke, dementia and AD, and no significant association was found in these cohorts.[2] However, discovering the underlying relationship between genes affecting WMH and cognitive function may require testing the association of a complete set of SNPs from these genes in large cohorts of participants at risk for cognitive impairment. Furthermore, WMH have been independently linked not only to vascular cognitive decline, but also more

specifically, to impairment and accelerated decline of executive functioning,[3] suggesting that use of a more specific phenotype may also aid in elucidating this relationship.

Summary

Cognitive impairment is a common and costly disorder with high prevalence and high projected incidence as the population in America ages. AD and VD, the two most common forms of cognitive impairment, are similar conditions with many overlapping clinical presentation and risk factors. While many mechanisms underlying cognitive impairment are not yet fully understood, many risk factors have been identified including age, *APOE* gene, hypertension, diabetes, BMI, depression, cholesterol, diet and exercise. MRI technology is further aiding the understanding of this pathology. Changes in white matter, including areas of WMHs, that can now be visualized and quantified, have been associated with cognitive performance, specifically executive functioning. Furthermore AD, VD and WMHs all share common vascular risk factors such as diabetes, hypertension, systolic blood pressure and stroke. Six genes were recently identified as being associated with WMH burden, and while the function of these genes is still poorly understood, it provides one step towards understanding the genetic basis of this neuroimaging phenotype, as well as candidate genes to study in association with cognitive phenotypes.

Research Question

Foranage *et al*. identified six genes in the 17q23 region associated with white matter hyperintensites (Table 3). Several studies have demonstrated WMH accumulation is associated with worse cognitive performance, particularly in the domain of executive function. Our primary hypothesis is that genes associated with WMH are

also associated with executive functioning, a cognitive skill that appears most affected by the accumulation of WMH. The following specific aims are investigated in this study:

1. Test whether there are associations between SNPs in these six genes and executive function as measured by Trails B.
2. Test whether there are associations between SNPs in these six genes and global cognitive function as measured by 3MS.
3. Compare and describe the SNPs in association with Trails B scores and 3MS scores in this MrOS sample.

METHODS

Data Management and Collection

This study was conducted using data from the Osteoporotic Fractures in Men Study (MrOS). MrOS is a multi-center prospective cohort study designed to examine potential causes of osteoporotic fracture in elderly men (65 years and older), and also includes additional measures related to many aspects of aging, of which several will be used for this analysis. It enrolled subjects between March 2000 and April 2002 at six study sites: Birmingham, Alabama; Minneapolis, Minnesota; Palo Alto, California; Monongahela Valley near Pittsburgh, Pennsylvania; Portland, Oregon; and San Diego, California. Because the primary focus of the MrOS study was osteoporosis, eligible men must have been able to walk without assistance, not have had bilateral hip replacements, and be community-dwelling, meaning none had been diagnosed with AD. All of these men had clinical and population characteristics measured at baseline and at follow-up studies. This study is limited to baseline data alone, which will be used for this cross-sectional analysis. Detailed descriptions of the MrOS cohort and its data collection have

been published elsewhere.[46,47] This study is funded by the National Institute of Aging and the National Institute of Arthritis and Musculoskeletal and Skin Diseases. I used the February 2012 data release. This protocol was approved by the OHSU IRB board. All data that I received was completely anonymized, prior to data being sent to me. I did not have the key to the code for the anonymized data**.** Ages over 90 were truncated to 90 before I receive the data and I did not know the key to find out specific ages for this subset. There were 21 individuals who were classified as 90 years or older.

Inclusion/Exclusion Criteria

For this analysis only data from five study sites was used: Minneapolis, Minnesota; Palo Alto, California; Monongahela Valley near Pittsburgh, Pennsylvania; Portland, Oregon; and San Diego, California because the consent form used in Birmingham did not include consent to have genotyping data analyzed for cognitive traits. Additionally, only Caucasian men had genotype data available, so analysis was limited to this racial group, which represented 77% of the cohort. It was also limited to those who had genotyping data and cognitive measures available at baseline. Additionally, those with a history of stroke were eliminated, as this is a known cause of cognitive decline, and the purpose of this study is to examine other causes of dementia (Figure 2).

Outcome Definitions

Cognitive Tests

There are many neuropsychological exams that have been validated for testing cognitive performance. These exams include some that tap global performance, and others are designed to assess language, attention, visuo-spatial skills, processing speed

and memory more distinctly. Two psychometrics used in this study are: The Trail Making Test Part B (Trails B) and Teng's Modified Mini-Mental (3MS) examination.

Trails B is commonly used as a measure of executive functioning,[5] and executive function is important to maintaining activities of daily living in older adults.[16] This timed test asks subjects to connect numbers and letters alternately and in order as quickly as possible and time is recorded as seconds. Various studies use different "end-points" for stopping the test when subjects have not yet completed the task, but the most common are to truncate at 200 or 300 seconds. This test has been validated as a measure of executive function by comparing it to a set-switching task known to provide an index of executive function.[5]

In the MrOS cohort, subjects were timed while performing the test and if a mistake was made, the test administrator corrected subjects, and it was assumed that the mistake was accounted for by the longer time needed to complete the task. Each subject in the MrOS study had the Trails B test administered at baseline and at two subsequent follow-up visits. Subjects in the MrOS study were given 300 seconds to complete the test. Those who did not complete the test in the allotted time had 300 seconds recorded for their time. Trails B was used as a continuous variable for this analysis.

The Teng's Modified Mini-Mental State Examination (3MS) is a modified version of the Folstein Mini-Mental State Examination (MMSE). The MMSE asks questions testing basic arithmetic, language use and comprehension and basic motor skills. The 3MS goes on to incorporate a few additional questions to cover a broader range of cognitive function. This test is graded on a scale of 0-100, with different weights given to various questions, and a score below 80 was considered to indicate cognitive impairment.[6] It is a

short and easily administrable exam that provides a measure of global cognitive function without examining a specific domain.[6] 3MS has been validated in cohort of individuals over age 65 as a reliable test of global cognitive performance.[48]

The 3MS was used to provide a measure for scoring global cognitive function in MrOS. Each subject in MrOS had the 3MS tests administered at baseline. 3MS score was used as a continuous variable for this analysis.

Genetic Data Generation in MrOS

Genomic DNA was extracted from blood using the Flexigene (Qiagen, Valencia, CA, USA) protocol at the University of Pittsburgh. Of the 5994 MrOS participants, 5551 provided blood samples and 5506 had sufficient DNA quantity for whole-genome sequencing. Genotyping was performed on members of the MrOS cohort for use in GWAS. These studies are used to scan the genomes for SNPs associated with an outcome of interest.[49] This can be very useful to identify novel regions of the genome that may be associated with an outcome, and is the method that was used to identify the six genes in the 17q26 region this study is examining. In MrOS, GWAS data was generated by whole-genome genotyping, creating a set of genotyped SNPs from which specific SNPs could also be used to test specific associations as I am doing in this study.

Whole-genome genotyping was performed using Illumina's HumanOmni1 Quad genotyping array at the Broad Institute. Genotypes were found using a clustering algorithm in Illumina's BeadStudio software.[50] In this process DNA is hybridized to probes matching specific genotypes where fluorescence is observed if the DNA is able to bind to a probe and is not observed if it has the opposite allele and cannot bind. Illumina's software identifies these reactions for the thousands of different SNPs being

tested to determine, or "call", each genotype in each individual. When DNA hybridizes to the probes on these arrays they form clusters, and the software that reads these results must distinguish between the various clusters to determine specific genotypes. The software has various built in measures to determine if these clusters are physically distinguishable thereby making the genotype calls more or less reliable. Samples were excluded due to poor quality if they had call rates lower than 97%, referring the frequency with which the software feels it can make an accurate call. GenTrain scores reflect the shape and inter-cluster distance on the arrays and scores lower than 0.6 were excluded. Cluster separation scores lower than 0.4, meaning the clusters were too close together to call, were also excluded. Additionally minor allele frequency lower than 0.01 were excluded as genotyping error would have a large effect on these SNPs. SNPs out of Hardy-Weinberg equilibrium were assumed to have genotyping error and were excluded (Pearson's chi-squared p-values $<10^{-4}$). Of the 1,016,423 SNPs on the array, 740,713 passed this quality control.

As further quality control, duplicates were genotyped for 81 samples and pairwise concordance was 100%. A diverse group of population samples were used for quality control (QC). Replicates of subjects used in HapMap trios of Utah residents with Northern and Western European ancestry (CEU) and Yoruba in Ibadan, Nigeria (YRI) populations and singletons from Han Chinese in Beijing, China (CHB) and Japanese in Tokyo, Japan (JPT) populations were genotyped along with MrOS samples.[51] Comparisons between these replicates and HapMap genotypes were determined to assess the quality of the genotyping within the MrOS cohort. Concordance between the control samples and HapMap genotypes was 99.7% and 95.0-99.7% for CEU and YRI samples

and CHB and JPT samples respectively indicating acceptable quality for the MrOS genotypes.

SNP Inclusion Criteria

Using dbSNP, a list of 1692 SNPs was composed that are in the six genes of interest: *WBP2, TRIM65, TRIM47, MRPL38, FBF1* and *ACOX1*. The analysis was based on all SNPs from this set that were also included on the Illumina 1M genotyping chip datasets that are available for study subjects. For each SNP, each subject had one of three genotypes: homozygote for the major allele, homozygote for the minor allele, or heterozygote. Of the 1692 SNPs identified in these target genes, only 45 were available in the MrOS GWAS dataset. Of these 45, 4 SNPs were excluded on the basis of having only one allele represented. Five SNPs did not have any individuals who carried the homozygote minor allele genotype. These SNPs could not be used for recessive effects models, but were able to be used for additive effects models (these models are explained below). SNPs in each gene were assessed for correlations with each other (Appendix 1), and highly correlated SNPs had one of the two removed at random. For this pruning only correlations within genes, not between genes, were considered. SNPs were considered highly correlated when $r^2 > 0.8$ and it was assumed that SNPs that were higher correlated were in strong linkage disequilibrium and would therefore be testing the same effects (Table 4).

Principal Component Analysis

Common ancestry can cause confounding of the relationship between SNPs and any outcome of interest as ancestral differences in cognition (attributable to genes not included in this study) could be mistakenly attributed to our specific SNPs. It is necessary

to account for this potential population stratification resulting from systematic ancestry differences in the population studied. For example, systematic ancestral differences may be responsible for causal SNPs underlying a phenotype of interest, but these same ancestral patterns may be responsible for non-causal SNPs being tested in relation to the phenotype. This circumstance may yield an association between these non-causal SNPs and the phenotype, that is not a true association but simply seen because of common ancestral patterns. Principal component analysis (PCA) is a method by which to model these ancestral differences and create variables that can be used to control for this stratification in other models. PCA can be used to reduce the complexity of high dimensional genomic data to lower dimensions while continuing to explain as much of the genetic variation in the entire sent of genes as possible. These resulting variables are principal components, called eigenvectors, and can be used in subsequent analysis within the same population to adjust the relationship between SNPs and phenotypes, thereby attempting to account for any systematic ancestral differences in the population.[52]

In MrOS PCA was used to model ancestry differences within this cohort using its GWAS data and including 35,769 SNPs in 4,637 men. There is a chance that some of the SNPs I am testing were included in this PCA, but because of the large number for total SNPs used for this analysis, our SNPs would not greatly influence the outcome of this analysis, and making it unlikely that adjusting for PCA would cause potential associations between our SNPs and cognitive measures to not be found. The first four of the resulting eigenvectors were then used in analyses to account for any systematic ancestry differences within the population and to ensure that spurious associations, due to common ancestries, are not observed.[52] While more than four eigenvectors were

generated, each successive one explains a smaller proportion of the variation, so the four of these eigenvectors should adequately account for any population stratification due to ancestry.

Covariates

Because of its strong link to cognitive decline, age is important to consider. Age is the strongest known predictor of dementia, and it was necessary to account for this covariate in any analysis. This cohort has a minimum age of 65 years. Subjects with ages greater than 90 years were censored to 90 years, as an age over 90 years is considered identifiable data and this type of data was not approved by the IRB.

Study site can also cause confounding through ancestry differences. This could occur through systematic ancestral differences by study site. While the principal components should account for this phenomenon, we do see highly significant differences between study sites even after adjustment for age. Therefore, we adjusted for it in each model, both base and multivariate.

Other than ancestry, confounding was not a concern in this analysis because no other variable, can influence a subject's genotype. However, because various other factors have been linked to cognitive decline, there was a potential for covariates to be influencing cognitive measures. The covariates that were examined and considered for a final model are: BMI, total cholesterol, LDL cholesterol, HDL cholesterol, hypertension status, anti-depressant use, selective serotonin reuptake inhibitor (SSRI) use, systolic blood pressure, diabetes status and education level. All variables were measured at baseline at the subjects' first visit, the same visit at which cognitive measures were assessed.

BMI is a calculated variable made by dividing each subject's mass in kilograms by their height in meters squared, as measured at the baseline visit. Total cholesterol, LDL cholesterol, HDL cholesterol and systolic blood pressure were measured in mg/dL at baseline. Diabetes status was measured at baseline using fasting glucose (≥8hrs) using a Hitachi 917 Autoanalyzer. Subjects were put into one of three categories: Diabetes (fasting glucose ≥126 mg/dl, or self-reported diabetes at baseline or use of hypoglycemic medications at baseline), impaired fasting glucose (fasting glucose ≥100 and <126 mg/dl) or normoglycemic (fasting or non-fasting glucose <100 mg/dl). Current and previous hypertension status was self-reported by participants as well as current and previous SSRI and antidepressant use.

The highest level of education was reported for each participant. There were eight categories possible: some elementary school, elementary school, some high school, high school, some college, college, some graduate school, and graduate school. All eight categories are presented with the population characteristics (Table 5). Due to small sample size for some of the levels of education, some of the groups were combined to create a new variable definition. They were combined based on similar Trials B scores to create an education variable with five levels: elementary school or less, some high school, high school or some college, college, some or all graduate school.

Models

To test for associations between SNPs in the six genes of interest, I built a variety of linear regression models. First each SNP was used to build additive regression models on each of my cognitive outcomes: 3MS and Trails B. An additive model uses three levels for the genotype: homozygous for the minor allele, heterozygous and homozygous

for the dominant allele. These were coded as 0, 1, and 2 respectively in this study. This allows one to test the effects of having zero, one or two copies of an allele separately. Each SNP was then also used to build recessive regression models on each cognitive outcome. Recessive models use only two levels for the genotype: those that contain no copies of the major allele, and those that contain the major allele. These are coded as 0 and 1 respectively. This allows one to test the effects of having a dominant allele versus not having a dominant allele. Using both methods allows one to look at different types of genetic associations. If the SNPs are related in a dominant fashion, and the heterozygous and homozygous dominant genotypes are phenotypically equivalent, then a recessive model will be best to identify associations. If a dose effect is present, where those with heterozygous genotypes fall phenotypically between the two homozygous genotypes, then additive regression will be best to identify the associations.

Each SNP was first used to build both additive and recessive regression models on each cognitive outcome adjusting for age and study site only. Second, each of these models was built again including the first four eigenvectors as covariates. Then, each model was also rebuilt using censored normal regression and adjusting for age, study site and the first four eigenvectors. Censored normal regression was used to account for the spike in scores at the top end of both the Trails B and 3MS distribution. Because participants were cut off after 300 seconds during the Trails B test, those who did not finish had a time of 300 seconds recorded, creating a spike at this final value (Appendix 2). Additionally, 3MS has a maximum score of 100 points, and therefore also has a spike at 100 points that represents all subjects with perfect scores (Appendix 2). It is assumed that among those with 300 seconds on Trails B or 100 points on 3MS that some variation

still exists, and censored normal regression can account for this. Using this method allowed a comparison between censored and non-censored models to see if these spikes were greatly influencing the outcomes.

Finally, linear regression models were built that not only adjusted for age, study site, the first four eigenvectors, but also the risk factors that were independently related to the cognitive phenotypes. These variables were chosen after looking at their associations with Trails B and 3MS in the first step of this analysis. This step was done to make sure that no factors independently related to the cognitive phenotypes were responsible for explaining the variation being attributed to the SNPs of interest.

Statistical Analysis

All statistical analyses were performed using SAS version 9.3 (SAS Institute Inc.) and used the following six-step process:

Step 1 – Summarize Population

Baseline demographic and clinical characteristics were summarized in all subjects Cognitive risk factors described are: age, education level, antidepressant and SSRI use. Vascular risk factors described are: hypertension status, total cholesterol, LDL cholesterol, HDL cholesterol, systolic blood pressure, and diabetes status. Means and standard deviations were reported for continuous variables. Population proportions were reported for dichotomous and categorical variables.

Step 2 – Covariate Analysis

Each covariate discussed above was analyzed for an association with Trails B time and 3MS score in the analytic dataset to identify potential effect modifiers. Non-parametric measures were used to test the relationships between covariates and both

Trails B and 3MS scores as both measures deviated from a normal distribution. Mean and standard deviation of each outcome variable (Trails B and 3MS) were calculated for each dichotomous covariate (hypertension, SSRI use and antidepressant use) and compared using a two-sided Wilcoxon rank sum test. Continuous covariates (age, BMI, total cholesterol, LDL cholesterol, HDL cholesterol, and systolic blood pressure) were divided into quartiles to create categories by which to compare Trails B and 3MS scores. Means of each of the quartiles as well as means from the other categorical variables (education and diabetes status) were compared using a two-way Kruskal-Wallis test. Continuous variables were also evaluated for correlation with both Trails B and 3MS and a Spearman rank correlation coefficient was reported. The Spearman rank correlation was repeated after adjusting for age. Age-adjusted p-values are also reported for non-continuous covariates. These were generated using linear regression analysis of each outcome of interest on each covariate while adjusting for age. Finally, expected genotype frequencies were calculated and each SNP was tested for Hardy-Weinberg Equilibrium using a Fisher's exact chi-squared test, to account for very low frequencies of certain alleles.

Step 3 – Initial Models

Linear regression models were built to test the additive effects of each of the 26 SNPs on Trails B and 3MS. Additive effects models considered three genotypes: homozygote for the minor allele (coded as zero), heterozygote (coded as one) and homozygote for the major allele (coded as 2). Five of the SNPs being used had no homozygote recessive genotypes present in this population, so only the remaining 21 SNPs were used to build recessive effects models. Recessive effects models consider

only two genotypes: homozygote for the minor allele (coded as zero), and those that contain at least one copy of the major allele (coded as one). These models included age, and study site, but no other covariates. Then a second set of models was build that included age, study site and the first four eigenvectors. Additionally, censored normal regression models were built for each outcome and both the additive and recessive effects models to account for the non-normal distributions of Trails B and 3MS.

Because each association was being tested with 26 SNPs, a false discovery rate (FDR) correction was used on each raw p-value.[53] This method accounted for multiple comparisons, while not being as strict as a family-wide error rate correction. A Bonferroni correction would mean that p-values for each SNP would need to be <0.0019 to be considered statistically significant. The FDR correction, however, adjusts the raw p-values by ranking them in ascending order, and then dividing each raw p-value by its rank, creating an adjustment that penalizes larger p-values less than smaller ones. [54]

Step 4 – Adjusted Models

Models were built adjusting for all covariates that showed a significant relationship to either Trails B or 3MS. Models of Trails B were adjusted for BMI, HDL cholesterol, systolic blood pressure, antidepressant use, diabetes status and education. These models were also adjusted for study site, age and the first four eigenvectors as in the base models. Models of 3MS were adjusted for age, BMI, HDL cholesterol, systolic blood pressure, diabetes status and education, study site and the first four eigenvectors. Linear regression models including these covariates were built to test both additive and recessive effects on Trails B and 3MS as described above. Additional censored normal regression models were also built using these covariates for both Trails B and 3MS

testing both additive and recessive effects to verify that deviation from normality was not affecting the linear regression (data not shown). For each model, p-values were adjusted using a FDR correction.

Step 5 – Model Diagnostics

Residual analysis was performed on models that showed SNPs significantly associated with an outcome. Plots showing raw residuals versus expected values were examined for unexpected patterns. Leverage graphs were also examined for outliers and Cook's distance was used to identify outliers and assess their influence and check for errors in the dataset.

Step 6 – Final Model

After looking at the associations above, the dataset was divided randomly into two groups. Using one of these groups we attempted to create a model that explained the most variation in Trails B using a subset of SNPs. Multivariate linear models were built using subsets of SNPs (as well as age and principal components) and Mallow's C_p, was used to compare the resulting multivariate models. Mallow's Cp is a calculated value, using the mean squared prediction error, which allows for comparisons between models using different subsets of all possible variables. The value of Mallow's Cp is expected to be close in to the number of predictors in the model, so a model with a Cp value mostly closely matching the number of predictors will be the best model. This method is used to select the model that allows the highest predictive ability without over adjusting for two many variables. The model that best explained variation in Trails B was then applied to the second half of the dataset to test if the identified SNPs were still significantly associated with Trails B.

RESULTS

Study Population

A total of 3552 and men with a mean age of 74 years (± 5.9) from the MrOS cohort met the inclusion criteria for this analysis. Subjects represented each of the 5 study sites with similar frequency. The mean Trails B score was 129.5 (± 56.3) and the mean 3MS score was 94.1 (± 5.0). Antidepressant use in this cohort was low with only 181 mean (5.5%) and even fewer subjects using SSRIs (n=78, 2.3%). In this group, 1636 (46.4%) subjects were normoglycemic, 1174 (33.3%) subjects experienced impaired fasting glucose, and 487 (13.8%) had diabetes. Close to half the dataset reported hypertension (41.2%). These men were highly educated (55% completed college or higher). Population and clinical characteristics are presented in Table 5.

Risk Factors and Cognitive Measures

Risk Factors and Trails B

Trails B was correlated with age ($r=0.34$, $p<0.001$), as expected, since age is known to be the greatest predictor of cognitive impairment. Trails B was also correlated, after adjusting for age, with systolic blood pressure ($r=0.38$, $p=0.03$), and was only weakly correlated with BMI ($r=0.05$, $p=0.002$) and very weakly correlated with HDL cholesterol ($r=-0.002$, $p=0.002$). Initially it appeared that high Trails B times (indicating poorer performance) were associated with lower LDL and total cholesterol values, but after adjusting for age, these correlations were no longer significant. As expected, Trails B times were highly associated with education level, diabetes status, and antidepressant use (all $p\leq0.001$). The associations between Trails B and other cognitive risk factors are presented in Table 6.

Risk Factors and 3MS

3MS was also correlated with age ($r=-0.29$, $p<0.001$), as expected. Also similar to Trails B, after adjustment for age, 3MS scores were correlated with BMI ($r=0.08$, $p<0.001$), HDL Cholesterol ($r=0.07$, $p<0.001$), and systolic blood pressure ($r=-0.08$, $p<0.001$), although all correlations were small and likely spurious and only due to the large population size. Also similar to Trails B, higher 3MS scores (indicating better performance) appeared initially associated with lower LDL and total cholesterol values, but after adjusting for age, these correlations were not longer significant. 3MS scores were highly associated with education level and diabetes status (both Kruskal-Wallis $p<0.001$), but unlike Trails B times, they were not associated with antidepressant use. The associations between 3MS and other cognitive risk factors are displayed in Table 7.

Study Sites and Cognitive Measures

Both Trails B and 3MS means varied significantly by study site ($p<0.001$ for both tests). These relationships were still significant for both cognitive measures after adjusting for age ($p<0.001$). Mean Trails B scores and mean 3MS scores are shown in Table 6. Differences by study site could indicate ancestral differences by study site, and could confound analysis. Therefore, study site was adjusted for in each model of SNPs on Trails B or 3MS.

Trails B and 3MS

Both Trails B and 3MS measure cognitive performance, but they have differences in the types of performance they measure making it useful to see if these measures outcomes are related to each other. The two measures were only moderately correlated with $r=0.43$ ($p<0.001$).

Linear Regression Models

SNPs

Twenty-six SNPs were selected for this analysis. All SNPs were also found to be in Hardy-Weinburg Equilibrium. Trails B and 3MS scores by genotype of each SNP are displayed in Appendix 3 and Appendix 4. To account for potential ancestral differences in the population, previously calculated principal components were used as covariates in both the base and adjusted models.

Use of Principal Components Analysis

Initial models were built with and without the inclusion of the eigenvectors from the principal component analysis. The models with and without the eigenvectors were very similar and they did not cause any noticeable change in the magnitude of effect or in which SNPs were significantly associated with outcomes. This suggests that confounding by ancestry was not a problem in this cohort.

Additive Effects Model on Trails B

An Additive Effects Model was built on Trails B for each of 26 SNPs (Table 8 and Appendix 5), adjusting for age, study site and the first four eigenvectors. SNPs in 3 different genes were found to have significant non-corrected β-values: rs7208173 in *TRIM65* (β=-45.2, p=0.024), rs9892372 in *MRPL38* (β=-19.8, p=0.011), and rs8082018 in *ACOX1* (β=-20.6, p=0.007). Negative beta-values correspond to a decreased Trails B score associated with the minor allele because the genotypes were coded such that a homozygote minor allele genotype is zero, a heterozygote genotype is one and a homozygote the major allele genotype is 2. For example, a β of -19.6 in the rs9892372 in *MRPL38* indicates that men who are heterozygous have a mean Trails B score that is 19.6

seconds longer than those with the homozygote major genotype, and 19.6 seconds shorter than those with the homozygote minor genotype. Considering the standard deviation of Trails B in this cohort is 56.3 seconds, the effect of one additional minor allele corresponds a decrease of 34.8% of one standard deviation in Trails B.

In this population, rs7208173 in *TRIM65* has a minor allele (C) frequency of 0.001, has no individuals with the CC genotype and only 7 individuals with the CT genotype, bringing into question whether or not this observed relationship is spurious. rs9892372 in *MRPL38* has only 2 individuals with the homozygote minor allele genotype (mean Trails B = 167±18.4) but has 39 heterozygous individuals (mean Trails B = 145.5 ±67.7) to compare with 3478 homozygous major allele individuals (mean Trails B = 129.3 ±56.1). rs8082018 in *ACOX1* has 2 homozygous minor allele individuals (Trails B = 167.0 ±18.4), 40 heterozygous individuals (mean Trails B = 145.7 ±66.8) and 3483 homozygous major allele individuals (mean Trails B = 129.3 ±56.1). While the minor allele frequency is low in each of these three SNPs, the three genotypes do show a dose-dependent effect with minor allele being associated with lower scores in the heterozygote genotypes, and even lower scores in the homozygote minor allele genotypes. Due to the similarity in the Trails B scores and allele frequencies at these two SNPs I went on to test the correlation between each of these SNPs and found that they are all highly correlated with each other ($r > 0.8$, $p<0.05$).

When the FDR correction was applied, none of these SNPs continued to show significant associations with Trails B times or 3MS scores. Including potential covariates (BMI, diabetes, education, systolic blood pressure and HDL cholesterol), as well as using censored normal regression models, identified the same three SNPs as significant

predictors for Trails B displaying only small differences in β-coefficients, but again, in each of these models the SNPs were no longer significant after using the FDR correction.

Recessive Effects Model on Trails B

The recessive effects model identified two additional SNPs in *ACOX1* with significant β-values when adjusted for only age, study site and the first four eigenvectors: rs6343 and rs3744033 (Appendix 6). However, after adjustments were made for the other cognitive risk factors that were also associated with Trails B (BMI, HDL cholesterol, systolic blood pressure, education, diabetes and anti-depressant use) the association with either of these SNPs did not remain significant.

Adjusting for the additional cognitive risk factors also identified, as significant at $p<0.05$ level, the same two similar SNPs from the additive effects models: rs9892372, and rs808218 although they were not found to be associated in the models only adjusted for age and principal components. This may not be a genuine association, due to the fact that at each of these SNPs only two individuals are homozygous for this minor allele, which is not enough to provide an accurate test. It is however, promising that it at least shows the same trend as the additive model. In these recessive effects models, using censored normal regression did not change the outcome. Applying the FDR correction did cause all SNPs to become insignificant predictors of Trails B in these models.

Additive Effects Model on 3MS

An Additive Effects Model was built on 3MS for each of 21 SNPs (Table 9 and Appendix 7). This model was adjusted for only age, study site and the four eigenvectors. In this model six SNPs were associated with 3MS scores, one in *MRPL38:* rs9892372, ($\beta=1.75$, $p=0.014$) and three in *ACOX1*: rs3643 ($\beta=0.46$, $p=0.029$), rs8082018 ($\beta=1.75$,

p=0.013), and rs7213998 (β=0.43, p=0.039). Because lower 3MS scores are associated with lower cognitive performance (as opposed to Trails B where higher scores are associated with lower cognitive performance), positive β-coefficients correspond to deleterious effects of a minor allele.

The adjustments for other cognitive risk factors did not change the significance of any SNPs in these models. Using censored normal regression also did not produce significantly different results in terms of which models significantly predicted 3MS scores. After using the FDR correction, no SNPs were significantly associated with 3MS scores in the additive models.

Furthermore, these SNPs, while associated with 3MS scores, are associated with very small actual changes in average 3MS scores. For instance, at rs8082018 those homozygous for the minor allele had a mean 3MS score of 92.0 ±2.8, while those with the heterozygous genotype had a mean 3MS score of 92.3 ±5.3 and individuals homozygous for the major allele had a mean 3MS score of 94.1 ±5.0. While these results may be statistically significant in a linear regression model, they may be clinically irrelevant, as these scores are all very similar and may indicate little difference in the real impact to an individual's daily life.

Recessive Effects Model on 3MS

Building recessive effects models on 3MS identified two additional SNPs in two genes as significantly associated with 3MS scores as well as identifying one SNP form the additive effects models with even higher predictive value (Appendix 8). rs7213548 in *FBF1* (β=3.31, p=0.009), and rs3643 (β=3.48, p=0.001) and rs11651351 (β=6.55, p<0.0001) in *ACOX1* were all found to be significantly associated with 3MS scores in the

model adjusting for age, study site and the four eigenvectors alone. Both SNPs in *ACOX1* remained significantly associated with 3MS after correction for multiple comparisons. After building multivariate models (adjusting for BMI, HDL cholesterol, systolic blood pressure, education, and diabetes), these same three SNPs are still significantly associated with 3MS scores. Each of these three SNPs have only a few individuals with the homozygous minor allele genotype, but their mean 3MS scores are noticeably different. For example rs3643 in *ACOX1* only has 20 individuals who are homozygous for the minor allele, but they have a mean 3MS score of 90.8 ± 6.5 compared to those who carry one of the major alleles who have a mean 3MS score of 94.1 ±5.03, and while these means are not different enough to be statistically significant, and may not even be clinically significant, a clear trend is observed. The SNP that is the most significantly associated with 3MS is rs11651351 in *ACOX1*. It only had 11 people who are homozygous at the minor allele and have an average 3MS score of 97.5 ±2.79 compared to those who carry the major allele who have a mean 3MS score of 94.1 ±4.77. The recessive effects model of rs3643 on 3MS was also significant after the FDR correction (p=0.14). These two SNPs were the only SNPs that were still significant after adjusting for all covariates in the multivariate model and using the FDR correction.

Interestingly, the adjustments for other cognitive decline risk factors in the recessive model on 3MS, yielded a noticeable change in some of the β-values associated with each significant SNP in its model. This effect was much less pronounced in either the additive or recessive effects models on Trails B, or in the additive models on 3MS. Further examination indicated that that education and systolic blood pressure had the greatest effect on decreasing the magnitude of the β-coefficient. Each of these two

covariates alone lowered the β-coefficient, and together they had an even greater effect. BMI and HDL cholesterol had very small effects on the β-coefficients in these models. As in the additive effects models, statistically significant associations corresponded to relatively small changes in 3MS scores that may have little clinical relevance.

Best-Fit Models

By regressing Trails B on age alone, we found that only 11.5% of the variation in Trails B is explained by age. When the four eigenvectors and the study site are included 13.2% of the variation in Trails B is explained. It would be useful to have a model that explained a greater portion of this variation. Towards this end, we attempted to select the best subset of SNPs to predict Trails B outcome. To do this, Mallow's C_p was calculated for various combinations of predictors. Each model included age, study site and the four principal components. Initially this analysis was performed in half of the group chosen at random. Using this approach the best model to predict Trails B did not include any SNPs and only included the required variables (age, eigenvectors and study site). The second best model included one SNP: rs8082018 in addition to age, study site and the four eigenvectors (C_p = 1.48, R=0.0973). When testing this model on the same half of the group, the β-value associated with this SNP was only marginally significant (p=0.053). This is likely due to the loss of power by only using half the dataset. Linear regression was then performed using this model in the second half of the dataset and similar results were observed. Using this model with the full dataset explained 13.5% of the variation in Trails B, barely more than without this SNP.

While it is interesting to know which SNP from this set is the best predictor of Trails B, this model does not provide a good way in which to predict which patients have

decreased cognitive performance. We know that this phenotype is multifaceted and likely controlled by many genes. Therefore, individuals SNPs are likely not going to be extremely useful for diagnostic tests. This shows that these genes are not alone a good set to use to develop a model that can explain a large amount of variation in cognition.

Residual Analysis

On each significant model, residual analysis was performed. A plot of residuals versus predicted values was created and assessed (Appendix 9). A pattern was observed in the graphs of all Trails B models. This was due to the spike in frequency for a time of 300 seconds. This was noted earlier as a deviation from normality in this measurement. We had attempted to use censored normal regression to account for this deviation, and it did not greatly change the results of the regression analyses. Graphs displaying leverage were used to identify outliers and additionally, using Cook's distance to assess influential points also generated a list of those people who had gotten 300 seconds as a Trails B time. Again, because the censored normal regression models that attempted to account for this phenomenon did differ greatly from the linear regression models, these results were not of concern. No other problems were observed in the residual analyses of these significant SNPs.

DISCUSSION

This study examined the relationship between six genes previously associated with WMH burden, *WBP2*, *TRIM47*, *TRIM65*, *MRPL38*, *FBF1* and *ACOX1*, and cognitive performance in elderly men. Trails B, a test of executive function, was used as the primary outcome, but 3MS, a measure of global cognition was also examined for associations with 26 SNPs in these six genes. This study found that SNPs in two genes,

MRPL38 and *ACOX1*, might be associated with executive function as measured by Trails B. Furthermore, SNPs in *MRPL38, ACOX1* and *FBF1* were found to be associated with global cognitive function as measured by 3MS. These models were adjusted for age, the largest known predictor of age-related cognitive impairment and study site, to account for potential confounding caused by difference in measurements. Adjusting models for principal components, other cognitive impairment risk factors or using a censored normal regression to account for non-normality in the outcome variables did not change the SNPs that were found to be associated, and had very little affect on the magnitude of the effects. Only one model, the recessive effects model of rs11651351 in *ACOX1* on 3MS score, remained significant after adjusting for covariates using a FDR correction to account for multiple testing. The recessive effects model of rs3643 in *ACOX1* and 3MS was also significant after applying the FDR correction as well as covariate adjustments, but had a p=0.051 when both the adjustments and FDR correction were applied. Overall, this study suggests that genes that predict white matter disease may also be associated with cognitive function in elderly men, but confirmation in larger cohorts, as well as in women, and a better understanding of these genes' functions are needed to support these findings.

Associations

Cognitive Risk Factors and Trails B and 3MS

This study confirmed that vascular risk factors such as HDL cholesterol, systolic blood pressure, BMI and diabetes are related to both tests of executive function and global cognition. As expected, age accounted for some of the variation seen between these measures and cognition, but significant relationships were still observed after

adjusting for age. Education level, a proxy measure of cognitive reserve, was highly associated with better Trails B and 3MS scores, even after adjusting for age. This population is highly educated with three quarters of men having attended at least some college and a quarter of them having finished graduate school. This selection bias is likely due to the selection bias resulting from men with greater education being more likely to volunteer for a study. Nevertheless, a clear association was observed between education level and cognitive measures of executive function and global cognition ($p<0.0001$ after age adjustment for both outcomes). This association between education level and AD has been shown consistently and is part of the cognitive reserve hypothesis that suggests that individuals with higher IQ, education or occupational complexity have decreased incidence and prevalence of AD.[55]

Depression is a known risk factor for cognitive decline,[56] but a direct measure of depression was not available in the MrOS dataset. Data was available on antidepressant use as well as use of the specific class of antidepressants, SSRIs. Antidepressant use was associated with increased Trails B times (indicating poorer performance), but it was not with 3MS scores, while SSRI use was not associated with either cognitive measure. It can be noted that antidepressant use may not accurately reflect those who experienced depression, as these drugs can be prescribed for other uses, or people may have experienced depression who did not seek treatment, or who used a non-pharmaceutical-based treatment. This likely misclassification would only cause failure to account for some of the variation in the outcome variables and should not affect the magnitude of any effects seen in this study, as it is not related to the genes themselves.

Genes Associated with Trails B and 3MS

Multiple SNPs were found in this 17q25 region that were related to either Trails B times or 3MS scores with nominally significant p values (<0.05) but most of these significant relationships were not present after FDR adjustment. These associated SNPs were identified in *ACOX1*, *MRPL38* and *FBF1*. No associations were found between either of the cognitive tests and any SNPs in *TRIM47, TRIM65* or *WBP2*.

MRPL38 showed one SNP, rs9892372, which was associated with Trails B and 3MS in additive models with unadjusted $0.006<p<0.03$ in all models but FDR adjusted $0.07<p<0.22$. This effect could not be reliably tested in the recessive models, as there were only two individuals homozygous for the minor allele. *MRPL38* encodes a protein in the mitochondrial ribosome subunit 39S. Mitochondrial dysfunction has been implicated in Alzheimer's disease pathology [57] as has oxidative stress [45] which has been shown to affect this specific subunit. [44] Both mitochondrial dysfunction and oxidative stress are areas of active research for AD therapies, and while the link between genes and these processes is still incomplete, these finding further support this area of research. *MRPL38* had only three SNPs to test in this cohort, all located in non-coding regions, so it is possible different associations could be observed with other known SNPs in this gene.

The association with *FBF1* was only seen in the recessive effects model on 3MS scores, with rs7213548 showing a significant association before the multiple comparison adjustment (p=0.009). Following the FDR correction this relationship was no longer significant (p=0.065). *FBF1* encodes for a protein that interacts with the Fas cell surface receptor, a regulator of apoptosis. Not only has apoptosis been implicated in AD [57], there is evidence of it occurring in white matter lesions, [58] suggesting some biologic

plausibility for this association, although further understanding of this gene's function is needed. By understanding the specific role of *FBF1* on apoptosis, therapies could be targeted to correctly regulate this process, and perhaps slow white matter lesion development.

SNPs in *ACOX1* showed associations with both Trails B and 3MS. rs8082018 was associated with both Trails B and 3MS in additive models before the FDR correction. Additionally, one more SNP (rs7213998) was associated with 3MS in the additive effects models and two more SNPs (rs3643 and rs11651351) were associated with 3MS in the recessive effects models before the FDR correction. rs3643 was still significant after the FDR correction in the age-adjusted model, but after the FDR correction in the multivariate model this relationship was no longer significant at a 0.05 level (p=0.051). rs11651351 was significantly associated with 3MS scores in the recessive model, and was the only relationship in this study to maintain its significance after accounting for FDR correction and all of the adjustments in the multivariate model (p=4.54 E-10). Because there were only eleven individuals homozygous for the recessive allele, a spurious association cannot be ruled out, but it is a relationship that at least warrants further study especially considering its magnitude (β=9.73 which represents almost a two standard deviation shift in 3MS scores).

The ACOX1 protein is responsible for the first, and rate-limiting, step of the very long chain fatty acid (VLCFA) beta-oxidation.[59] This process occurs in an organelle called the peroxisome when fatty acids are too long to be oxidized directly by the mitochondria.[59] Following this initial peroxisomal oxidation of VLCFA, fatty acids can undergo additional oxidation in the mitochondria and can subsequently be used to

generate energy in the Krebs cycle. ACOX1 deficiency leads to the accumulation of VLCFA in the plasma and tissues and results in a rare disease called pseudoneonatal adrenoleukodystrophy (P-NALD), which leads to neuroinflammation and neurodegeneration.[60] One characteristic of P-NALD is demyelination of axons, which, as discussed earlier, is thought to be a cause the white matter hyperintensities seen on MRI.[60] Therefore, functional mutations in the *ACOX1* gene could plausibly lead to abnormally high VLCFA levels and perhaps their associated neuroinflammation and degeneration which could be observed as WMHs. However, because this SNP is in a non-coding region of *ACOX1,* it is hard to tell its exact physiological affect on the protein for which this gene encodes, as it is likely only a marker near a functional variant, but its relationship with fatty acid beta-oxidation does provide biologic plausibility for the relationship between this gene and cognitive function.

Additionally, when ACOX1 catalyzes the first step in the long chain fatty acid beta-oxidation pathway, electrons are donated to molecular oxygen creating hydrogen peroxide, a strong oxidizing agent, as a byproduct.[42] Increased peroxisomal *ACOX1* expression is known to lead to oxidative DNA damage.[59] Oxidative stress has been repeatedly linked to AD pathology and may also be induced by beta amyloid as the disease progresses, suggesting the possibility of great benefit to AD patients if therapies could be targeted to reduce oxidative stress.[61]

Interestingly, a greater number of SNPs were associated with 3MS than with Trails B. We predicted that there would be stronger and more frequent associations with Trails B as it is a test of executive function, and WMHs have been specifically linked to executive function[3,4], and used this as our primary outcome measure in this study. Since

these genes were linked to WMHs by Fornage *et al.*[2] we hypothesized that they might also be associated with executive dysfunction, the cognitive phenotype linked to WMH accumulation. This relationship was observed between *MRPL38* and *ACOX1*, but these genes, and additionally *FBF1,* were also related to 3MS scores and the relationship existed for a greater number of SNPs. It is not surprising that there is overlap here, as both tests are cognitive tests that have been associated with aging and are therefore not entirely independent. This was demonstrated in this study when a small correlation was found between the two tests ($r=0.43, p<0.001$). 3MS is a broader test getting at generalized cognitive performance, so it might follow that a broader range of SNPs would be associated with this measure.

Study Limitations

The main limitations of this study were sample size, the population included only males, study design and *APOE* allele status was not available. From the MrOS cohort, 3552 subjects had genotype data available for analysis. Because many of the SNPs being studied here have very low minor allele frequencies, a larger study population is needed to be able to observe enough instances of the minor allele. Many of the SNPs had fewer than twenty individuals exhibiting the homozygosity for the minor allele, making it very hard to build dependable recessive effects models. Even with the additive effects models, some of the heterozygote genotypes were too small to have a great amount of confidence in their predictions. For instance, rs7208173 in *TRIM65* showed an association with Trails B (in models without an FDR correction), but because there were only eight individuals who carried at least one copy of the minor allele, and within these eight individuals there was high variability in their Trails B score, these results cannot be

interpreted as anything other than spurious in this cohort. Validation in another cohort for these candidate SNPs would be useful. For instance, Study of Osteoporotic Fractures (SOF) the sister study to MrOS, would be a good cohort in which these SNPs could be validated. A future study in a larger population may also help further elucidate if these relationships, or lack thereof, are spurious or accurate. The original GWAS study that identified these genes as related to WMHs was a meta-analysis of studies that included 9,361 individuals, and therefore, may have been more able to identify significant relationships with these genes.[2] Due to the biologic plausibility underlying the relationship between white matter abnormalities with both executive function and cognitive performance, it seems reasonable that the relationships observed between these genes and Trails B or 3MS are accurate, and further testing could be used to substantiate these findings.

In this study we were not able to control for *ApoE e*4 Allele status. This gene has been repeatedly and strongly linked to AD, and is one of the strongest predictors of its occurrence, after age with the e4 allele being associated with worse outcomes.[19] This cohort did not have the SNPs needed to determine the *ApoE*4 genotype in the study subjects. Because this gene is a strong independent predictor of cognitive decline, it may be necessary, like age, to account for it in models used to predict a cognitive outcome. This was not possible in this population and represents a limitation to this study. Because *APOE* is on chromosome 19, it is likely not in LD with the genes being tested in this study, and is less likely to be a confounder. Hopefully, this means this adjustment would not greatly change the results, but it could still be an important covariate, and these results would be stronger if it could be considered.

The biggest limitation of this analysis was study design. Because this was a cross-sectional study, change over time could not be examined, which is necessary to really observe cognitive deficit. It is possible that someone with low cognitive functioning, who maintains that same low function over time, will appear to have low cognitive measures when they're not experiencing dementia-related symptoms. Additionally, someone who had very high cognitive functioning, but has experienced some decline, may appear just average when taking a cross-sectional view. These two cases could create a scenario where some who is in the process of experiencing decline would appear cognitively healthier than someone who is maintaining their cognitive ability. When studying dementia, the outcome of interest is really *decrease* in cognitive ability, not simply how subjects perform at baseline. Because of this, the ideal study design would be a longitudinal study that examines change in cognitive scores over time in relation to genes.

Finally, because this study was only conducted in males of Caucasian descent, it cannot be generalized to other populations. Conducting this analysis in women would be useful counterpart, as well as in other populations with a broader range of ancestry. A larger, more diverse sample would improve generalizability.

Potential Confounders and Biases

Confounding

Because genotypes are fixed and cannot be influenced by any known factors, confounding is unlikely to be influencing these results. The only thing that could be causing confounding is ancestral patterns within the population. This, however, was accounted for through the use of principal component analysis and adjustment for the

first four eigenvectors. The models with and without the inclusion of eigenvectors were very similar, so they were likely not needed to account for confounding. By definition, a confounder must be associated with the primary predictor and also the primary outcome while remaining clear from the biological causal pathway between the predictor and outcome, and in this case no other factor is able to change an individual's genetic code. There are, however, potential sources of bias that could be influencing the results of this study.

Selection Bias

Selection bias was present in this data set as all men were community dwelling at baseline, suggesting that they were primarily cognitively intact at baseline. As discussed above, executive function can be impaired before clinical dementia symptoms are observed, but in general this group of men were highly functioning cognitively. If these genes truly are associated with low cognitive function, those who are affected might not be represented in the cohort of relatively healthy volunteers. There still existed variation in cognitive performance against which to test the effects of these genes, but having subjects with a wider range of cognitive abilities may further help to elucidate this potential relationship.

Another source of selection bias may be present in the education levels of the men in this cohort. It is a highly educated group of men, which over half processing an undergraduate degree, and over a quarter possessing a graduate degree. As mentioned above this is likely due to the fact that higher educated individuals and also may place more value on participating in a study such as MrOS resulting in a group of men that does not closely match the general population. We know that greater cognitive reserve

has been linked to higher cognitive performance in the elderly, this analysis could be experiencing bias towards the null, which men who would otherwise be experiencing a lowered cognitive performance due to a specific genotype, may have those effects masked by having greater cognitive reserve. While we tried to account for this by adjusting for level of education, this measure is only a proxy for cognitive reserve, and may not have completely eliminated the effect of this type of bias.

Finally, only a subset of men within this cohort was used for this analysis. Not only was the dataset limited to Caucasian men due to genotype data availability, some men had not consented for genotype analysis related to the study of cogitation. This caused one whole study site, Birmingham, Alabama, to be excluded from this analysis. Summary statistics for the MrOS cohort showed that for 3MS the average score for the entire cohort was 93.3 ± 5.9, whereas in my cohort average 3MS score was 94.1 ± 5.0. Trails B time for the entire cohort was 134.5 ± 58.9 seconds whereas in my cohort the average Trails B time was 129.5 ± 56.3. For each of these measures scores were very similar, with very slightly "healthier" scores in my dataset. These differences are very small, well within one standard deviation of each other, and, because it was an entire study site being excluded, those who were not included in the dataset were not systematically excluded based on their exposure or outcome, so this should not affect the results of this analysis.

Information Bias in the Exposure

A limitation of this study, and many genetic association studies, is that the SNPs tested are not expected to be causal. Most are intronic or in untranslated regions, with only few exceptions (rs7213548 in *FBF1* is responsible for a missense mutation changing

Cys to Ser [42]). Instead, these SNPs mark a genomic region that might contain causative variants. This means that there is known misclassification in the exposure but, because we do not know which SNPs in this region are the causative variants, this misclassification cannot be avoided. However, it is still useful to identify SNPs that may mark a functional gene or gene region, but it must still be noted that we are not testing the functionality of these genes directly.

In any study, there is a chance for inaccurate measurements leading to information bias in the exposure. Inaccurate genotyping may have occurred, but since it is likely to be random and not systematic it is unlikely to have had a large effect on our associations identified in this large group. Futhermore, genotype data was subjected to stringent quality control as described above.

Information Bias in the Covariates

Covariates that were clinically measured at baseline are likely more reliable than the self-reported measures due to recall ability. Cholesterol measurements, blood pressure, anti-depressant use and diabetes were all measured at baseline, the same visit in which Trails B and 3MS tests were administered. Self-reported measures are slightly less reliable, such as stroke and hypertension status. It is hard to postulate how variation in any of these measures would affect the resulting data, and while measurement error can be a concern, because these are all covariates, and not primary predictors, hopefully any minor error did not contribute significantly to the results of this study.

Information Bias in the Outcome

Because Trails B and 3MS were measured variables they could be subject to interviewer bias, but since this is unrelated to a subjects genotype, it is unlikely to have

any affect in the study outcomes. Additionally, variation could occur by study site, but this is thought to arise from true differences in cognition at each site rather than measurement error since strict protocols were used to reduce this in MrOS. Regardless of the reason, Trails B and 3MS scores were significantly different by study site in this cohort. To account for this, the study site variable was included in each of the analysis, so each model includes adjustment for location.

While measurements of Trails B and 3MS likely did not introduce a large amount of misclassification error into this study, their representation of cognitive phenotypes may have be a source of non-differential misclassification bias. Because it is very difficult to differentiate between AD and VD they were group together conceptually for the purposes of this study. The two cognitive tests, Trails B and 3MS, did not distinguish between AD and VD at all (or any other cause of cognitive impairment). AD and VD may have great overlap and be controlled by the same genes, but they also maybe controlled by a different set of genes, or two overlapping sets of genes. Using these two overlapping phenotypes in genetic association studies introduces some ambiguity into the results, as associated genes may be associated to one or both specific outcomes. Because the genes that were chosen for this analysis were related to white matter, we hypothesized that they were also related to Trails B as a text of executive function. WMH is a vascular pathology, but executive functioning may be caused by multiple pathologies, such as other forms of cognitive decline that are not vascular disease mediated. This means that the outcome being measured by Trails B is somewhat ambiguous. Because there may be multiple pathways by which to reach decline in executive functioning, it would make the results much stronger if it were possible measure if subjects had decreased executive

functioning due to vascular disease or due to some other pathology using MRI, but this measure was outside the scope of this study. If present, this bias would be responsible for underestimating the true association between SNPs and vascular-mediated decline in executive functioning, causing bias towards the null and potentially missing some true associations, and therefore should not invalidate the results of this study.

Overall Validity of Data

While multiple potential sources of bias exist, the single greatest source seems to be that all individuals were healthy at the time this data was collected. This means that there was little variation in the data to be explained by the SNPs of interest creating large selection bias. Also, the ambiguity of the phenotype may also be masking a true association between SNPs and a more specific phenotype. Because both of these situations would cause bias towards the null, it means that, if anything, the SNPs suggested by this study, could have a stronger association with cognitive measures in a more variable population with more directly measured phenotypes. Therefore, there were no sources of bias that should automatically make this data invalid, but it is necessary to address these questions in a larger, more diverse elderly population. Additionally, a different study design where longitudinal data is used and incident cognitive decline can be observed would also help to observe a greater variation in genotypes and substantiate associations found here. MrOS does have data available for both 3MS and Trails B at times after the initial visit, so this analysis could be performed in a future study.

Public Health Implications

AD is not only a leading cause of death in the elderly, but it is also responsible for a sharp decline in quality of life in those who are affected and has a tremendous burden

on caregivers and health care economics. Because strategies to prevent or slow the disease are unavailable, more research in this area is needed to discover new treatments. Recently much work has been put forth to explain how "vascular risk factors" can modulate risk for age related dementia incidence. Because many vascular risk factors (diabetes, cholesterol and blood pressure) are modifiable, targeting these may be useful in helping to delay the onset of Alzheimer's. However, caution still needs to be taken in making global recommendations in this field. For example, lowering blood pressure in later ages has been associated with worse brain outcomes some subjects.[62] Identifying genes that link cognitive decline to other measures of vascular health will permit the discovery of new disease mechanisms or phenotypes where preventative strategies can be developed. Additionally, prevention strategies will need to be implemented along with prediction strategies to stay ahead of the curve of aging elders at risk for dementia. Once both of these are in place, those that can be identified as more likely to develop age-related cognitive impairment could take measures to reduce their risk thereby preventing or delaying onset of cognitive dysfunction.

Ethical Concerns

There are various ethical problems associated with the study of genetic determinants of any disease, but specifically cognitive impairment as it is not currently treatable. When predictive genes are discovered, researchers may use them to gain further understanding of the diseases and their pathology, leading to potential treatments and/or cures. While these advances are biologically useful, there is typically a lag before they are clinically useful. Being able to identify a genetic risk factor in a patient brings up a host of ethical concerns if there is no prevention or treatment available to him or her.

Whether or not to test patients for genetic risk factors for age-related cognitive impairment is an important ethical question that clinicians now face, because it is unclear if patients should know about a predisposition to a disease they cannot treat or prevent.

Additionally, when testing subjects for research studies, it must be determined whether or not patients should know what their test results are. If studies such as these yield genes that may be associated to cognitive outcomes, researchers must decide whether or not to inform the subjects of their genetic status. Again, this is a difficult question because subjects may not be able to do anything with the information if there is no treatment available. However, if this motivates individuals to take necessary steps to improve their lifestyle and overall health then perhaps this type of information can have clinical utility on population health. As we learn more about the genetic profile of individuals who develop cognitive decline, or any genetic disease, these questions become increasingly important to answer, and clinicians and researchers must carefully consider these important issues.

CONCLUSIONS

In conclusion, this study identified three genes related to cognitive measures in elderly men. Only the association between *ACOX1* and 3MS remained significant after a FDR adjustment was used to account for multiple comparisons. It was initially hypothesized that these genes would be related to executive function as measured by Trails B, due to their associations with WMH burden. Although two of these six genes were associated with Trails B, these associations did not remain significant after adjustment for multiple comparisons. One SNP in *ACOX1* was significantly associated with 3MS after adjustments. This may be due to the small sample size resulting in few

occurrences of the minor alleles in the SNPs being tested. White matter burden has been linked to decreased cognitive function (including executive function) as well as to these six genes, so the relationships observed in this study may not be spurious, but further studies are needed to verify these relationships.

References

1. Gorelick PB. Scuteri A. Black SE. Decarli C. Greenberg SM. Iadecola C. Launer LJ. Laurent S. Lopez OL. Nyenhuis D. Petersen RC. Schneider JA. Tzourio C. Arnett DK. Bennett DA. Chui HC. Higashida RT. Lindquist R. Nilsson PM. Roman GC. Sellke FW. Seshadri S. American Heart Association Stroke Council, Council on Epidemiology and Prevention, Council on Cardiovascular Nursing, Council on Cardiovascular Radiology and Intervention,and Council on Cardiovascular Surgery and Anesthesia. Vascular contributions to cognitive impairment and dementia: A statement for healthcare professionals from the american heart association/american stroke association. *Stroke*. 2011;42(9):2672-2713.

2. Fornage M, Debette S, Bis JC, et al. Genome-wide association studies of cerebral white matter lesion burden: The CHARGE consortium. *Ann Neurol*. 2011;69(6):928-939.

3. Prins ND, van Dijk EJ, den Heijer T, et al. Cerebral small-vessel disease and decline in information processing speed, executive function and memory. *Brain*. 2005;128(Pt 9):2034-2041.

4. Zheng JJJ, Lord SR, Close JCT, et al. Brain white matter hyperintensities, executive dysfunction, instability, and falls in older people: A prospective cohort study. *J Gerontol A Biol Sci Med Sci*. 2012.

5. Arbuthnott K, Frank J. Trail making test, part B as a measure of executive control: Validation using a set-switching paradigm. *Journal of Clinical & Experimental Neuropsychology: Official Journal of the International Neuropsychological Society*. 2000;22(4):518-528.

6. Teng EL, Chui HC. The modified mini-mental state (3MS) examination. *J Clin Psychiatry*. 1987;48(8):314-318.

7. Alzheimer's disease facts and figures. *Alzheimer's & Dementia*. 2012;8(2):14.

8. Jellinger KA, Attems J. Neuropathological evaluation of mixed dementia. *J Neurol Sci*. 2007;257(1-2):80-87.

9. Hebert LE, Beckett LA, Scherr PA, Evans DA. Annual incidence of alzheimer disease in the united states projected to the years 2000 through 2050. *Alzheimer Disease & Associated Disorders*. 2001;15(4):169-173.

10. Moorhouse P, Rockwood K. Vascular cognitive impairment: Current concepts and clinical developments. *Lancet Neurology*. 2008;7(3):246-255.

11. Birks J. Cholinesterase inhibitors for alzheimer's disease. *Cochrane Database of Systematic Reviews*. 2006(1):005593.

12. Vickrey BG, Mittman BS, Connor KI, et al. The effect of a disease management intervention on quality and outcomes of dementia care: A randomized, controlled trial. *Ann Intern Med*. 2006;145(10):713-726.

13. Chan RC, Shum D, Toulopoulou T, Chen EY. Assessment of executive functions: Review of instruments and identification of critical issues. *Archives of Clinical Neuropsychology*. 2008;23(2):201-216.

14. What is executive function. http://www.ncld.org/ld-basics/ld-aamp-executive-functioning/basic-ef-facts/what-is-executive-function. Updated 2012. Accessed 07/02, 2012.

15. Clark LR, Schiehser DM, Weissberger GH, Salmon DP, Delis DC, Bondi MW. Specific measures of executive function predict cognitive decline in older adults. *Journal of the International Neuropsychological Society*. 2012;18(1):118-127.

16. Martyr A, Clare L. Executive function and activities of daily living in alzheimer's disease: A correlational meta-analysis. *Dement Geriatr Cogn Disord*. 2012;33:189-203.

17. Jack CR,Jr, Knopman DS, Jagust WJ, et al. Hypothetical model of dynamic biomarkers of the alzheimer's pathological cascade. *Lancet Neurology*. 2010;9(1):119-128.

18. Heston L, Mastri A, Anderson V, White J. Dementia of the alzheimer type. *Arch Gen Psychiatry*. 1981;38:1085-1090.

19. Jun G, Vardarajan B, Buros J, et al. Comprehensive search for alzheimer disease susceptibility loci in the APOE region. *Arch Neurol*. 2012;online.

20. Bertram L, McQueen M, Mullin K, Blacker D, Tanzi R. The AlzGene database. Alzheimer Research Forum Web site. http://www.alzgene.org. Accessed 11/12, 2012.

21. Tanzi R, Bertram L. Twenty years of the alzheimer's disease amyloid hypothesis: A genetic perspective. *Cell*. 2005;120(4):545-555.

22. Daw EW, Payami H, Nemens EJ, et al. The number of trait loci in late-onset alzheimer disease. *Am J Hum Genet*. 2000;66(1):196-204.

23. Bertram L, Tanzi RE. The genetics of alzheimer's disease. *Progress in Molecular Biology & Translational Science*. 2012;107:79-100.

24. Reitz C, Brayne C, Mayeux R. Epidemiology of alzheimer disease. *Nature Reviews Neurology*. 2011;7(3):137-152.

25. Debette S, Seshadri S, Beiser A, et al. Midlife vascular risk factor exposure accelerates structural brain aging and cognitive decline. *Neurology*. 2011;77(5):461-468.

26. Scahill RI, Schott JM, Stevens JM, Rossor MN, Fox NC. Mapping the evolution of regional atrophy in alzheimer's disease: Unbiased analysis of fluid-registered serial MRI. *Proc Natl Acad Sci U S A*. 2002;99(7):4703-4707.

27. Yue NC, Arnold AM, Longstreth WT,Jr, et al. Sulcal, ventricular, and white matter changes at MR imaging in the aging brain: Data from the cardiovascular health study. *Radiology*. 1997;202(1):33-39.

28. de Leeuw FE, de Groot JC, Achten E, et al. Prevalence of cerebral white matter lesions in elderly people: A population based magnetic resonance imaging study. the rotterdam scan study. *Journal of Neurology, Neurosurgery & Psychiatry*. 2001;70(1):9-14.

29. Decarli C. Clinically asymptomatic vascular brain injury: A potent cause of cognitive impairment among older individuals. *J Alzheimers Dis* [Epub ahead of print]. 2012.

30. Kearney-Schwartz A, Rossignol P, Bracard S, et al. Vascular structure and function is correlated to cognitive performance and white matter hyperintensities in older hypertensive patients with subjective memory complaints. *Stroke*. 2009;40(4):1229-1236.

31. Schmidt R, Grazer A, Enzinger C, et al. MRI-detected white matter lesions: Do they really matter? *J Neural Transm*. 2011;118(5):673-681.

32. Knopman DS, Penman AD, Catellier DJ, et al. Vascular risk factors and longitudinal changes on brain MRI: The ARIC study. *Neurology*. 2011;76(22):1879-1885.

33. Dufouil C, Chalmers J, Coskun O, et al. Effects of blood pressure lowering on cerebral white matter hyperintensities in patients with stroke: The PROGRESS

(perindopril protection against recurrent stroke study) magnetic resonance imaging substudy. *Circulation*. 2005;112(11):1644-1650.

34. Debette S, Markus HS. The clinical importance of white matter hyperintensities on brain magnetic resonance imaging: Systematic review and meta-analysis. *BMJ*. 2010;341:3666.

35. Debette S, Beiser A, DeCarli C, et al. Association of MRI markers of vascular brain injury with incident stroke, mild cognitive impairment, dementia, and mortality: The framingham offspring study. *Stroke*. 2010;41(4):600-606.

36. Silbert LC, Howieson DB, Dodge H, Kaye JA. Cognitive impairment risk: White matter hyperintensity progression matters. *Neurology*. 2009;73(2):120-125.

37. Verhaaren BF, de Boer R, Vernooij MW, et al. Replication study of chr17q25 with cerebral white matter lesion volume. *Stroke*. 2011;42(11):3297-3299.

38. Meroni G, Diez-Roux G. TRIM/RBCC, a novel class of 'single protein RING finger' E3 ubiquitin ligases. *Bioessays*. 2005;27(11):1147-1157.

39. Brown WR, Moody DM, Challa VR, Thore CR, Anstrom JA. Apoptosis in leukoaraiosis lesions. *J Neurol Sci*. 2002;203-204:169-171.

40. Gulyas B, Brockschnieder D, Nag S, et al. The norepinephrine transporter (NET) radioligand (S,S)-[18F]FMeNER-D2 shows significant decreases in NET density in the human brain in alzheimer's disease: A post-mortem autoradiographic study. *Neurochem Int*. 2010;56(6-7):789-798.

41. The Jackson Laboratory. Mouse phenome database web site, the jackson laboratory. http://phenome.jax.org. Updated 2012. Accessed November/18, 2012.

42. National Center for Biotechnology Information. Gene Web site. http://www.ncbi.nlm.nih.gov/gene. Published 2012. Updated 2012. Accessed 09/25, 2012.

43. Montine J, Morrow J. Fatty acid oxidation in the pathogenesis of alzheimer's disease. *Am J pathol*. 2005;166(5):1283-1283-1289.

44. Soreghan BA, Lu BW, Thomas SN, et al. Using proteomics and network analysis to elucidate the consequences of synaptic protein oxidation in a PS1 + AbetaPP mouse model of alzheimer's disease. *J Alzheimer's Dis*. 2005;8(3):227-241.

45. Venkateshappa C, Harish G, Mahadevan A, Srinivas Bharath MM, Shankar SK. Elevated oxidative stress and decreased antioxidant function in the human hippocampus and frontal cortex with increasing age: Implications for neurodegeneration in alzheimer's disease. *Neurochem Res*. 2012;37(8):1601-1614.

46. Blank JB, Cawthon PM, Carrion-Petersen ML, et al. Overview of recruitment for the osteoporotic fractures in men study (MrOS). *Contemporary Clinical Trials*. 2005;26(5):557-568.

47. Orwoll ES, Blank JB, Connor-Barrett E, et al. Design and baseline characteristics of the osteoporotic fractures in men (MrOS) study - A large observational study of the determinants of fracture in older men. *Contemporary Clinical Trials*. 2005;26(5):569-569-585.

48. Bassuk SS, Murphy JM. Characteristics of the modified mini-mental state exam among elderly persons. *J Clin Epidemiol*. 2003;56(7):622-628.

49. Manolio TA. Genomewide association studies and assessment of the risk of disease. *N Engl J Med*. 2010;363(2):166-176.

50. Illumina. Illumina BeadStudio analysis software module. Illumina Web site. http://www.illumina.com/Documents/products/datasheets/datasheet_beadstudio.pdf. Updated 2010. Accessed 1/12, 2013.

51. The HapMap Consortium. The international HapMap project. *Nature*. 2003;426:789-789-796.

52. 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(8):904-909.

53. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J R Statist Soc*. 1995;57(1):289-289-300.

54. Curran-Everett D. Multiple comparisons: Philosophies and illustrations. *American Journal of Physiology - Regulatory Integrative & Comparative Physiology*. 2000;279(1):R1-8.

55. Meng X, D'Arcy C. Education and dementia in the context of the cognitive reserve hypothesis: A systematic review with meta-analyses and qualitative analyses. *PLoS ONE [Electronic Resource]*. 2012;7(6):e38268.

56. Yaffe K, Blackwell T, Gore R, Sands L, Reus V, Browner WS. Depressive symptoms and cognitive decline in nondemented elderly women: A prospective study. *Arch Gen Psychiatry*. 1999;56(5):425-430.

57. Crews L, Masliah E. Molecular mechanisms of neurodegeneration in alzheimer's disease. *Hum Mol Genet*. 2010;19(R1):R12-20.

58. Brown WR, Moody DM, Challa VR, Thore CR, Anstrom JA. Apoptosis in leukoaraiosis lesions. *J Neurol Sci*. 2002;203-204:169-171.

59. Varanasi U, Chu R, Chu S, Espinosa R, LeBeau MM, Reddy JK. Isolation of the human peroxisomal acyl-CoA oxidase gene: Organization, promoter analysis, and chromosomal localization. *Proc Natl Acad Sci U S A*. 1994;91(8):3107-3111.

60. El Hajj HI, Vluggens A, Andreoletti P, et al. The inflammatory response in acyl-CoA oxidase 1 deficiency (pseudoneonatal adrenoleukodystrophy). *Endocrinology*. 2012;153(6):2568-2575.

61. Sultana R, Butterfield DA. Role of oxidative stress in the progression of alzheimer's disease. *J Alzheimer's Dis*. 2010;19(1):341-353.

62. Whitmer RA, Sidney S, Selby J, Johnston SC, Yaffe K. Midlife cardiovascular risk factors and risk of dementia in late life. *Neurology*. 2005;64(2):277-281.

63. Ashford JW, Mortimer JA. Non-familial alzheimer's disease is mainly due to genetic factors. *J Alzheimer's Dis*. 2002;4(3):169-177.

64. Wilson RS, Barral S, Lee JH, et al. Heritability of different forms of memory in the late onset alzheimer's disease family study. *J Alzheimer's Dis*. 2011;23(2):249-255.

65. Lee JH, Flaquer A, Stern Y, Tycko B, Mayeux R. Genetic influences on memory performance in familial alzheimer disease. *Neurology*. 2004;62(3):414-421.

66. Swan GE, Carmelli D. Evidence for genetic mediation of executive control: A study of aging male twins. *Journals of Gerontology Series B-Psychological Sciences & Social Sciences*. 2002;57(2):P133-43.

67. Mayeux R, Stern Y. Epidemiology of alzheimer disease. *Cold Spring Harb Perspect Med*. 2012;2(a006239).

68. Solomon A, Kivipelto M, Wolozin B, Zhou J, Whitmer RA. Midlife serum cholesterol and increased risk of alzheimer's and vascular dementia three decades later. *Dementia & Geriatric Cognitive Disorders*. 2009;28(1):75-80.

69. Weuve J, Kang JH, Manson JE, Breteler MM, Ware JH, Grodstein F. Physical activity, including walking, and cognitive function in older women. *JAMA*. 2004;292(12):1454-1461.

70. Erickson KI, Voss MW, Prakash RS, et al. Exercise training increases size of hippocampus and improves memory. *Proc Natl Acad Sci U S A*. 2011;108(7):3017-3022.

71. Bertoni-Freddari C, Fattoretti P, Casoli T, et al. Neuronal apoptosis in alzheimer's disease: The role of age-related mitochondrial metabolic competence. *Ann N Y Acad Sci*. 2009;1171:18-24.

Figures

Figure 1: Approximate progression of dementias

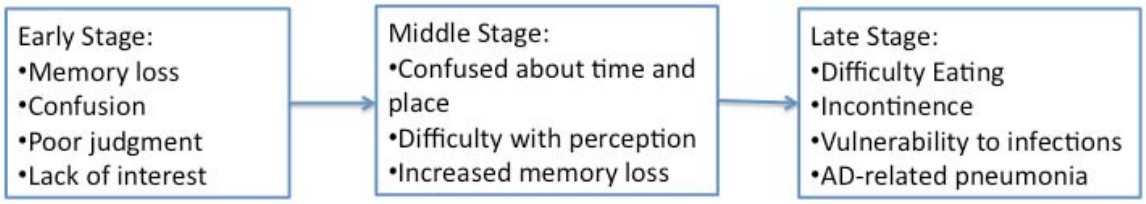
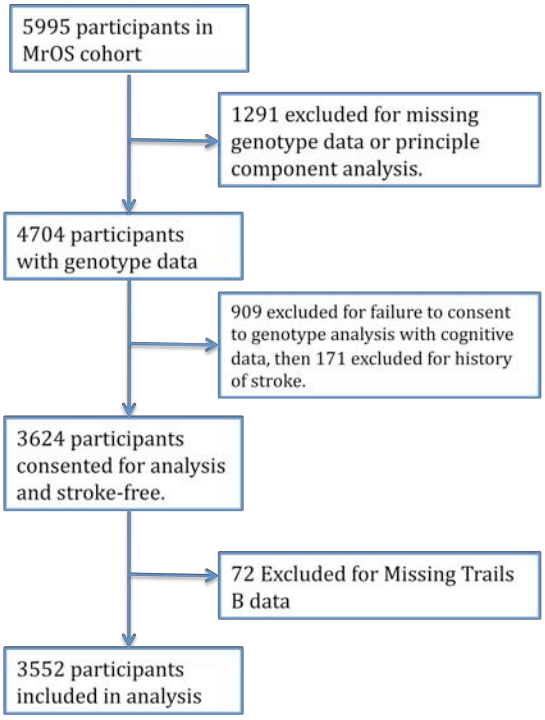


Figure 2: Exclusion criteria for analytic dataset



Tables

Table 1: Heritability of various cognitive measures

Phenotype	h^2	Citation
Alzheimer's in Swedish Twins	0.74	Ashford et al [63]
Alzheimer's in Norwegian Twins	0.61	Ashford et al [63]
Episodic memory	0.49 (p<0.0001)	Wilson et al [64]
Semantic memory	0.32 (p<0.0001)	Wilson et al [64]
Working memory	0.34 (p<0.0001)	Wilson et al [64]
MMSE	0.25 (p<0.0000001)	Lee et al [65]
Executive function in Twins	0.79	Swan et al. [66]
Trails B	0.5	Swan et al. [66]
Total recall	0.32 (p<0.0000001)	Lee et al [65]

MMSE=Mini-mental state examination
h^2=estimate of heritability (estimated proportion of variation between individuals in a population that is influenced by genetic factors)

Table 2: Common risk and protective factors for cognitive decline

Factor	Direction of effect	Description
Age	Increased risk	Over the age of 65 there is an exponential increase in occurrence of VD. [1] Age is the strongest predictor for cognitive decline.
APOE e*4*	Increased risk	ApoE e4 allele is associated with increased risk of AD and cardiovascular risk factors. [1] Individuals carrying one or more copy of *APOE* e4 have 3.68 times the odds (95% CI 3.30 – 4.11) of developing AD than non-carriers. [19]
Head trauma LOC and *APOE4*	Increased risk	Loss of Consciousness and traumatic brain injury lead to increased effect size in *ApoE* e4 carriers, meaning that those with the e4 allele and LOC injury are at even greater risk than those simply carrying the e4 allele. [67]
Hypertension	Increased risk	HTN prior to being 65 (defined as dbp over 90mmHg or sbp over 140mmHg) is a risk factor for late-life cognitive decline. Hazard ratio = 1.24 (95% CI 1.04-1.48). High SBP has been associated with greater late-life cognitive decline. Effects of diastolic blood pressure are unclear for late life hypertension. [62,67]
Type II	Increased risk	A meta-analysis of risk of type II diabetes on AD found

Diabetes		a relative risk of 1.54 (95% CI 1.33-1.79). Also, a longer duration of diabetes is associated with poorer cognitive function. [1,67]
High Body Weight	Increased risk	Both high and low BMI have been associated with AD, but the strongest effect is seen between obesity and AD. Meta-analysis between obesity and AD risk yields an odds ratio of 1.59 (95% CI 1.02-2.5). [67]
Depression	Increased risk	People with 3-5 depressive symptoms compared with 0-2 symptoms had increased cognitive decline. Adjusted Odds ratio of 1.6 (95% CI 1.2 – 2.1). [56]
Cholesterol	Increased risk	High cholesterol in midlife predicts cognitive impairment in late life. Results are more ambiguous for cholesterol measured concurrently with cognition. High midlife cholesterol was associated with an increased risk of AD 3 decades later. HR = 1.57 (95% CIL 1.23-2.01). [68]
Cognitive Reserve	Protective	Education and leisure activities are protective against AD onset. Increased brain use resulting from these activities is referred to as "cognitive reserve". Meta-analysis of high cognitive reserve compared to low reserve gives an odds ratio of 0.54 (95% CI 0.49-0.59). [67]
Diet	Protective	Diets high in Vitamins B, C, D and E and the Omega-3 fatty acids are associated with decreased risk of cognitive decline and AD dementia. High adherence to self-reported Mediterranean diet is associated with decreased risk of cognitive decline and AD dementia (HR=0.85, CI 0.42-0.87) compared with those in the lowest. [67]
Exercise	Protective	Long-term, regular exercise is strongly associated with lower levels of cognitive decline. Exercise has also been shown to be protective when combined with a Mediterranean diet. Walking 4 hours per week is associated with less cognitive decline. Exercising 3 times a week for 30 minutes can increase the size of the hippocampus compared to non-aerobic stretching exercise over 1-year. [1,67,69,70]

Table3: Function of genes associated with WMH burden [42]

Gene	Function
WBP2	This gene encodes a WW domain binding protein, meaning that it binds another protein at its WW, a specific structure present in a variety of proteins. Increased *WBP2* expression has been associated with decreased expression of sodium-dependent noradrenaline transporter in mice, while decreased noradrenaline transporters have been associated with AD in humans. [40-42]
TRIM65 TRIM47	This gene codes for a protein that is member of a superfamily of proteins (with TRIM47) involved in biological processes including innate immunity, apoptosis, cell cycle regulation, vesicular trafficking and neuroprotection. Increased apoptosis leading to neuronal death and decreased neuronal density has been associated with Alzheimer's disease. [2,71]
MRPL38	This gene encodes a protein in the mitochondrial ribosome subunit 39S. 39S was found to be an oxidatively modified protein, and oxidative stress has been associated with Alzheimer's disease. [42,44,45]
FBF1	This gene codes for a protein that interacts with the Fas cell surface receptor, a regulator of apoptosis, which has been implicated in AD pathologies. [2,71]
ACOX1	This gene encodes a protein that is the first enzyme of the fatty acid beta-oxidation pathway. Oxidation products of fatty acids have been associated with neurodegeneration. [42,43]

Table 4: Pruned SNPs and their relationship to analyzed SNPs

Gene	SNP in Analysis	SNPs in LD not in analysis	r2
TRIM47	rs3744017	rs990862	0.8439
		rs1105917	0.8645
TRIM65	rs1551619	rs34974290	0.8535
TRIM65	rs3760128	rs3744021	0.8762
FBF1	rs7213548	rs7218738	0.9670
		rs3744002	0.9964
FBF1	rs2608881	rs9891076	0.8338
		rs2305913	0.8198
		rs1135889	0.8627
		rs2608882	0.8610
ACOX1	rs3643	rs7217955	0.8918
ACOX1	rs3682	rs10852766	0.9967
ACOX1	rs8082018	rs806546	0.9879
ACOX1	rs12603572	rs3744035	0.9701
ACOX1	rs7213998	rs7218656	0.9796

Table 5: Population and clinical characteristics

	mean	SD
Trail Making Test B score	129.5	56.3
3MS score	94.1	5.0
Age (years)	74.0	5.9
BMI (kg/m^2)	27.3	3.8
Total Cholesterol (mg/dL)	192.9	3.8
LDL Cholesterol (mg/dL)	113.9	30.4
HDL Cholesterol (mg/dL)	94.5	14.9
Systolic Blood Pressure (mmHg)	141.4	9.2
	N	%
Self-reported Hypertension	1453	41.2
Anti-Depressant Use	181	5.5
SSRI Use, n	78	2.3
Diabetes		
Normoglycemic	1636	46.4
Impaired Fasting Glucose	1174	33.3
Diabetes	487	13.8
Education		
Some Elementary School	2	0.06
Elementary School	54	1.53
Some High School	131	3.71
High School	613	17.37
Some College	789	22.35
College	639	18.1
Some Graduate School	395	11.19
Graduate School	907	25.69
Site:		
Minneapolis	722	20.3
Palo Alto	701	19.7
Pittsburgh	847	23.9
Portland	664	18.7
San Diego	618	17.4

Table 6: Mean Trails B scores by covariates

	r (p-value)	r Age-adjusted (p-value)	Q1	Q2	Q3	Q4	K-W p-value
Age, years			under 70	70-73	74-78	above 78	
Mean Trails B	0.3389 (<0.0001)		107.3	121.2	138.4	154.3	<0.0001
BMI, kg/m^2			under 24.7	24.7-26.7	26.8-29.4	above 29.4	
Mean Trails B	-0.01123 (0.5048)	0.0549 (0.0019)	131.96	125.6	130.02	130.35	0.137
Total Cholesterol, mg/dL			under 169	169-191.9	192-215	above 215	
Mean Trails B	-0.569 (0.0011)	-0.01040 (0.5567)	132.7	132.6	130.2	124.0	0.001
LDL Cholesterol, mg/dL			under 94	94-113	114-134	above 134	
Mean Trails B	-0.0583 (0.0008)	-0.01999 (0.2587)	134.7	130.2	129.0	125.1	6E-04
HDL Cholesterol, mg/dL			under 40	41-47	48-57	above 57	
Mean Trails B	-0.0323 (0.0642)	-0.0023 (0.0023)	133.8	131.3	123.6	130.4	2E-04
Systolic BP, mm/Hg			under 131	131-140	141-152	above 152	
Mean Trails B	0.0964 (<0.0001)	0.3800 (0.0318)	123.1	127.9	131.8	134.8	<0.0001
	Mean Trails B	SD	K-W Test p-value	Age-adjusted p-value			
Education			<0.0001	<0.0001			
Elementary School or Less	203.3	73.2					
Some High School	179.6	68.4					
High School or Some College	137.6	57.4					
College	125.4	53.5					
All/Some Grad School	115.4	47.4					
Site							
Minneapolis	120.0	55.3	<0.0001	<0.001			
Palo Alto	124.0	51.2					
Pittsburgh	128.2	58.9					
Portland	129.0	54.8					
San Diego	124.2	38.9					
Diabetes			<0.0001	<0.0001			
No Impairment	139.2	55.1					
Impaired Glucose	130.0	56.9					
Diabetes	126.8	59.6					
	mean trails B	mean trails B (unaffected)	Wilcoxon p-value	age-adjusted p-value			
Hypertension	132.0	127.7	0.010	0.1482			
Antidepressant use	142.8	129.1	0.0116	0.0008			
SSRI use	136.9	129.5	0.4166	0.0835			

Table 7: Mean 3MS scores by covariates

	r (p-value)	r Age-adjusted (p-value)	Q1	Q2	Q3	Q4	K-W p-value
Age, years			under 70	70-73	74-78	above 78	
Mean 3MS	-0.2896 (<0.0001)		95.6	94.7	93.7	92.0	<0.0001
BMI, kg/m^2			under 24.7	24.7-26.7	26.8-29.4	above 29.4	
Mean 3MS	-0.1816 (0.2809)	-0.0753 (<0.0001)	94.1	94.4	94.1	93.9	0.006
Total Cholesterol,			under 169	169-191.9	192-215	above 215	
Mean 3MS	-0.1816 (0.2809)	0.0088 (0.6176)	93.8	93.9	94.4	94.3	0.063
LDL Cholesterol, mg/dL			under 94	94-113	114-134	above 134	
Mean 3MS	-0.1816 (0.2809)	0.0137 (0.4407)	93.8	94.0	94.2	94.3	0.072
HDL Cholesterol,			under 40	41-47	48-57	above 57	
Mean 3MS	0.0512 (0.0034)	0.0696 (<0.0001)	93.9	93.8	94.4	94.3	0.002
Systolic BP, mm/Hg			under 131	131-140	141-152	above 152	
Mean 3MS	-0.1219 (<0.0001)	-0.0826 (<0.0001)	95.0	94.1	93.7	93.4	<0.0001
	Mean 3MS	SD	K-W Test p-value	Age-adjusted p-value			
Education			<0.0001	<0.0001			
Elementary School	87.9	6.6					
Some High School	90.4	5.6					
High School or Some College	93.1	5.0					
College	94.4	5.7					
All/Some Grad	95.7	3.7					
Site							
Minneapolis	95.1	4.2	<0.0001	<0.001			
Palo Alto	94.2	6.3					
Pittsburgh	92.9	5.0					
Portland	94.4	4.6					
San Diego	93.9	4.5					
Diabetes			<0.0001	<0.0001			
No Impairment	93.4	4.5					
Impaired Glucose	93.9	5.8					
Diabetes	94.4	4.6					
	mean 3MS	mean 3MS (unaffected)	Wilcoxon p-value	age-adjusted p-value			
Hypertension	93.83	94.27	0.0006	0.0553			
Antidepressant use	93.73	94	0.6293	0.5435			
SSRI use	94.23	94.02	0.4006	0.0956			

Table 8: Linear regression of SNPs on Trails B

					Age-Adjusted[a]				Multivariate Model[b]			
gene	SNP	major allele	Minor allele	MAF	B	SE(B)	p-value	P-value with FDR	B	SE(B)	p-value	P-value with FDR
TRIM65	rs7208173	C	T	0.001	-45.212	19.978	**0.024**	0.205	-47.974	19.549	**0.014**	0.122
MRPL38	rs9892372	G	A	0.006	-19.833	7.767	**0.011**	0.139	-22.748	8.231	**0.006**	0.074
FBF1	rs7213548	A	T	0.069	1.931	2.473	0.435	0.960	2.613	2.648	0.324	0.959
ACOX1	rs3643	T	C	0.078	0.118	2.336	0.960	0.960	0.459	2.494	0.854	0.959
ACOX1	rs8082018	G	A	0.006	-20.578	7.686	**0.007**	0.139	-23.102	8.131	**0.004**	0.074
ACOX1	rs7213998	C	T	0.082	-0.961	2.261	0.671	0.960	-0.123	2.407	0.959	0.959
ACOX1	rs11651351	C	T	0.056	-0.415	2.726	0.879	0.960	-1.820	2.921	0.533	0.959
ACOX1	rs3744033*	A	G	0.191	9.671	4.569	**0.034**	0.297	8.787	4.835	0.069	0.363

SNPs in this table were significant before FDR corrections in at least one model on Trails B or 3MS

*Recesive effects model shown, not significant in the additive model

[a]Also adjusted for study site

[b]Adjusted for age, principal components, BMI, diabetes, education, systolic blood pressure and HDL cholesterol

Table 9: Linear regression of SNPs on 3MS

					Age-Adjusted[a]				Multivariate Model[b]			
gene	SNP	major allele	Minor allele	MAF	B	SE(B)	p-value	P-value with FDR	B	SE(B)	p-value	P-value with FDR
TRIM65	rs7208173	C	T	0.235	1.548	1.839	0.400	0.780	1.780	1.670	0.286	0.531
MRPL38	rs9892372	G	A	0.006	1.754	0.714	**0.014**	0.182	1.550	0.690	**0.026**	0.221
FBF1	rs7213548*	A	T	0.069	3.309	1.271	**0.009**	0.065	2.535	1.257	**0.044**	0.184
ACOX1	rs3643*	T	C	0.078	3.477	1.081	**0.001**	**0.014**	2.912	1.034	**0.005**	0.051
ACOX1	rs8082018	G	A	0.006	1.748	0.706	**0.013**	0.182	1.430	0.690	**0.038**	0.244
ACOX1	rs7213998	C	T	0.082	0.430	0.208	**0.039**	0.308	0.460	0.200	**0.021**	0.221
ACOX1	rs11651351*	C	T	0.056	6.552	1.451	**6.53E-06**	1.37E-04	9.732	1.448	**2.16E-11**	**4.54E-10**
ACOX1	rs3744033	A	G	0.191	-0.117	0.145	0.421	0.780	-0.100	0.140	0.494	0.676

SNPs in this table were significant before FDR corrections in at least one model on Trails B or 3MS

*Recesive effects model shown, not significant in the additive model

[a]Also adjusted for study site

[b]Adjusted for age, principle components, BMI, diabetes, education, systolic blood pressure and HDL cholesterol

Appendix 1
Correlations between SNPs within genes

Correlations Between SNPs in *TRIM47*			
	rs3744017	**rs3744021**	**rs9908862**
rs3744021	r=0.738		
	p<.0001		
rs9908862	r=0.840	r=0.638	
	p<.0001	p<.0001	
rs1105917	r=0.860	r=0.648	r=0.980
	p<.0001	p<.0001	p<.0001

Correlations Between SNPs in *TRIM65*			
	rs3744017	**rs3744021**	**rs9908862**
rs3744021	r=0.738		
	p<.0001		
rs9908862	r=0.840	r=0.638	
	p<.0001	p<.0001	
rs1105917	r=0.860	r=0.648	r=0.980
	p<.0001	p<.0001	r<.0001

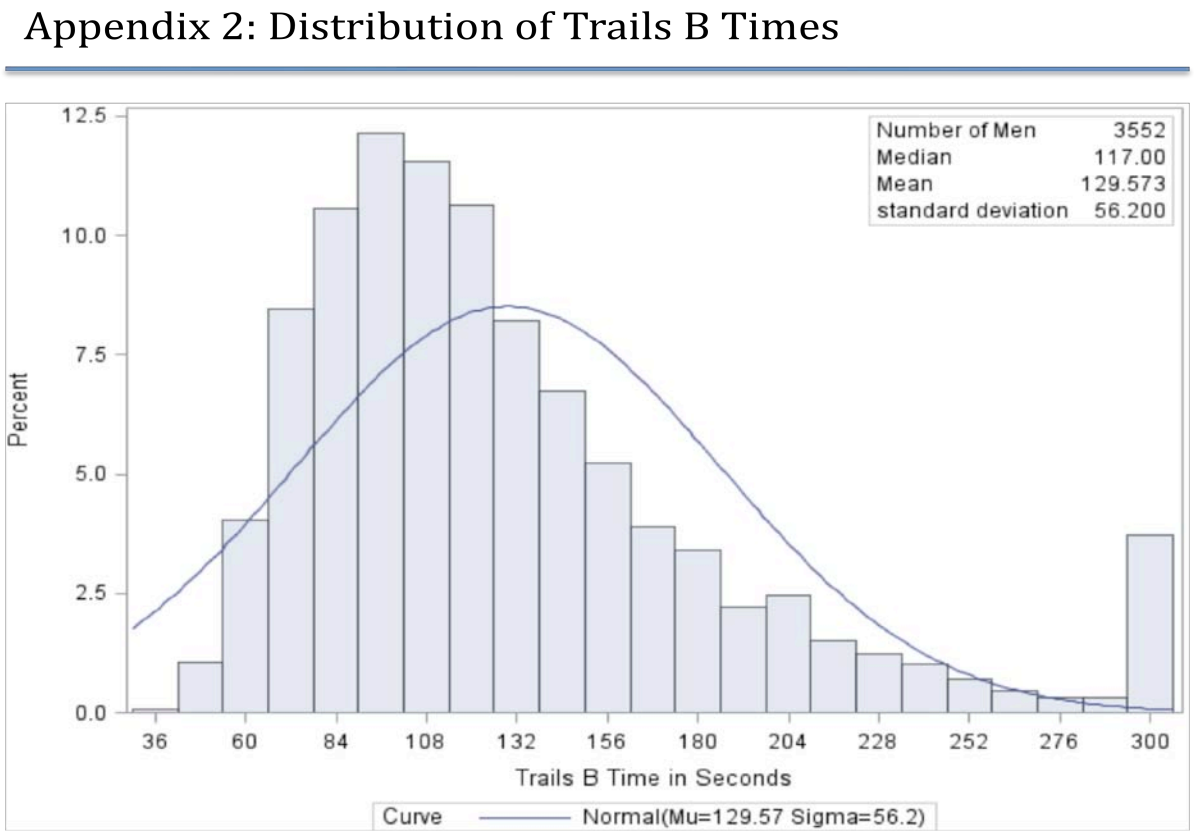
Correlations Between SNPs in *WBP2*		
	rs2305914	
rs936393	r=0.489	
	p<.0001	

Correlations Between SNPs in *MRPL38*		
	rs7370	**rs9892372**
rs9892372	p=-0.035	
	p=0.0395	
rs34136221	r=-0.066	r=-0.009
	p<.0001	p=0.576

Correlations Between SNPs in *FBF1*								
	rs9891076	**rs7213548**	**rs7218738**	**rs2305913**	**rs6501840**	**rs1135889**	**rs2608882**	**rs2608881**
rs7213548	r=0.396							
	p<.0001							
rs7218738	r=0.374	r=0.967						
	p<.0001	p<.0001						
rs2305913	r=0.984	r=0.394	r=0.402					
	p<.0001	p<.0001	p<.0001					
rs6501840	r=0.078	r=-0.019	r=-0.009	r=0.084				
	p<.0001	p=0.268	p=0.580	p<.0001				
rs1135889	r=0.724	r=-0.119	r=-0.125	r=0.725	r=-0.018			
	p<.0001	p<.0001	p<.0001	p<.0001	p=0.295			
rs2608882	r=0.722	r=-0.115	r=-0.121	r=0.723	r=-0.018	r=0.996		
	p<.0001	p<.0001	p<.0001	p<.0001	p=0.290	p<.0001		
rs2608881	r=0.834	r=-0.131	r=-0.136	r=0.821	r=-0.029	r=0.867	r=0.866	
	p<.0001	p<.0001	p<.0001	p<.0001	p=0.088	p<.0001	p<.0001	
rs3744002	r=0.394	r=0.994	r=0.967	r=0.394	r=-0.019	r=-0.120	r=-0.116	r=-0.130
	p<.0001	p<.0001	p<.0001	p<.0001	p=0.271	p<.0001	p<.0001	<.0001

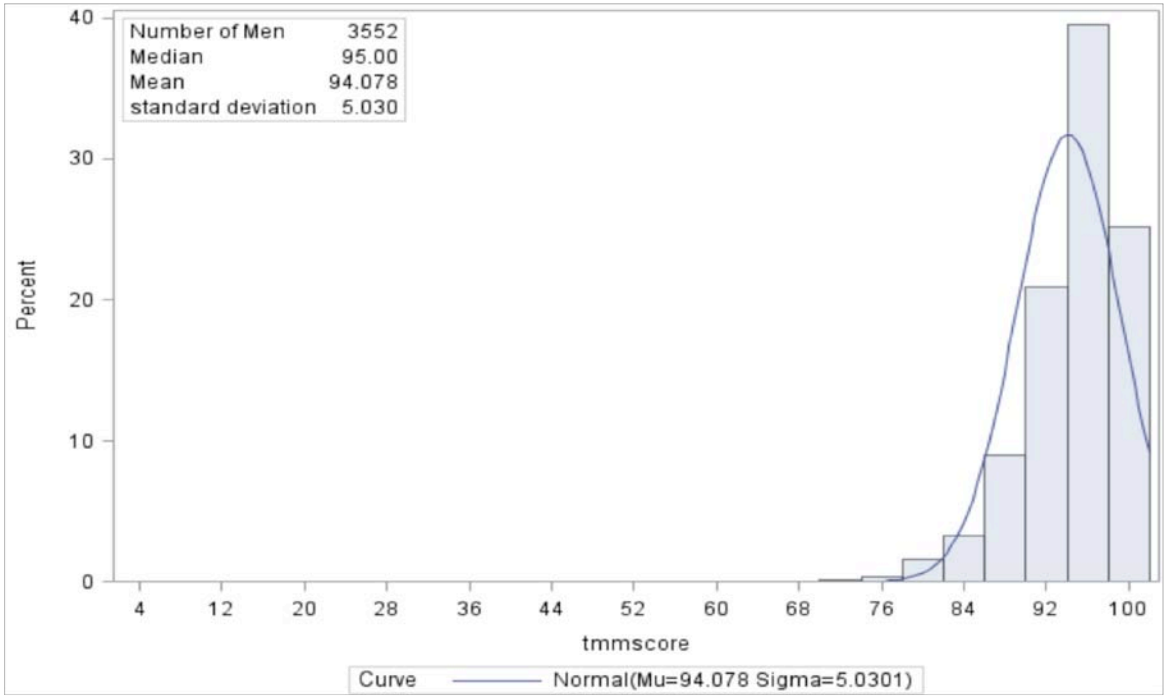
Correlations Between SNPs in *ACOX1*															
	rs3643	**rs12430**	**rs3682**	**rs8082018**	**rs8065946**	**rs7217955**	**rs10852766**	**rs11077799**	**rs12603572**	**rs7213998**	**rs7219716**	**rs7218656**	**rs11651351**	**rs3744032**	**rs3744033**
rs12430	r=-0.067														
	p<.0001														
rs3682	r=0.425	r=0.469													
	p<.0001	p<.0001													
rs8082018	r=0.278	r=-0.033	r=0.116												
	p<.0001	p=0.047	p<.0001												
rs8065946	r=0.282	r=-0.033	r=0.120	r=0.989											
	p<.0001	p=0.052	p<.0001	p<.0001											
rs7217955	r=0.980	r=-0.065	r=0.417	r=0.278	r=0.282										
	p<.0001	p=0.0001	p<.0001	p<.0001	p<.0001										
rs10852766	r=0.424	r=0.468	r=0.996	r=0.116	r=0.120	r=0.416									
	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001									
rs11077799	r=-0.105	r=-0.087	r=-0.251	r=-0.018	r=-0.017	r=-0.098	r=-0.251								
	p<.0001	p<.0001	p<.0001	p=0.294	p=0.322	p<.0001	p<.0001								
rs12603572	r=-0.015	r=-0.052	r=0.143	r=-0.017	r=-0.017	r=-0.018	r=0.146	r=-0.049							
	p=0.368	p=0.002	p<.0001	p=0.320	p=0.326	p=0.278	p<.0001	p=0.003							
rs7213998	r=0.575	r=-0.092	r=0.280	r=-0.008	r=-0.007	r=0.585	r=0.282	r=0.025	r=-0.018						
	p<.0001	p<.0001	p<.0001	p=0.649	p=0.690	p<.0001	p<.0001	p=0.144	p=0.283						
rs7219716	r=0.463	r=-0.129	r=0.381	r=0.105	r=0.107	r=0.470	r=0.384	r=-0.032	r=0.281	r=0.749					
	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p=0.055	p<.0001	p<.0001					
rs7218656	r=0.567	r=-0.096	r=0.288	r=-0.009	r=-0.008	r=0.575	r=0.290	r=0.024	r=-0.020	r=0.981	r=0.767				
	p<.0001	p<.0001	p<.0001	p=0.600	p=0.640	p<.0001	p<.0001	p=0.158	p=0.242	p<.0001	p<.0001				
rs11651351	r=-0.072	r=-0.034	r=-0.162	r=-0.007	r=-0.006	r=-0.071	r=-0.162	r=-0.054	r=-0.042	r=-0.075	r=-0.102	r=-0.078			
	p<.0001	p=0.046	p<.0001	p=0.690	p=0.724	p<.0001	p<.0001	p=0.001	p=0.012	p<.0001	p<.0001	p<.0001			
rs3744032	r=-0.043	r=-0.070	r=0.214	r=-0.003	r=-0.002	r=-0.042	r=0.216	r=-0.079	r=0.477	r=-0.065	r=0.568	r=-0.061	r=-0.063		
	p=0.010	p<.0001	p<.0001	p=0.866	p=0.902	p=0.013	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p=0.0003	p=0.0002		
rs3744033	r=-0.066	r=0.616	r=0.189	r=-0.016	r=-0.015	r=-0.059	r=0.189	r=-0.104	r=-0.073	r=-0.147	r=-0.212	r=-0.155	r=0.502	r=-0.130	
	p<.0001	p<.0001	p<.0001	p=0.328	p=0.375	p=0.0004	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	p<.0001	
rs3744035	r=-0.019	r=-0.046	r=0.149	r=-0.017	r=-0.017	r=-0.021	r=0.152	r=-0.052	r=0.971	r=-0.021	r=0.283	r=-0.023	r=-0.038	r=0.491	r=-0.071
	p=0.271	p=0.006	p<.0001	p=0.305	p=0.311	p=0.202	p<.0001	p=0.002	p<.0001	p=0.202	p<.0001	p=0.171	p=0.022	p<.0001	p<.0001

Appendix 2
Distributions of Cognitive Measures



Appendix 3: Distribution of Trails B times in seconds with normal curve superimposed. Notice the spike at 300 seconds indicating those subjects who did not complete the test in this time period.

Appendix 2: Distribution of 3MS scores



Appendix 3: Distribution of Trails B times in seconds with normal curve superimposed. Notice the spike at 300 seconds indicating those subjects who did not complete the test in this time period.

Appendix 3

Allele Frequencies and Average Trails B Scores by Genotype											
Gene	SNP	Major Allele	Minor Allele	Minor Allele Frequency	Minor Allele homozygote		Heterozygote		Major Allele homozygote		HWE Chi-Squared p-value
					N	Mean Trails B (SD)	N	Mean Trails B (SD)	N	Mean Trails B (SD)	
WBP2	rs2305914	G	T	0.0727	15	144.7(57.1)	483	125.9(53.6)	3028	130(56.7)	0.7769
WBP2	rs936393	A	G	0.1908	127	131.3(60)	1091	128.7(55.1)	2307	129.8(56.6)	0.9958
TRIM47	rs3744017	G	A	0.1915	126	131.3(60.3)	1098	129.1(55.7)	2300	129.5(56.3)	0.9689
TRIM65	rs1551619	C	T	0.2351	195	128.7(58.6)	1267	128.8(55.2)	2062	129.9(56.6)	0.9999
TRIM65	rs3760128	A	G	0.3482	437	131.1(59.6)	1565	127.8(53.7)	1500	130.6(57.7)	0.8151
TRIM65	rs4393627	A	G	0.1061	45	121.2(44.9)	658	130.9(57)	2822	129.3(56.3)	0.8198
TRIM65	rs7208173	C	T	0.0010	0	na	7	154.4(74)	3516	129.5(56.2)	0.9999
TRIM65	rs9914840	C	T	0.0024	0	na	17	137.9(62)	3507	129.4(56.2)	0.9999
TRIM65	rs9916659	A	G	0.0026	0	na	18	138.8(60.3)	3450	129.6(56.3)	0.9999
MRPL38	rs7370	G	A	0.2645	238	134.3(63.4)	1389	128.3(54.6)	1898	129.7(56.6)	0.8636
MRPL38	rs9892372	G	A	0.0061	2	167(18.4)	39	145.5(67.7)	3478	129.3(56.1)	0.4803
MRPL38	rs34136221	G	A	0.0206	1	86(na)	143	128.4(56.8)	3381	129.5(56.3)	0.9999
FBF1	rs7213548	A	T	0.0691	14	151.7(58.1)	459	125.4(52.9)	3051	130(56.7)	0.8514
FBF1	rs6501840	C	T	0.0028	0	na	20	121.8(59.2)	3505	129.5(56.3)	0.9999
FBF1	rs2608881	G	C	0.2692	248	133.6(62.4)	1401	128.4(54.3)	1875	129.8(56.9)	0.9074
ACOX1	rs3643	T	C	0.0781	20	153.2(63)	510	126.3(53.8)	2993	129.8(56.5)	0.9837
ACOX1	rs12430	G	A	0.1011	42	119.5(45.8)	627	130.4(56.4)	2846	129.4(56.4)	0.7451
ACOX1	rs3682	A	G	0.3474	448	132.2(60)	1552	127.7(53.7)	1523	130.4(57.6)	0.4916
ACOX1	rs8082018	G	A	0.0062	2	167(18.4)	40	145.7(66.8)	3483	129.3(56.1)	0.4846
ACOX1	rs11077799	A	G	0.1039	37	136.7(75.1)	658	131.1(56.7)	2827	129(55.9)	0.9955
ACOX1	rs12603572	T	C	0.0119	0	na	84	130.1(51.2)	3440	139.5(56.4)	0.9999
ACOX1	rs7213998	C	T	0.0825	26	119.6(51.2)	529	129.9(57.6)	2968	129.9(57.6)	0.9562
ACOX1	rs7219716	C	A	0.1372	73	125.5(50.7)	820	129.7(56.5)	2627	129.6(56.4)	0.7893
ACOX1	rs11651351	C	T	0.0558	11	124.9(39.2)	371	131.7(59.1)	3141	129.3(56)	0.9999
ACOX1	rs3744032	C	T	0.0498	8	124.4(40.3)	335	128.8(53.8)	3181	129.6(56.6)	0.9999
ACOX1	rs3744033	A	G	0.1906	137	121(46.3)	1070	131.2(58.1)	2318	129.2(55.9)	0.7942

Appendix 4

Allele Frequencies and Average 3MS Scores by Genotype											
Gene	SNP	Major Allele	Minor Allele	Minor Allele Frequency	Minor Allele homozygote		Heterozygote		Major Allele homozygote		HWE Chi-Squared p-value
					N	Mean 3MS (SD)	N	Mean 3MS (SD)	N	Mean 3MS (SD)	
WBP2	rs2305914	G	T	0.0727	15	91.8(7)	483	93.9(4.6)	3028	94.1(5.1)	0.7769
WBP2	rs936393	A	G	0.1908	127	93.7(4.9)	1091	94.1(4.5)	2307	94.1(5.2)	0.9958
TRIM47	rs3744017	G	A	0.1915	126	93.5(4.9)	1098	94.1(4.6)	2300	94.1(5.2)	0.9689
TRIM65	rs1551619	C	T	0.2351	195	93.9(4.7)	1267	94(4.6)	2062	94.2(5.3)	0.9999
TRIM65	rs3760128	A	G	0.3482	437	93.8(4.9)	1565	94.1(4.5)	1500	94.1(5.6)	0.8151
TRIM65	rs4393627	A	G	0.1061	45	93.2(5.6)	658	94.1(4.5)	2822	94.1(5.1)	0.8198
TRIM65	rs7208173	C	T	0.0010	0	na	7	93.7(5)	3516	94.1(5)	0.9999
TRIM65	rs9914840	C	T	0.0024	0	na	17	93.8(2.9)	3507	94.1(5)	0.9999
TRIM65	rs9916659	A	G	0.0026	0	na	18	93.6(2.9)	3450	94.1(5)	0.9999
MRPL38	rs7370	G	A	0.2645	238	94(4.7)	1389	94.1(4.5)	1898	94.1(5.4)	0.8636
MRPL38	rs9892372	G	A	0.0061	2	92(2.8)	39	92.2(5.5)	3478	94.1(5)	0.4803
MRPL38	rs34136221	G	A	0.0206	1	na	143	94.2(4.6)	3381	94.1(5)	0.9999
FBF1	rs7213548	A	T	0.0691	14	90.6(7.3)	459	94.1(4.5)	3051	94.1(5.1)	0.8514
FBF1	rs6501840	C	T	0.0028	0	na	20	92.7(5.5)	3505	94.1(5)	0.9999
FBF1	rs2608881	G	C	0.2692	248	94(4.8)	1401	94.1(4.5)	1875	94.1(5.4)	0.9074
ACOX1	rs3643	T	C	0.0781	20	90.8(6.5)	510	93.9(4.6)	2993	94.1(5.1)	0.9837
ACOX1	rs12430	G	A	0.1011	42	93.5(5.2)	627	94.2(4.5)	2846	94.2(5.1)	0.7451
ACOX1	rs3682	A	G	0.3474	448	93.7(4.8)	1552	94.1(4.5)	1523	94.2(5.6)	0.4916
ACOX1	rs8082018	G	A	0.0062	2	92(2.8)	40	92.3(5.3)	3483	94.1(5)	0.4846
ACOX1	rs11077799	A	G	0.1039	37	92.8(7.8)	658	94.2(4.8)	2827	94.2(5)	0.9955
ACOX1	rs12603572	T	C	0.0119	0	na	84	94.3(4.1)	3440	94.1(5)	0.9999
ACOX1	rs7213998	C	T	0.0825	26	94.2(5.6)	529	93.7(5)	2968	94.2(5)	0.9562
ACOX1	rs7219716	C	A	0.1372	73	94.2(4.4)	820	93.9(4.9)	2627	94.2(5.1)	0.7893
ACOX1	rs11651351	C	T	0.0558	11	97.5(2.79)	371	94.4(4.5)	3141	94.1(4.8)	0.9999
ACOX1	rs3744032	C	T	0.0498	8	95.6(3.3)	335	94.2(4.5)	3181	94.1(5.1)	0.9999
ACOX1	rs3744033	A	G	0.1906	137	93.8(8.8)	1070	94.3(4.6)	2318	94(4.9)	0.7942

Appendix 5

Linear Regression of Trails B on SNPs - Additive Effects																	
		Adjusted for Age and Study Site				Adjusted for Age, Study Site and PCA				Adjusted for Age, Study Site, PCA (Censored)				Fully Adjusted Model			
Gene	SNP	B	se(B)	p-value	p-value with FDR	B	se(B)	p-value	p-value with FDR	B	SE	p-value	p-value with FDR	B	se(B)	p-value	p-value with FDR
WBP2	rs2305914	2.184	2.418	0.366	0.982	2.056	2.425	0.396	0.960	2.102	2.509	0.402	0.996	2.035	2.579	0.430	0.959
WBP2	rs936393	0.618	1.584	0.696	0.982	0.539	1.586	0.734	0.960	0.537	1.641	0.743	0.996	0.105	1.685	0.950	0.959
TRIM47	rs3744017	0.288	1.584	0.856	0.982	0.199	1.586	0.900	0.960	0.157	1.642	0.924	0.996	-0.201	1.684	0.905	0.959
TRIM65	rs1551619	1.122	1.467	0.444	0.982	1.068	1.468	0.467	0.960	1.014	1.519	0.504	0.996	0.783	1.562	0.616	0.959
TRIM65	rs3760128	0.889	1.300	0.494	0.982	0.765	1.303	0.557	0.960	0.769	1.349	0.568	0.996	0.915	1.390	0.510	0.959
TRIM65	rs4393627	-0.046	2.005	0.982	0.982	-0.200	2.010	0.921	0.960	-0.086	2.079	0.967	0.996	0.836	2.142	0.696	0.959
TRIM65	rs7208173	-41.179	19.818	**0.038**	0.327	-45.212	19.978	**0.024**	0.205	-45.010	20.649	**0.029**	0.254	-47.974	19.549	**0.014**	0.122
TRIM65	rs9914840	-8.886	12.768	0.486	0.982	-11.024	12.890	0.392	0.960	-10.323	13.323	0.438	0.996	-14.117	13.416	0.293	0.959
TRIM65	rs9916659	-10.805	12.442	0.385	0.982	-12.656	12.564	0.314	0.960	-11.960	12.986	0.357	0.996	-16.176	13.011	0.214	0.959
MRPL38	rs7370	-0.102	1.421	0.943	0.982	-0.148	1.422	0.917	0.960	-0.193	1.472	0.896	0.996	0.114	1.507	0.940	0.959
MRPL38	rs9892372	-18.127	7.698	**0.019**	0.241	-19.833	7.767	**0.011**	0.139	-19.622	8.034	**0.015**	0.190	-22.748	8.231	**0.006**	0.074
MRPL38	rs34136221	0.389	4.396	0.929	0.982	0.255	4.397	0.954	0.960	-0.025	4.552	0.996	0.996	-2.321	4.696	0.621	0.959
FBF1	rs7213548	2.019	2.473	0.414	0.982	1.931	2.473	0.435	0.960	1.993	2.559	0.436	0.996	2.613	2.648	0.324	0.959
FBF1	rs6501840	8.732	11.745	0.457	0.982	6.672	11.873	0.574	0.960	6.530	12.289	0.595	0.996	8.109	12.195	0.506	0.959
FBF1	rs2608881	0.234	1.413	0.869	0.982	0.183	1.414	0.897	0.960	0.171	1.463	0.907	0.996	0.397	1.496	0.791	0.959
ACOX1	rs3643	0.396	2.326	0.865	0.982	0.118	2.336	0.960	0.960	0.179	2.417	0.941	0.996	0.459	2.494	0.854	0.959
ACOX1	rs12430	0.401	2.049	0.845	0.982	0.314	2.053	0.879	0.960	0.416	2.124	0.845	0.996	1.426	2.181	0.513	0.959
ACOX1	rs3682	0.320	1.290	0.804	0.982	0.191	1.293	0.883	0.960	0.201	1.339	0.881	0.996	0.492	1.384	0.722	0.959
ACOX1	rs8082018	-18.873	7.617	**0.013**	0.241	-20.578	7.686	**0.007**	0.139	-20.374	7.950	**0.010**	0.190	-23.102	8.131	**0.004**	0.074
ACOX1	rs11077799	-3.921	2.047	0.055	0.360	-3.873	2.046	0.058	0.380	-4.115	2.119	0.052	0.339	-4.173	2.195	0.057	0.372
ACOX1	rs12603572	3.899	5.786	0.500	0.982	3.682	5.788	0.525	0.960	3.958	5.988	0.509	0.996	4.450	6.240	0.476	0.959
ACOX1	rs7213998	-0.860	2.258	0.703	0.982	-0.961	2.261	0.671	0.960	-1.200	2.341	0.608	0.996	-0.123	2.407	0.959	0.959
ACOX1	rs7219716	-0.080	1.799	0.964	0.982	-0.190	1.804	0.916	0.960	-0.304	1.867	0.870	0.996	0.461	1.922	0.810	0.959
ACOX1	rs11651351	-0.475	2.720	0.861	0.982	-0.415	2.726	0.879	0.960	-0.522	2.822	0.853	0.996	-1.820	2.921	0.533	0.959
ACOX1	rs3744032	1.880	2.875	0.513	0.982	1.863	2.877	0.517	0.960	1.950	2.977	0.513	0.996	2.760	3.077	0.370	0.959
ACOX1	rs3744033	0.569	1.572	0.717	0.982	0.517	1.576	0.743	0.960	0.544	1.630	0.739	0.996	0.182	1.678	0.914	0.959

Appendix 6

Linear Regression of Trails B on SNPs - Recessive Effects

		Adjusted for Age and Study Site				Adjusted for Age, Study Site and PCA				Adjusted for Age, Study Site, PCA (Censored)				Fully Adjusted Model			
Gene	SNP	B	se(B)	p-value	p-value with FDR	B	se(B)	p-value	p-value with FDR	B	SE	p-value	p-value with FDR	B	se(B)	p-value	p-value with FDR
WBP2	rs2305914	-10.401	13.550	0.443	0.757	-10.292	13.549	0.447	0.750	-11.937	13.960	0.393	0.749	-14.753	14.226	0.300	0.770
WBP2	rs936393	-1.104	4.715	0.815	0.861	-1.201	4.715	0.799	0.869	-1.155	4.829	0.811	0.946	-0.181	4.996	0.971	0.971
TRIM47	rs3744017	-0.901	4.730	0.849	0.861	-1.033	4.731	0.827	0.869	-0.511	4.836	0.916	0.948	0.833	4.994	0.868	0.959
TRIM65	rs1551619	0.676	3.853	0.861	0.861	0.632	3.854	0.870	0.870	0.260	3.955	0.948	0.948	1.689	4.089	0.680	0.793
TRIM65	rs3760128	-1.806	2.672	0.499	0.757	-2.102	2.680	0.433	0.750	-2.110	2.746	0.442	0.750	-2.088	2.844	0.463	0.770
TRIM65	rs4393627	6.361	7.857	0.418	0.757	5.456	7.884	0.489	0.750	8.477	7.983	0.288	0.605	4.103	8.396	0.625	0.772
MRPL38	rs7370	-3.883	3.515	0.269	0.707	-4.064	3.517	0.248	0.650	-3.928	3.600	0.275	0.605	-3.492	3.672	0.342	0.770
MRPL38	rs9892372	-58.927	37.057	0.112	0.587	-72.370	37.924	0.056	0.297	-66.156	38.410	0.085	0.448	-77.106	37.177	**0.038**	0.363
MRPL38	rs34136221	18.861	52.409	0.719	0.839	18.675	52.407	0.722	0.842	23.256	53.054	0.661	0.868	3.805	51.253	0.941	0.971
FBF1	rs7213548	-19.241	14.021	0.170	0.595	-18.978	14.019	0.176	0.615	-16.444	14.193	0.247	0.605	-11.742	15.464	0.448	0.770
FBF1	rs2608881	-2.390	3.450	0.489	0.757	-2.584	3.452	0.454	0.750	-2.383	3.532	0.500	0.750	-2.091	3.597	0.561	0.772
ACOX1	rs3643	-25.680	11.719	**0.028**	0.313	-26.781	11.746	**0.023**	0.297	-24.041	11.893	**0.043**	0.448	-23.668	12.471	0.058	0.363
ACOX1	rs12430	6.764	8.137	0.406	0.757	5.820	8.164	0.476	0.750	8.926	8.267	0.280	0.605	5.564	8.514	0.513	0.770
ACOX1	rs3682	-2.994	2.645	0.258	0.707	-3.321	2.654	0.211	0.632	-3.543	2.721	0.193	0.605	-2.608	2.828	0.356	0.770
ACOX1	rs8082018	-58.953	37.048	0.112	0.587	-72.280	37.914	0.057	0.297	-66.055	38.399	0.085	0.448	-76.981	37.173	**0.038**	0.363
ACOX1	rs11077799	-12.157	8.664	0.161	0.595	-12.058	8.665	0.164	0.615	-11.798	8.885	0.184	0.605	-13.591	9.409	0.149	0.624
ACOX1	rs7213998	6.884	10.321	0.505	0.757	6.973	10.346	0.500	0.750	7.391	10.566	0.484	0.750	7.450	10.763	0.489	0.770
ACOX1	rs7219716	2.741	6.204	0.659	0.839	2.554	6.221	0.681	0.842	3.627	6.339	0.567	0.794	6.406	6.607	0.332	0.770
ACOX1	rs11651351	7.244	15.818	0.647	0.839	7.503	15.819	0.635	0.842	4.305	16.478	0.794	0.946	-8.512	17.098	0.619	0.772
ACOX1	rs3744032	6.857	18.536	0.711	0.839	6.778	18.540	0.715	0.842	3.228	19.317	0.867	0.948	15.213	20.948	0.468	0.770
ACOX1	rs3744033	9.909	4.561	**0.030**	0.313	9.671	4.569	**0.034**	0.297	9.319	4.689	**0.047**	0.448	8.787	4.835	0.069	0.363

Appendix 7

Linear Regression of 3MS on SNPs - Additive Effects

		Adjusted for Age and Study Site				Adjusted for Age, Study Site and PCA				Adjusted for Age, Study Site, PCA (Censored)				Fully Adjusted Model			
Gene	SNP	B	se(B)	p-value	p-value with FDR correction	B	se(B)	p-value	p-value with FDR correction	B	se(B)	p-value	p-value with FDR correction	B	se(B)	p-value	p-value with FDR correction
WBP2	rs2305914	0.382	0.219	0.081	0.423	0.355	0.220	0.106	0.552	0.345	0.230	0.134	0.617	0.427	0.214	**0.046**	0.352
WBP2	rs936393	0.040	0.144	0.781	0.953	0.032	0.144	0.825	0.933	0.030	0.150	0.843	0.984	0.115	0.139	0.408	0.624
TRIM47	rs3744017	0.134	0.144	0.350	0.718	0.126	0.144	0.380	0.759	0.132	0.150	0.379	0.751	0.171	0.139	0.218	0.499
TRIM65	rs1551619	0.104	0.133	0.434	0.745	0.095	0.133	0.478	0.823	0.102	0.139	0.464	0.801	0.181	0.129	0.160	0.447
TRIM65	rs3760128	0.072	0.118	0.542	0.829	0.069	0.119	0.559	0.856	0.073	0.124	0.558	0.853	0.109	0.115	0.341	0.601
TRIM65	rs4393627	0.047	0.182	0.796	0.953	0.058	0.182	0.750	0.933	0.051	0.190	0.789	0.984	0.000	0.177	0.999	0.999
TRIM65	rs7208173	1.334	1.800	0.459	0.745	1.484	1.814	0.413	0.767	1.313	1.914	0.493	0.801	1.562	1.660	0.347	0.601
TRIM65	rs9914840	-0.076	1.159	0.947	0.953	-0.266	1.170	0.820	0.933	-0.176	1.216	0.885	0.984	0.324	1.139	0.776	0.877
TRIM65	rs9916659	0.201	1.128	0.859	0.953	0.036	1.138	0.975	0.985	0.129	1.184	0.913	0.984	0.689	1.104	0.533	0.712
MRPL38	rs7370	-0.008	0.129	0.953	0.953	-0.002	0.129	0.985	0.985	-0.004	0.135	0.974	0.984	0.036	0.125	0.771	0.877
MRPL38	rs9892372	1.549	0.699	**0.027**	0.205	1.524	0.705	**0.031**	0.236	1.571	0.735	**0.033**	0.251	1.330	0.690	0.054	0.352
MRPL38	rs34136221	-0.081	0.399	0.839	0.953	-0.092	0.399	0.818	0.933	-0.102	0.417	0.807	0.984	-0.059	0.385	0.878	0.951
FBF1	rs7213548	0.320	0.224	0.154	0.667	0.319	0.224	0.155	0.672	0.302	0.235	0.199	0.617	0.333	0.217	0.126	0.447
FBF1	rs6501840	1.242	1.066	0.244	0.703	1.195	1.077	0.267	0.695	1.335	1.120	0.233	0.617	1.165	1.012	0.250	0.499
FBF1	rs2608881	-0.046	0.128	0.717	0.953	-0.044	0.128	0.734	0.933	-0.047	0.134	0.728	0.984	-0.010	0.124	0.935	0.973
ACOX1	rs3643	0.469	0.211	**0.026**	0.205	0.463	0.212	**0.029**	0.236	0.458	0.222	**0.039**	0.251	0.457	0.205	**0.026**	0.340
ACOX1	rs12430	-0.026	0.186	0.888	0.953	-0.010	0.186	0.956	0.985	-0.024	0.194	0.900	0.984	-0.063	0.180	0.728	0.877
ACOX1	rs3682	0.108	0.117	0.359	0.718	0.108	0.117	0.360	0.759	0.102	0.123	0.405	0.751	0.132	0.114	0.245	0.499
ACOX1	rs8082018	1.487	0.692	**0.032**	0.205	1.460	0.698	**0.036**	0.236	1.511	0.727	**0.038**	0.251	1.179	0.682	0.084	0.364
ACOX1	rs11077799	0.193	0.186	0.300	0.709	0.189	0.186	0.309	0.730	0.223	0.194	0.251	0.617	0.255	0.180	0.156	0.447
ACOX1	rs12603572	-0.579	0.525	0.270	0.703	-0.590	0.525	0.262	0.695	-0.617	0.549	0.261	0.617	-0.426	0.508	0.402	0.624
ACOX1	rs7213998	0.499	0.205	**0.015**	0.205	0.484	0.205	**0.018**	0.236	0.522	0.214	**0.015**	0.251	0.467	0.198	**0.018**	0.340
ACOX1	rs7219716	0.203	0.163	0.214	0.695	0.190	0.163	0.245	0.695	0.209	0.171	0.220	0.617	0.276	0.158	0.082	0.364
ACOX1	rs11651351	0.032	0.247	0.896	0.953	0.066	0.247	0.790	0.933	-0.005	0.259	0.984	0.984	0.328	0.240	0.172	0.447
ACOX1	rs3744032	-0.329	0.261	0.207	0.695	-0.332	0.261	0.204	0.695	-0.366	0.273	0.180	0.617	-0.153	0.255	0.548	0.712
ACOX1	rs3744033	-0.116	0.143	0.417	0.745	-0.095	0.143	0.507	0.823	-0.143	0.149	0.338	0.733	-0.084	0.138	0.544	0.712

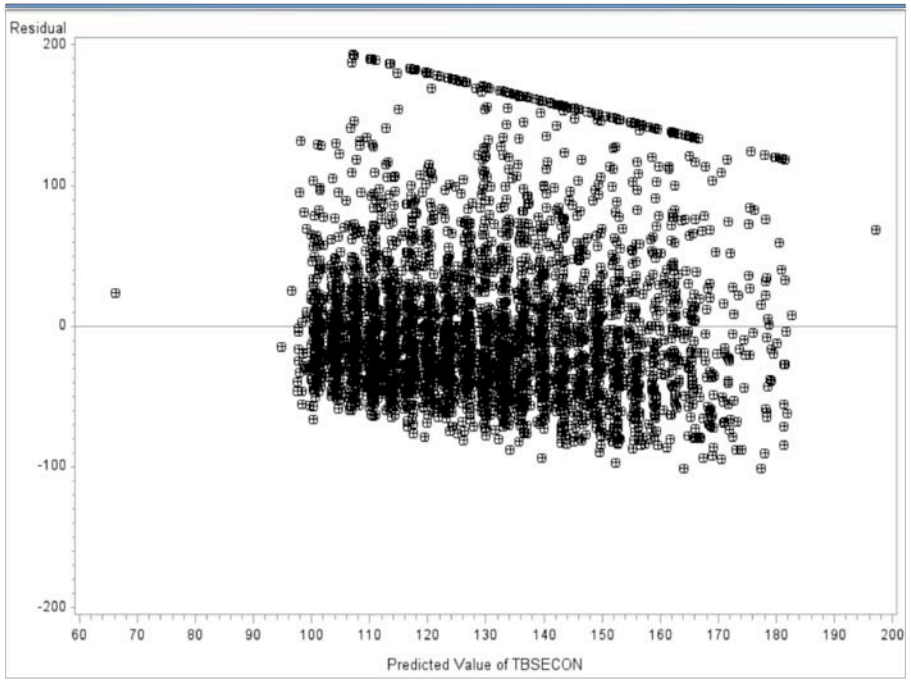
Appendix 8

Linear Regression of 3MS on SNPs - Recessive Effects

		Adjusted for Age and Study Site				Adjusted for Age, Study Site and PCA				Adjusted for Age, Study Site, PCA (Censored)				Fully Adjusted Model			
Gene	SNP	B	se(B)	p-value	p-value with FDR correction	B	se(B)	p-value	p-value with FDR correction	B	SE	p-value	p-value with FDR correction	B	se(B)	p-value	p-value with FDR correction
WBP2	rs2305914	1.813	1.230	0.140	0.421	1.802	1.229	0.143	0.428	1.773	1.283	0.167	0.388	2.599	1.207	**0.031**	0.164
WBP2	rs936393	0.320	0.428	0.455	0.683	0.305	0.428	0.475	0.696	0.315	0.447	0.481	0.704	0.387	0.415	0.350	0.525
TRIM47	rs3744017	0.535	0.429	0.213	0.468	0.516	0.429	0.229	0.436	0.548	0.448	0.221	0.388	0.518	0.414	0.212	0.370
TRIM65	rs1551619	0.195	0.350	0.578	0.758	0.173	0.350	0.621	0.811	0.168	0.366	0.647	0.849	0.210	0.338	0.535	0.668
TRIM65	rs3760128	0.303	0.243	0.213	0.468	0.302	0.244	0.215	0.436	0.319	0.255	0.210	0.388	0.269	0.235	0.252	0.408
TRIM65	rs4393627	1.086	0.713	0.128	0.421	1.127	0.715	0.115	0.428	1.278	0.744	0.086	0.360	0.936	0.694	0.178	0.370
MRPL38	rs7370	-0.015	0.319	0.963	0.963	-0.006	0.319	0.985	0.985	0.003	0.333	0.992	0.992	-0.002	0.307	0.994	0.994
MRPL38	rs9892372	3.910	3.365	0.245	0.468	4.339	3.443	0.208	0.436	4.664	3.581	0.193	0.388	4.510	3.153	0.153	0.370
MRPL38	rs34136221	1.239	4.757	0.795	0.878	1.276	4.755	0.788	0.843	1.527	4.944	0.757	0.872	3.097	4.352	0.477	0.667
FBF1	rs7213548	3.328	1.272	**0.009**	0.062	3.309	1.271	**0.009**	0.065	3.432	1.322	**0.009**	0.066	2.535	1.257	**0.044**	0.184
FBF1	rs2608881	-0.086	0.313	0.784	0.878	-0.080	0.313	0.799	0.843	-0.070	0.327	0.830	0.872	-0.053	0.301	0.860	0.950
ACOX1	rs3643	3.370	1.064	**0.002**	**0.016**	3.384	1.066	**0.002**	**0.016**	3.521	1.109	**0.001**	**0.016**	2.843	1.029	**0.006**	0.060
ACOX1	rs12430	0.811	0.737	0.271	0.475	0.852	0.739	0.249	0.436	1.009	0.769	0.189	0.388	0.964	0.711	0.175	0.370
ACOX1	rs3682	0.362	0.240	0.133	0.421	0.368	0.241	0.127	0.428	0.376	0.252	0.136	0.388	0.307	0.233	0.188	0.370
ACOX1	rs8082018	3.912	3.364	0.245	0.468	4.341	3.441	0.207	0.436	4.665	3.579	0.192	0.388	4.509	3.152	0.153	0.370
ACOX1	rs11077799	1.826	0.786	**0.020**	0.106	1.782	0.786	**0.023**	0.122	1.822	0.821	**0.026**	0.139	1.789	0.761	**0.019**	0.131
ACOX1	rs7213998	-0.089	0.937	0.924	0.963	-0.234	0.939	0.803	0.843	-0.234	0.981	0.812	0.872	-0.226	0.895	0.801	0.934
ACOX1	rs7219716	-0.167	0.563	0.766	0.878	-0.251	0.564	0.656	0.811	-0.202	0.588	0.732	0.872	0.007	0.547	0.989	0.994
ACOX1	rs11651351	6.638	1.431	**3.53E-06**	**7.41E-05**	6.702	1.431	**2.80E-06**	**6.00E-05**	6.700	1.493	**7.24E-06**	**1.52E-04**	9.940	1.441	**5.22E-12**	**1.10E-10**
ACOX1	rs3744032	-1.636	1.682	0.331	0.535	-1.711	1.682	0.309	0.499	-1.899	1.771	0.284	0.458	-1.087	1.779	0.541	0.668
ACOX1	rs3744033	0.246	0.414	0.553	0.758	0.282	0.415	0.497	0.696	0.290	0.433	0.503	0.704	0.513	0.403	0.204	0.370

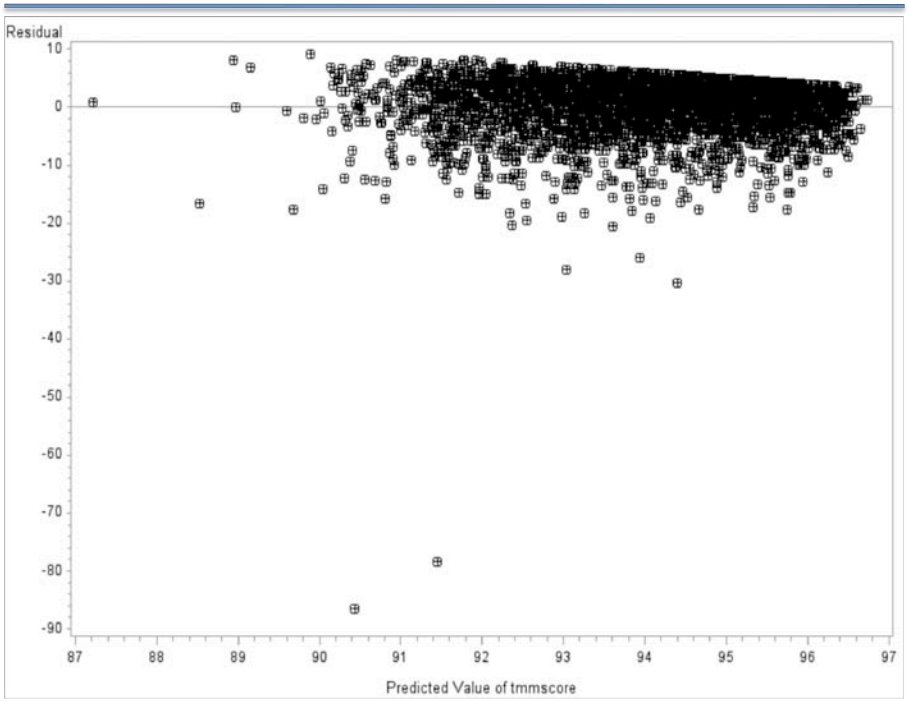
Appendix 9
Sample residual plots

Appendix 9: Example Plot, Residual Vs. Predicted Values for Regression of rs8082018 on Trails B



Appendix 10: After regression of each SNP on Trails B, predicted vs. residual values were plotted. This graph is an example representative of these plots. The pattern observed near the top of this graph represents individuals who did did not finish the Trails B test and therefore got a time of 300 seconds. It represents a deviation from normality, but this was expected and results did not change substantially after using censored normal regression to account for this phenomenon.

Appendix 9: Example Pot, Residual vs. Predicted for regression of rs3643 on 3MS



Appendix 10: After regression of each SNP on 3MS, predicted vs. residual values were plotted. This graph is an example representative of these plots. The outlier observed were representative of subjects who scored very poorly on this test (<60pts).