Heritability of Cognition and Memory Loss in the Cache County Memory Study Cohort

Cassidy P. Allen
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/gradreports

Part of the Physical Sciences and Mathematics Commons

Recommended Citation
Allen, Cassidy P., "Heritability of Cognition and Memory Loss in the Cache County Memory Study Cohort" (2008). All Graduate Plan B and other Reports. 1289. https://digitalcommons.usu.edu/gradreports/1289
HERITABILITY OF COGNITION AND MEMORY LOSS IN THE
CACHE COUNTY MEMORY STUDY COHORT

by

Cassidy P. Allen

A report submitted in partial fulfillment
of the requirements for the degree
of
MASTER OF SCIENCE
in
Statistics

UTAH STATE UNIVERSITY
Logan, Utah
2008
ABSTRACT

Heritability of Cognition and Memory Loss in the Cache County Memory Study Cohort

by

Cassidy P. Allen, Master of Science

Utah State University, 2008

Major Professor: Dr. Chris Corcoran
Department: Mathematics and Statistics

The Cache County Memory Study (CCMS) is a longitudinal study of the elderly in Cache County, Utah, that was initiated over twelve years ago to explore the role of APOE genotype and other environmental factors in dementia risk and cognitive function. Collaboration between the CCMS and investigators with the Utah Population Database (UPDB) at the University of Utah has revealed a significant number of sibships among the original 5,092 CCMS participants, along with complex pedigrees that include additional thousands of ancestors and relatives not enrolled in the Cache Study. Information about these families and pedigrees raises the potential for studies of the genetic epidemiology of age-related traits. Until biological samples from the Cache cohort can be leveraged for genetic association studies of cognition, a useful first step is to explore the heritability of cognition and cognitive decline (i.e., the degree to which these quantitative traits are inherited).
The objectives of this project are to:

1. Compare and contrast results from previous studies of the heritability of cognition and cognitive change among the elderly.

2. Use software tools to compute heritability measures with respect to cognitive function for the nuclear families within the CCMS. Additionally account for potential confounding factors – such as APOE genotype and age – and explain their effects.

3. Identify potential methods for more complex analyses with the CCMS data that could include information about complex pedigrees and repeated cognitive measures.

(42 pages)
ACKNOWLEDGMENTS

I would especially like to thank my committee members, Drs. Chris Corcoran, Richard Cutler, and Ron Munger, for their support and assistance throughout the entire process.

I give special thanks to Tyler and Lincoln, and to my family, friends, and colleagues for their encouragement, moral support, and patience as I worked my way from the project beginning to this final document. I could not have done it without each of you.

Cassidy P. Allen
1. INTRODUCTION AND BACKGROUND

When family members are more alike than pairs of unrelated individuals, the question arises as to how much of the similarities and differences – or variance – is due to genes, and what portion of the variability can be attributed to the environment. Correlation or covariation among family members quantify familial resemblance, while by definition heritability measures the degree of variation between individuals in a population. 

Conceptually, heritability is used “to quantify the level of predictability of passage of a biologically interesting phenotype [i.e., physical trait] from parent to offspring” (Feldman, 151). Heritability is often evaluated by careful experimentation and complex statistical analysis.

Formally, heritability is defined as a ratio of variances: the proportion of total variation in a population relative to a continuous-type trait – measured at a particular age or point in time – which can be attributed to variation in the population’s genetic makeup. As a summary statistic, heritability estimates the proportion of variation between individuals in a population that is attributable to their genotypes. However, heritability is conventionally computed having observed only familial relationships and phenotypes among study subjects.

An important distinction is generally made regarding the phenotypic variance considered in calculating heritability. There are several variants of heritability, with seemingly small but important differences. A broad distinction is often made between genetic and cultural heritability (or environmental variance). According to Rice and Borecki (2001), heritability originally centered on a polygenic model, or one in which a large
number of genes each have a small additive effect on phenotypic variability. In a polygenic model, the phenotype is simply a function of genetic and environmental effects, thus

\[ P = G + E. \]

Expressed in terms of variance components: \( V_P = V_G + V_E. \)

As genetic heritability is nearly always the focus of inference in modern studies of complex disease, it has been further divided into narrow-sense heritability (the most common meaning of heritability, denoted by \( h^2 \)) and broad-sense heritability \( (H^2) \). While broad-sense heritability is the fraction of the total variation in phenotype due to genetic effects \( (H^2 = V_G/V_P) \), the genetic variation can be decomposed further. Generally, the genetic variance can be partitioned into three parts: the variance of additive genetic effects, of dominant genetic effects, and of epistatic genetic effects (Visscher, Hill, and Wray 2008). Additive genetic effects are those that are transmissible from parent to child, as opposed to the part that is results from a unique patterning of genes in each generation (Neisser, et al. 1996). Narrow-sense heritability differs from the broad-sense in that it only estimates the proportion of phenotypic variation that is strictly due to additive genetic effects: \( h^2 = V_A/V_P. \) We will focus on this definition of narrow-sense heritability in our analyses of the Cache County data.

Heritability is not a measure of genetic control, but rather a statistic that measures how that control can vary. Heritability cannot be interpreted with respect to individual risk, and in that respect heritability measures are sometimes misunderstood. For example, given a heritability value of 0.6 for a personality trait, it would be incorrect to say that an individual inherits 60% of that trait from his parents and 40% from the environment. If all
people were genetic clones then all traits would have a heritability index of zero, meaning that all variability between individuals would be due only to environmental factors. Contrary to common misinterpretations of the heritability index, when populations become more egalitarian (i.e., their environmental exposures become more homogeneous) heritability indices rise. In that setting, variability between individuals can increasingly be explained only by genetic factors. Above all, estimates of heritability must be considered relative to the environment in which the trait is observed. Studies of the same trait carried out in different populations often yield different (sometimes inconsistent) results.

The fundamental questions in biology identified in the nature-nurture debate may have quantitative answers that can be expressed in terms of heritability. A highly genetically loaded trait, such as hair color, is still partly determined by environmental effects such as temperature and oxygen in the atmosphere. A more useful distinction than “nature v. nurture” is “obligate v. facultative” in which some traits are more obligate (e.g. everyone has eyes) and other more facultative, or sensitive to environmental variation, such as a person’s first language.

Studies of IQ illustrate some of these issues underlying heritability and its interpretation. An assortment of such studies have estimated the IQ heritability in the U.S. to be between 0.4 and 0.8 (Plomin, Pedersen, Lichtenstein, and McClearn 1994; Neisser, et al. 1996; Bouchard, Lykken, McGue, Segal, and Tellegen 1990); that is, depending on the study, variation in genes contributed somewhat less than half to considerably more than half of the variance in IQ among the children studied. The remainder of the variation would nominally be attributed to environmental effects and errors in IQ measurements. These
studies primarily involved twins under the age of 18, measured at two time periods, as much as 3 years apart. Heritability estimates between 0.4-0.8 indicate that IQ is a considerably heritable trait. It seems logical that genetic determinants of IQ should become less significant with age, as individuals become more experienced and educated. Surprisingly, the opposite is true. Until recently heritability of IQ was studied primarily among children, and so resulting estimates of heritability were misleading. It is now known that the heritability of IQ increases with age. Heritability estimates in more recent studies have been found to be as low as 0.20 in infancy and increase to about 0.40 in middle childhood, and still grow to as high as 0.80 as a person matures into adulthood (Plomin, DeFries, McClearn, and McGuffin 2001; Plomin, DeFries, Craig, and McGuffin 2003).

These seeming inconsistencies can be explained by understanding how genetic epidemiologists interpret heritability. First, a high heritability does not mean that the environment has no effect on the development of a trait, or that learning is not involved. For example, vocabulary size has shown substantial heritability (and high correlation with general intelligence) although an individual's vocabulary is wholly learned. In a society in which plenty of words are available in everyone's environment, especially for individuals who are motivated to seek them out, the number of words that individuals actually learn depends to a considerable extent on their genetic predispositions (Neisser, et al. 1996).

Another common error is to assume that because something is heritable it is necessarily unchangeable. Heritability does not imply immutability. As previously noted, heritable traits can depend on learning, and they may be subject to other environmental effects as well. The value of heritability can change if the distribution either of environmental factors
or genotypes changes substantially. For example, an impoverished or suppressive environment could fail to support the development of a trait, and hence restrict individual variation. Differences in variation of heritability are found between developed and developing nations. This could affect estimates of heritability (Neisser, et al. 1996).

There can be environmental changes that do not change heritability. If exposure to a negatively correlated environment risk factor declines equally for all members of a population, the mean value of the trait will rise without any change in its heritability, since relative environmental differences between individuals will remain unchanged. For example, Neisser, et al. (1996) report that heritability estimates of stature remain high and somewhat unchanged while average height continues to increase.

**Estimating Heritability**

There are various methods for calculating heritability that are specific to individual study designs. Such designs commonly require the sampling of twins, nuclear families, complex pedigrees, or adopted children. In some settings, the genetic and environmental effects are not resolvable and are estimated as a single heritable component that represents all additive effects that are transmitted from parents to offspring. The assumptions underlying these designs are also critical (Rice and Borecki 2001).

The twin study provides one of the earliest proposed and most straightforward designs for resolving genetic and cultural inheritance (Eaves 1977). The distinction between monozygotic (MZ) and dizygotic (DZ) twins provides the rationale to separate the genetic and environmental components of variance. By definition, MZ twins share 100% of
their genes, and DZ twins share only about 50% – the same as any non-twin sibling pair. In addition to analytic simplicity, a significant strength of this design is the inherent control for environmental factors, since both types of twins are assumed to share a common environment. In this setting, broad-sense heritability is estimated as twice the difference between the respective correlation coefficients for the quantitative phenotype among MZ and DZ twins (Ott 1999, p309).

An alternative to studies of twins is the ascertainment of parent-child clusters or other subsets of nuclear families (including sibships). For parents and offspring, that narrow-sense heritability can be estimated as twice the correlation coefficient between parent and offspring, as demonstrated by Fisher (1970) as well as Snedecor and Cochran (1967). For sibships with unknown parents, Snedecor and Cochran (1967) suggested estimating broad-sense heritability using the twice the intraclass correlation coefficient, which can be obtained from a one-way analysis of variance.

Assumptions about the underlying distributional properties of the data, as well as linearity and additivity are fundamental assumptions of most polygenic or multifactorial models, and complex traits are likely to be affected by factors that act non-additively. In the traditional simple pedigree of nuclear family data, familial resemblance cannot be resolved into genetic and environmental components, since family members share both genes and environment. The most straightforward proposed estimator of heritability in this setting is twice the correlation for first-degree relatives, since on average they share half of their genes (Rice and Borecki 2001).
An additional design of interest requires complex pedigrees, or multigenerational families. A simple pedigree is defined as any family in which no loops occur and in which, for any mating, at most one of the mates has his or her parents in the pedigree (Ott 1999, p183-4). A loop is said to exist in a pedigree when at least one path consists of a complete circle, leaving an individual by one line and returning by a different line. A complex pedigree is any family tree with one or more generations beyond a nuclear family, and comprises all non-simple pedigrees. To illustrate, Figure 1 below contains (a) a complex pedigree with a marriage loop in the third generation in which siblings in one family have married siblings of another family, and (b) a complex pedigree in which first cousins have married in the third generation, forming a so-called consanguineous loop.

Large families are ideal for Mendelian traits and also for studying allele-sharing (identical-by-descent, or IBD) relationships across generations (Ott, 1999). The data structure permits estimation of the genetic heritability without contamination by shared
environmental influences for at least two reasons. First, extended family members are unlikely to share environmental influences to any great extent; second, with a reasonable variety of relationships of differing degrees, there is a precise structuring of the expected covariances under a polygenic model (Rice, 2008). This is a significant advantage of the multigenerational design. When the cohort includes only nuclear families (with parents and non-twin children), it is impossible to separate out shared environment from certain types of genetic effects (Gur et al. 2007).

Unlike studies of twins, sib pairs, and nuclear families, complex pedigrees require more sophisticated variance components techniques. Such methods separate the variance of quantitative traits into components such as major genes, polygenes, random effects, and so forth (Amos, 1994; Almasy and Blangero, 1998). Most implementations of these methods rely on maximum likelihood, under the assumption that the phenotypes of all family members within a pedigree jointly follow a multivariate normal distribution (Hopper and Mathews, 1982). Hypotheses are then evaluated using the likelihood ratio test (Rice, 2008).

In many previously published studies, estimates of familial correlation have primarily been made for each family relationship (for example, parent–offspring, sibling–sibling), resulting in multiple estimates with differing power based on the number of pairs. In this approach, larger pedigrees have a distinct advantage over nuclear families in that more pair-wise combinations containing a broader array of genetic relationships are possible. The familial correlation (derived from the covariance among relatives) then estimates a fraction of the heritability of the trait (Fogarty, Rich, Hanna, Warram, and
However, in a larger pedigree, the increasingly large number of estimates seems to result in less clarity and more questions.

To overcome the deficiencies of that method, the before mentioned variance components approach is used instead. Variance component methods generally assume that traits are normally distributed and are particularly sensitive to kurtosis in the trait distribution (Allison, et al. 1999). Estimates of the means and variances of components of the models are obtained by maximum likelihood methods (Lange, Weeks, and Boehnke, 1988).

Once the variance components are estimated, it is fairly simple to calculate heritability. The variance of the random effect for the polygenic relationship is divided by the summed total of the variance of the continuous trait. This calculation is illustrated in a model in the Methods section.

**Heritability of Cognition and Cognitive Change**

Cognition and cognitive change have long been examined with respect to heritability. Table 1 summarizes the published findings from many of these studies, organizing their results in terms of study design. Cognitive performance can be measured using various tools. As described in the following section, the CCMS has used the Modified Mini-Mental State (3MS) exam. Metrics used by other studies are specified in Table 1.

Among the cited studies of heritability of cognition and heritability of change in cognition, many of the same confounders were controlled for, though nearly all of the studies found differing results as to which confounders had a significant effect on
Table 1 - Summary of heritability studies of cognition and cognitive change among the elderly.

<table>
<thead>
<tr>
<th>Design</th>
<th>Study</th>
<th>Cognitive Metric</th>
<th>Heritability Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Twins</td>
<td>Finkel, et al. 1998</td>
<td>Verbal, spatial, memory, and perceptual speed tests. (Covariates: Cohort, Time.)</td>
<td>0.60 - 0.80</td>
</tr>
<tr>
<td></td>
<td>Carmelli et al. 2002</td>
<td>MMSE, Digit Symbol Substitution, ISBMD, Color-Word Inference Test, TMT A and TMT B, CVLT. (Covariates: Age, Education).</td>
<td>0.32 - 0.78</td>
</tr>
<tr>
<td></td>
<td>Read et al. 2006</td>
<td>Synonyms test, Block Design test, Digit Span Forward and Backward, Thurstone's Picture Memory, Symbol Digit. (Covariates: Sex, Age, Education, Health)</td>
<td>0.34 - 0.68</td>
</tr>
<tr>
<td></td>
<td>McClearn et al. 1997</td>
<td>Verbal Meaning, Figure Logic, Block Design, Digit Span, Picture Memory, Information, Symbol Digit. (Covariates: Age, Gender)</td>
<td>0.32 - 0.62</td>
</tr>
<tr>
<td></td>
<td>Brandt et al. 1993</td>
<td>TICS-m (Covariate: Education)</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>McGue and Christensen 2001</td>
<td>MMSE (Covariates: Age (not significant))</td>
<td>0.42 - 0.52</td>
</tr>
<tr>
<td>Nuclear Families</td>
<td>Johnson et al. 2007</td>
<td>WAIS similarities subtest, Mattis Dementia Rating Scale, Selective Reminding Test, and BVRT (Covariates: Age, Sex, Education, APOE)</td>
<td>0.326</td>
</tr>
<tr>
<td>Complex Pedigrees</td>
<td>Gur et al. 2007</td>
<td>Penn Conditional Exclusion, Continuous Performance, Word Memory, and Face Memory tests; Visual Object Learning test, Judgement of Line Orientation</td>
<td>0.0 - 0.69</td>
</tr>
</tbody>
</table>
test, Emotion Intensity Discrimination test and a sensorimotor dexterity test. (Covariates: Age, Sex, Age², Age*Sex)

<table>
<thead>
<tr>
<th>Study</th>
<th>Cognitive Metric</th>
<th>Heritability Estimates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reynolds et al. 2002</td>
<td>Koh's Block Design, Thurstone’s Picture Memory, and Symbol Digit</td>
<td>0.71 - 0.77, 0.59 - 0.67, 0.82 - 0.93</td>
</tr>
<tr>
<td>Finkel et al. 2005</td>
<td>Verbal, spatial, memory, and processing speed tests. (Covariates: Sex, Age, Education, Self-Rated Health)</td>
<td>0.77 - 0.79</td>
</tr>
<tr>
<td>McGue and Christensen 2002</td>
<td>MMSE (Covariates: Age, Gender)</td>
<td>0.76 for cognition, 0.06 for linear change</td>
</tr>
</tbody>
</table>

heritability. The most common are Age, Sex, and Education. Each of the studies in Table 1 involved a sample population of elderly adults who were administered memory-related tests at least one time. The second half of the table describes those studies in which subjects were given tests of memory multiple times in order to chart the change in cognition. These studies have previously been conducted only under a Twin Study design.

These studies published in current literature all demonstrate that genetic factors explain more than 30% to as much as 80% of the variation in most cognitive abilities in adults. But, very few studies have been previously carried out involving change in cognitive function in complex pedigrees.
Investigations with simplex and multiplex families have supported an additive model in which increased genetic risk is accompanied by increased impairment in language, intelligence, verbal memory, visual reproduction, visual working memory, verbal learning, delayed visual recall, and perceptual- and pure-motor speed. In 2007, one study was published in which the neurocognitive profile of multiplex multigenerational families with schizophrenia was characterized in order to find heritability estimates of neurocognitive measures (Gur et al. 2007). Though this study involves persons aged 19 to 84, and is not simply a cohort of the elderly, it is the only study involving heritability of cognition in larger pedigrees.

The Cache County Memory Study

The Cache County Study on Memory, Health, and Aging was initiated in 1994 to investigate the association of APOE genotype and environmental exposures on cognitive function and dementia. A cohort initially comprised of 5,092 Cache County, Utah, residents (representing 90% of the Cache population aged 65 or older at the study's outset) has been followed continually for over twelve years. NIH-funded efforts have thus far supported four triennial waves of data collection for all surviving participants, along with additional clinical assessments for those at high risk for dementia. This combination of repeated general and clinical examinations has yielded a wealth of data, including measures of cognitive performance and dementia, along with risk factor data that provides information about demographic variables, medical conditions, medication and supplement use, and other environmental exposures.
The Cache County Family-Based Cohort Study on Aging (R01 AG18712; Professor Ronald Munger, PI) was one of supporting projects for the CCMS, funded through the NIA RFA AG-99-007, Exploratory Projects for Longitudinal Genetic Epidemiologic Studies on Aging. The purpose of this project was to build upon the existing resources of the Cache County study for genetic epidemiologic studies in part by linking the original 5092 Cache cohort members to the computerized genealogical records of the Utah Population Database (UPDB), gathering additional blood samples for future genetics studies, and contacting siblings of original cohort members to assess the feasibility of an expanded family-based study. The collaboration between the Cache County Study and investigators with the UPDB at the University of Utah has serendipitously uncovered the family structure underlying the original Cache sample: by linking original Cache participants with the comprehensive genealogical and vital records contained in the UPDB, a significant number of sibships was revealed that provide the potential for studies of the genetic epidemiology of age-related traits. The UPDB is a remarkable, integrated tool for exploring the genetic similarities among millions of members of the Utah Population. Its holdings include more than one million genealogical records dating back to the 1700s, a complete set of Utah birth certificates since 1947, all Utah death certificates since 1904, statewide cancer incidence data from 1966-present (from the Utah SEER registry), and current follow-up data on Utah residents supplied from Utah Driver License records and Medicare data (Wylie and Mineau, 2003; Skolnick, 1980). All told, the database contains over 8 million records, and many of the people represented by these records are linked together into pedigrees spanning as many as 10 generations by state-of-the-art record linkage software.
As the sibships identified by the UPDB comprise the data we will use for heritability estimation, their distribution is summarized in Table 1 of the following section. As already indicated, these sibships can be leveraged along with existing DNA samples for genetic studies of complex age-related traits, such as dementia or cognitive function. Estimating heritability for cognition and cognitive change is an important first step in describing any intra-familial correlation among Cache subjects with respect to memory loss, particularly after accounting for APOE genotype (a known risk factor for dementia). In addition to the sibships within the original Cache cohort, the UPDB further revealed complex pedigrees that include these sibships. An open research question is how knowledge of such second or higher degree relationships might increase power either for testing heritability or for future genetic association studies. Although the analyses described in the following sections focus only the sibships, an important objective of this project is to identify methods or tools that can be applied as well to the complex pedigrees.

<table>
<thead>
<tr>
<th>Parental Data</th>
<th>Sibship Size</th>
<th>Frequency</th>
<th>Total Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Parents</td>
<td>2</td>
<td>439</td>
<td>878</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>117</td>
<td>351</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>32</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>10</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>One Parent</td>
<td>2</td>
<td>7</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>612</td>
<td>1462*</td>
</tr>
</tbody>
</table>

* Three study subjects are individually represented in two different nuclear families (as a parent and as a sibling).
2. METHODS

As described in the previous section, our heritability analyses will utilize the sibships identified by the UPDB from among the original 5092 members of the Cache County Memory Study cohort. The distribution of these sibships is described in Table 2.

Continuous cognitive phenotypes are available from the Modified Mini-Mental State Exam or 3MS, (Teng and Chui, 1987; Tschanz et al., 2002) completed at the baseline interview by the majority (94%) of the 5092 participants. The 3MS is a 100-point adaptation of the Mini-Mental State Exam (MMSE; see Folstein, Folstein, and McHugh 1975) that renders the instrument more sensitive than the MMSE to change at upper and lower ends of a distribution. This instrument provides a measure of global cognitive ability that is sensitive to detecting dementia, (Hayden et al., 2003; Khachaturian, Gallo, and Breitner 2000) and consists of items that assess specific cognitive domains including learning, memory, executive functioning, orientation, abstract reasoning, expressive language ability, constructional praxis, and concentration. Global 3MS scores completed at all waves of the CCSMHA will provide the continuous cognitive phenotypes for analysis.

Additional environmental and demographic factors have also been collected on all 5092 cohort members at study entry, and updated where necessary for all surviving participants at each subsequent study wave. Each participant (or when severe dementia is present, a knowledgeable informant) provided information regarding demographic information, medical history, medication and supplement use, dietary intake, and other health habits. Table 3 provides summary information about the distributions of 3MS and other factors that we consider for these analyses, including age, sex, body mass index.
### Table 3 - Summary statistics and correlations between model covariates of interest

#### Summary Statistics

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Available</th>
<th>Missing</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proportion</td>
<td>Number</td>
<td>Proportion (#)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.4384</td>
<td>641</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.5609</td>
<td>820</td>
<td>0.00068(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Apoe e4 alleles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.6806</td>
<td>995</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.2661</td>
<td>389</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.0233</td>
<td>34</td>
<td>0.03009(44)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S.D.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>76.07</td>
<td>6.96</td>
<td>0.00068(1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transformed Age</td>
<td>5.81</td>
<td>2.92</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI</td>
<td>25.84</td>
<td>4.373</td>
<td>0.092(134)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>12.64</td>
<td>2.44</td>
<td>0.0048(7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline 3MS</td>
<td>88.61</td>
<td>9.68</td>
<td>0.056(82)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transformed Baseline 3MS</td>
<td>0.574</td>
<td>0.228</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change in 3MS</td>
<td>-0.4438</td>
<td>1.774</td>
<td>0.047(69)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transformed Change in 3MS</td>
<td>-0.920</td>
<td>2.526</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### Correlations Between Fixed Covariates and Cognitive Measures

<table>
<thead>
<tr>
<th>Correlation Coefficient / p-value (n)</th>
<th>3MS Score at Baseline</th>
<th>Change in 3MS Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.45774 &lt;0.0001 (1380)</td>
<td>-0.11119 &lt;0.0001 (1393)</td>
</tr>
<tr>
<td>BMI</td>
<td>0.10623 0.0001 (1326)</td>
<td>0.06974 0.0111 (1326)</td>
</tr>
<tr>
<td>Education</td>
<td>0.30004 &lt;0.0001 (1378)</td>
<td>0.02288 0.3939 (1391)</td>
</tr>
</tbody>
</table>
(BMI), years of education, and number of APOE ε4 alleles.

Before heritability analysis is possible, both computer programs that will be used necessitate that all variables considered in the analysis satisfy the assumption of normality. For this reason, several of the covariates required a linear or non-linear transformation. The measure of cognition, 3MS score, was transformed by standardizing the variable, as the score is on a 100 point scale. The change in 3MS score was normalized by a natural logarithm transformation, and the original age of the subjects at baseline was fairly skewed, and so was transformed by a positive root.

The majority of studies involving heritability of cognition in twins have used the structural equation modeling program Mx, and use the raw maximum likelihood estimation procedure. However, most studies involving non-twin pairs or complex pedigrees (for quantitative traits) use SOLAR. Studies involving nuclear families seem to be fairly divided mainly between these two programs.

As mentioned in the previous section, in practice the variance components approach is used both for nuclear families and for complex pedigrees. There are two widely used implementations of variance components freely available through the internet: the Statistical Analysis for Genetic Epidemiology (SAGE, 2008) and Sequential Oligogenic Linkage Analysis Routines (SOLAR, 2007) packages. The procedures for computing heritability in SAGE are implemented in its ASSOC module (SAGE, 2008). In the SAGE implementation, familial correlations and arbitrary covariates are incorporated into the likelihood by assuming the correlation structure described in Elston, George, and Severtson
(1992) under the multivariate normal regression model described by George and Elston (1987). This model for a continuous trait includes five random effects: a polygenic effect ($G_i$), a nuclear family effect ($F_i$), a marital effect ($M_i$), a sibship effect ($S_i$), and an individual (person-specific) environmental effect ($E_i$). Note that an individual belonging to numerous families (as a spouse, child, sibling, parent, or even through multiple marriages, and so on) will share a nuclear family effect with each family of which he is a member, a marital effect with each spouse, and a sib effect with his own siblings. Thus, for an individual, $i$, a continuous trait $Y_i$, (with vector of covariates $x_i$) the model is of the form:

$$h(Y_i) = h(\beta^T x_i) + G_i + F_i + M_i + S_i + E_i,$$

(1)

where $h$ is a transformation, either by Box-Cox or by George-Elston to form, in the case of the dependent variable, a response trait. It then follows that the total variance in the trait is equal to the of the variances of the five random effect components

$$V[h(Y)] = \sigma^2_{G} + \sigma^2_{F} + \sigma^2_{M} + \sigma^2_{S} + \sigma^2_{E}. $$

(2)

The respective components divided by the total variance can be interpreted as the intraclass correlations for families, sibships correlation, and environmental correlation; or interclass correlation in case of the marital correlation. Assuming polygenic variance, the polygenic effect, $\sigma^2_{G}$, divided by the total variance is termed heritability. The numerical algorithm for model fitting is described by Elston et al. (1992). Estimation of parameters is performed by jointly analyzing the independent pedigrees, maximizing the natural logarithm of the likelihood numerically, and summing the log likelihoods over all pedigrees.
The SOLAR implementation can be used for pedigrees of arbitrary size and complexity, applying a mixed effect model that incorporates random additive genetic effects along with fixed effects for potential confounders. (Lunetta, et al. 2007) Model fitting is based on maximum likelihood procedures. The observed covariances among family members are compared to the expected covariances based on shared DNA, and SOLAR estimates the proportion of the population variance for the trait resulting from additive genetic sharing. Likelihood ratio tests are used to assess the significance of both fixed and random effects. Heritability can be calculated as the additive genetic component (this assumes there is no gene-gene interaction) of variance (or polygenic component) from the covariance among relatives in SOLAR using the POLYGENIC procedure. While the polygenic heritability statistic yielded by SOLAR may harmonize with estimates produced by SAGE, as we demonstrate with the Cache sibships the results will differ due in part to how missing values are handled in each package.

3. RESULTS

Heritability analyses for the transformed cognition measure using SAGE and SOLAR yielded somewhat disparate results. Results for each with covariates are summarized in Table 4. As shown in Table 4, SAGE reported statistically significant random, polygenic, and sibling effects, but no significant family effect. The resulting heritability estimate was $0.465 (se = 0.067, p < 10^{-8})$. The final model included the significant covariates transformed Age, Sex, and APOE. These results are based on the inclusion of 1367 individuals from the 2735 in the pedigrees containing CCMS sibships. In order to quantify the effect of a nuclear family only, that is, to not include information that would make the
pedigree complex, the pedigree used for this analysis contains all ID numbers for parents of sibships, but no other information is included for anyone who is not part of a sibship. A few study participants had both children and siblings in the study, in which case, a person’s information as a parent is retained. For this reason, the analysis can only include individuals without a missing value for any one variable of interest.

When performing these analyses in SOLAR, the program output a warning that the estimated trait standard deviation was below 0.5, and recommended the trait be multiplied by a factor of 5.4. This was done and the analyses proceeded. SOLAR reported a heritability of 0.300 (std err = 0.066, p < 10^{-6}). After certain covariates (BMI, interactions between age, sex, education, and APOE) were tested and yielded insignificant p-values, the final model included covariates transformed age, APOE, and sex. The model estimates for these variables are included in Table 4. Though SAGE does not produce a like estimate, SOLAR reported that the proportion of variance due to these final covariates is 0.3489181. These results are based on the inclusion of 1420 individuals.

Table 4 – Results from each package on heritability of the transformed 3MS score

<table>
<thead>
<tr>
<th>Results from SAGE on transformed 3MS score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimate</td>
</tr>
<tr>
<td>Heritability (with covariates)</td>
</tr>
</tbody>
</table>
### Heritability (without covariates)

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.01e-06</td>
<td>8.14e-08</td>
<td>&lt; 10^{-6}</td>
</tr>
</tbody>
</table>

### Variance Components

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>0.020360</td>
<td>0.002460</td>
<td>&lt; 10^{-6}</td>
</tr>
<tr>
<td>Polygenic</td>
<td>0.017686</td>
<td>0.002821</td>
<td>&lt; 10^{-6}</td>
</tr>
<tr>
<td>Sibling</td>
<td>1.00x10^{-7}</td>
<td>Unavailable</td>
<td>Unavailable</td>
</tr>
</tbody>
</table>

### Covariates

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transformed Age</td>
<td>-0.035248</td>
<td>0.001890</td>
<td>&lt; 10^{-6}</td>
</tr>
<tr>
<td>APOE</td>
<td>-0.043234</td>
<td>0.010644</td>
<td>4.87x10^{-5}</td>
</tr>
<tr>
<td>Sex</td>
<td>-0.049969</td>
<td>0.010187</td>
<td>&lt; 10^{-6}</td>
</tr>
</tbody>
</table>

### Results from SOLAR on transformed 3MS score

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heritability (with covariates)</td>
<td>0.2995</td>
<td>0.066</td>
<td>&lt; 10^{-6}</td>
</tr>
<tr>
<td>Heritability (without covariates)</td>
<td>0.4584438</td>
<td>0.0650868</td>
<td>&lt; 10^{-6}</td>
</tr>
</tbody>
</table>

### Covariates

<table>
<thead>
<tr>
<th></th>
<th>Estimate</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transformed Age</td>
<td>-0.1629396</td>
<td>0.0154376</td>
<td>&lt; 10^{-6}</td>
</tr>
<tr>
<td>Transformed Age^2</td>
<td>-0.0138373</td>
<td>0.0027930</td>
<td>0.02091</td>
</tr>
<tr>
<td>Education</td>
<td>0.1432207</td>
<td>0.0123364</td>
<td>&lt; 10^{-6}</td>
</tr>
<tr>
<td>APOE</td>
<td>-0.3415432</td>
<td>0.0575035</td>
<td>0.0000176</td>
</tr>
<tr>
<td>Sex</td>
<td>0.3136087</td>
<td>0.0582426</td>
<td>&lt; 10^{-6}</td>
</tr>
<tr>
<td>Transformed Age*Sex</td>
<td>-0.0653245</td>
<td>0.0199166</td>
<td>0.04936</td>
</tr>
</tbody>
</table>
As discussed in Section 2, many previous studies have focused specifically on rate of
cognitive change. Rapid memory loss is highly correlated with dementia onset, and thus
represents an additional phenotype of interest. Analyzing repeated measures with respect
to heritability is an open research area. Few if any family-based studies have attempted to
use anything more sophisticated than a summary measure as a focus of inference. An
objective of this project is to identify more sophisticated methods for estimating
heritability of 3MS change. Burton, Scurrah, Tobin, and Palmer (2005) suggest a promising
Bayesian approach that we will pursue in future work. Their implementation uses the
Bayesian-oriented WinBUGS 1.4 software package (Spiegelhalter, Thomas, Best, and Lunn
2003). For this analysis, we confined ourselves to a two-stage summary response
approach, where a line is fit for each individual, representing change in 3MS with respect to
time, and the fitted slope is used as the quantitative trait.

Results for rate of memory loss are summarized in Table 5. When evaluated in the
S.A.G.E. package, with inclusion of the same covariates, the heritability index was estimated
as 0.215456 (std err = 0.0873, p =0.006836). Under this model, there was again significant
random, polygenic, and sibling effects, but the family effect was insignificant and excluded.
One major difference in the results when using the transformed slope value is that the sex
covariate was not found to be statistically significant. Further, this analysis included only
964 individuals with complete information.

When the same evaluation was made using the SOLAR package, 966 individuals
were included and the heritability was calculated as 0.1582733 (std err = 0.083, p
Table 5 - Results of heritability of the transformed change in 3MS score

<table>
<thead>
<tr>
<th>Heritability</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(with covariates)</td>
<td>0.215456</td>
<td>0.087379</td>
<td>0.006836</td>
</tr>
<tr>
<td>(without covariates)</td>
<td>1.62e-08</td>
<td>Unavailable</td>
<td>Unavailable</td>
</tr>
</tbody>
</table>

Variance Components

<table>
<thead>
<tr>
<th>Source</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>4.253210</td>
<td>0.491698</td>
<td>&lt;10^{-6}</td>
</tr>
<tr>
<td>Polygenic</td>
<td>1.168042</td>
<td>0.482610</td>
<td>0.007755</td>
</tr>
<tr>
<td>Sibling</td>
<td>1.00x10^{-7}</td>
<td>Unavailable</td>
<td>Unavailable</td>
</tr>
</tbody>
</table>

Covariates

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transformed Age</td>
<td>0.286489</td>
<td>0.026688</td>
<td>&lt;10^{-6}</td>
</tr>
<tr>
<td>APOE</td>
<td>0.341442</td>
<td>0.132151</td>
<td>0.009774</td>
</tr>
<tr>
<td>Sex</td>
<td>-0.122558</td>
<td>0.133860</td>
<td>0.359891</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Heritability (with covariates)</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(with covariates)</td>
<td>0.15827</td>
<td>0.088</td>
<td>0.03139</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Heritability (without covariates)</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(without covariates)</td>
<td>0.2089784</td>
<td>0.0899317</td>
<td>&lt;10^{-6}</td>
</tr>
</tbody>
</table>

Covariates

<table>
<thead>
<tr>
<th>Covariate</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transformed Age</td>
<td>0.0154376</td>
<td>0.0000152</td>
<td></td>
</tr>
</tbody>
</table>
=0.03139). The model included covariates transformed age at first evaluation, years of education, and the age at first evaluation by sex interaction. The proportion of variance due to all final covariates was 0.1199559.

4. DISCUSSION

From the Cache sibships, we estimated heritability of cross-sectional cognition in old age to be between 29% to 47%, and heritability for cognitive decline to be between 16-22%. The disparities between SAGE and SOLAR seem to arise from how missing values are handled and how individuals are counted. SAGE results for cognition were based on 1367 study participants, while SOLAR used 1420 observations. For instance, when a family member has a missing value for any variable of interest, that individual is entirely excluded from the analysis in SAGE, although information regarding that individual’s familial relationships is still used. In SOLAR, all data on all participants in any valid pedigree is retained, but if any combination of missing values renders a pedigree invalid, then all members of the pedigree are excluded from the analysis.

Though the results from SAGE and SOLAR are not close enough to one another to say that the estimate of cognitive heritability is become more refined, the estimates are definitely within reasonable range of many of the results found through previous work.
These data provide additional support to earlier findings that patterns of change in total variance for cognitive abilities in late adulthood are measure dependent (Finkel, Pedersen, McClearn, Plomin, and Berg 1996; Morse, 1993; Finkel, Pedersen, Plomin, and McClearn 1998). We have used the global 3MS scale, ignoring the subscales that comprise the summands for 3MS. These individual subscales purportedly measure more specific aspects of cognitive function, and most previous heritability analyses have focused on these subscales (e.g., such as verbal memory, sensorimotor, or emotional testing components). Examination of these more refined metrics is an area for future work.

Some have expressed concern about the confounding effects of age and gender when exploring heritability of neurocognitive traits (Gur, et al. 2007). The incorporation of these effects in the present analysis by inclusion of covariates in the model development has accounted for this concern, and the effects observed were consistent with the current literature. Most importantly, our results indicate that genetic variability can in fact be established after accounting for these effects.

Another issue to address is the use of repeated measures. The heritability of rates of change as phenotypes has received surprisingly little attention in the literature. As mentioned previously, there are one or two recently developed approaches that treat longitudinal data in a more appropriately sophisticated way – applying these methods to the Cache families is another priority. There are obvious limitations with the conventional summary response approach we have used in treating fitted slopes as phenotypes. The assumption of linearity may be tenuous. Moreover, we are giving each individual equal weight, although some have more 3MS measures than others.
A final significant objective for further analysis is the incorporation of complex pedigrees. We have found that SAGE will use familial information regardless of whether phenotype or environmental data are missing. The question is whether additional knowledge regarding higher-degree familial relationships – without knowledge of phenotypes – improves either the power or other properties of heritability estimates. For example, the UPDB at present can determine who among the Cache cohort are cousins by identifying common ancestors not in the Cache study, but the ancestors’ traits are unknown. The potential efficiency gained from this additional information must be systematically explored.

**Acknowledgements**

The SAGE software package is supported by a U.S. Public Health Service Resource Grant (RR03655) from the National Center for Research Resources.
5. LITERATURE CITED


APPENDIX - SAGE AND SOLAR COMMANDS AND RESULTS

SAGE Parameter File Commands:

```python
pedigree,
{
  delimiters=""
  delimiter_mode="single"
  individual_missing_value=""
  sex_code_male="m",female="f",trait,unknown=""
#Variables in the pedigree file follow
  individual_id="UPDBID"
  pedigree_id="PEDIGREE"
  parent_id="PARENT1ID"
  parent_id="PARENT2ID"
  sex_field="SEX"
  trait=DEMENTIA,binary,affected="1",unaffected="0",missing="".
  trait=ONSETAGE,continuous,missing="".
  trait=MA,binary,affected="1",unaffected="0",missing="".
  trait=ONSETAGEMA,continuous,missing="".
  trait=V1MSADJ,continuous,missing="".
  trait=V1ABSTR,continuous,missing="".
  trait=V1ATTN,continuous,missing="".
  trait=V1COMMAND,continuous,missing="".
  trait=V1DELREC,continuous,missing="".
  trait=V1FLUENCY,continuous,missing="".
  trait=V1LTM,continuous,missing="".
  trait=V1PENT,continuous,missing="".
  trait=V1PLACE,continuous,missing="".
  trait=V1NAME,continuous,missing="".
  trait=V1REG,continuous,missing="".
  trait=V1REPEAT,continuous,missing="".
  trait=V1SENT,continuous,missing="".
  trait=V1TIME,continuous,missing="".
  trait=MMMS,continuous,missing="".
  trait=ABSTR,continuous,missing="".
  trait=ATTN,continuous,missing="".
  trait=COMMAND,continuous,missing="".
  trait=DELREC,continuous,missing="".
  trait=FLUENCY,continuous,missing="".
  trait=LTM,continuous,missing="".
  trait=PENT,continuous,missing="".
  trait=PLACE,continuous,missing="".
  trait=NAME,continuous,missing="".
  trait=REG,continuous,missing="".
  trait=REPEAT,continuous,missing="".
  trait=SENT,continuous,missing="".
  trait=COGTIME,continuous,missing="".
  covariate=APOENUM,categorical,missing="".
  covariate=V1AGE,continuous,missing="".
  covariate=EDUC,categorical,missing="".
  covariate=BM1,continuous,missing="".
  covariate=transAGE,continuous,missing="".
  trait=transMS,continuous,missing="".
  trait=transMMMS,continuous,missing="".
  trait=newtransMS,continuous,missing="".
  covariate=agesq,continuous,missing="".
}
```
Results from SAGE on transformed3MS:

COPYRIGHT (C) 2005 CASE WESTERN RESERVE UNIVERSITY

Results

Analysis description

Title 7.15 with vars in solar
Primary Trait transformed3MS (Quantitative)
Allow averaging Disabled
Limit output Disabled (output includes complete model information)

Transformation configuration

Method George-Elston
Lambda1 is estimated
Lambda2 is fixed at 0.000000

Transformed components: Intercept transAGE APOENUM SEX_CODE
Non-transformed components:

Model 'Baseline'

Sample description

Number of individuals in dataset 2735
Number of constituent pedigrees in dataset 607
Number of singletons in dataset 58
Number of constituent pedigree members with complete information 1367
Number of singletons with complete information 1

Note: missing data replaced by averages are considered complete.

<table>
<thead>
<tr>
<th>Covariates</th>
<th>Name</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>transAGE</td>
<td>5.704880</td>
<td>2.882638</td>
<td>0.000000</td>
<td>0.000000</td>
<td>14.389677</td>
</tr>
<tr>
<td>APOENUM</td>
<td>0.309211</td>
<td>0.504515</td>
<td>0.000000</td>
<td>0.000000</td>
<td>2.000000</td>
</tr>
<tr>
<td>SEX_CODE</td>
<td>0.558480</td>
<td>0.496568</td>
<td>0.000000</td>
<td>0.000000</td>
<td>1.000000</td>
</tr>
</tbody>
</table>

MAXIMIZATION RESULTS Baseline without test covariates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>S.E.</th>
<th>P-value</th>
<th>Deriv</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variance components</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>0.020360</td>
<td>0.002460</td>
<td>&lt; 1.00e-07</td>
<td>-0.0065751781</td>
</tr>
<tr>
<td>Polygenic</td>
<td>0.017686</td>
<td>0.002821</td>
<td>&lt; 1.00e-07</td>
<td>-0.0032390359</td>
</tr>
<tr>
<td>Sibling</td>
<td>1.00e-07</td>
<td>Ind. func. fixed @ bnd</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other parameters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total variance</td>
<td>0.038046</td>
<td>0.001510</td>
<td>&lt; 1.00e-07</td>
<td>-0.0002330552</td>
</tr>
<tr>
<td>Heritability</td>
<td>0.464854</td>
<td>0.067023</td>
<td>&lt; 1.00e-07</td>
<td>-0.0004251100</td>
</tr>
<tr>
<td>Correlations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full sibs</td>
<td>0.232429</td>
<td>0.033512</td>
<td>&lt; 1.00e-07</td>
<td>0.0000000000</td>
</tr>
<tr>
<td>Half sibs</td>
<td>0.116213</td>
<td>0.016756</td>
<td>&lt; 1.00e-07</td>
<td>0.0000000000</td>
</tr>
<tr>
<td>Parent offspring</td>
<td>0.232427</td>
<td>0.033512</td>
<td>&lt; 1.00e-07</td>
<td>0.0000000000</td>
</tr>
<tr>
<td>Sibship</td>
<td>4.91e-06</td>
<td>5.93e-07</td>
<td>&lt; 1.00e-07</td>
<td>0.0000000000</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.586876</td>
<td>0.006374</td>
<td>&lt; 1.00e-07</td>
<td>0.0000000000</td>
</tr>
<tr>
<td>Covariates</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>transAGE</td>
<td>-0.035248</td>
<td>0.001890</td>
<td>&lt; 1.00e-07</td>
<td>0.0000000000</td>
</tr>
<tr>
<td>APOENUM</td>
<td>-0.043234</td>
<td>0.010644</td>
<td>4.87e-05</td>
<td>0.0000000000</td>
</tr>
<tr>
<td>SEX_CODE</td>
<td>0.049969</td>
<td>0.010187</td>
<td>&lt; 1.00e-06</td>
<td>0.0000000000</td>
</tr>
<tr>
<td>Transformation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lambda1</td>
<td>2.106346</td>
<td>0.166495</td>
<td>&lt; 1.00e-07</td>
<td>0.0000000000</td>
</tr>
<tr>
<td>Lambda2</td>
<td>0.000000</td>
<td>Fixed</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Final ln likelihood: -18384.509346
Commands for SOLAR on newtransformed3MS:

% solar
solar> load pedigree nuclear_family_phen.csv
solar> load phenotypes nuclear_family_phen.csv
solar> trait newtransformed3MS
solar> covariate transAGE transAGE^2 BMI EDUC APOENUM sex transAGE*sex transAGE^2*sex
solar> polygenic -screen

Results from SOLAR on newtransformed3MS:

Pedigree: nuclear_ped.csv
Phenotypes: nuclear_phen.csv
Trait: newtransformed3MS Individuals: 1420

H2r is 0.2995234 p = 0.0000009 (Significant)
H2r Std. Error: 0.0660259

- transAGE p = 1.9220993e-24 (Significant)
- transAGE^2 p = 0.0209089 (Significant)
- BMI p = 0.5542592 (Not Significant)
- EDUC p = 1.1679937e-30 (Significant)
- APOENUM p = 0.0000176 (Significant)
- sex p = 2.2346077e-07 (Significant)
- transAGE*sex p = 0.0493649 (Significant)
- transAGE^2*sex p = 0.6018304 (Not Significant)
- transAGE*BMI p = 0.6813912 (Not Significant)
- transAGE*EDUC p = 0.4867394 (Not Significant)
- transAGE*APOENUM p = 0.8111754 (Not Significant)

The following covariates were removed from final models:
BMI transAGE^2*sex transAGE*BMI transAGE*EDUC transAGE*APOENUM

Proportion of Variance Due to All Final Covariates is
0.3489181

From SAGE on transchange3MS:

COPYRIGHT (C) 2005 CASE WESTERN RESERVE UNIVERSITY

Results

====================
Analysis description

Title on transMMMS no fe
Primary Trait transchange3MS (Quantitative)
Allow averaging Disabled
Limit output Disabled (output includes complete model information)

Transformation configuration

Method George-Elston
Lambda1 is estimated
Lambda2 is fixed at 0.000000

Transformed components: Intercept transAGE APOENUM SEX_CODE
Non-transformed components:

Model 'Baseline'

Sample description

Number of individuals in dataset 2735
Number of constituent pedigrees in dataset 607
Number of singletons in dataset 58
Number of constituent pedigree members with complete information 964
Number of singletons with complete information 1

Note: missing data replaced by averages are considered complete.

Covariates

Name Mean Std. dev. Min Max
--------- ----------- ----------- --------- ---------
transAGE 5.146962 2.687002 0.000000 14.389677
APOENUM 0.310881 0.501537 0.000000 2.000000
SEX_CODE 0.573057 0.494634 0.000000 1.000000

MAXIMIZATION RESULTS Baseline without test covariates
### Parameter Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>S.E.</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variance components</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>4.253210</td>
<td>0.491698</td>
<td>&lt;1.00e-07</td>
</tr>
<tr>
<td>Polygenic</td>
<td>1.168042</td>
<td>0.482610</td>
<td>0.007755</td>
</tr>
<tr>
<td>Sibling</td>
<td>1.00e-07</td>
<td>Ind. func. fixed</td>
<td>bnd</td>
</tr>
<tr>
<td>Other parameters</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total variance</td>
<td>5.421251</td>
<td>0.248338</td>
<td>&lt;1.00e-07</td>
</tr>
<tr>
<td>Heritability</td>
<td>0.215456</td>
<td>0.087379</td>
<td>0.006836</td>
</tr>
<tr>
<td>Correlations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full sibs</td>
<td>0.107728</td>
<td>0.043690</td>
<td>0.013672</td>
</tr>
<tr>
<td>Half sibs</td>
<td>0.053864</td>
<td>0.021845</td>
<td>0.013672</td>
</tr>
<tr>
<td>Parent offspring</td>
<td>0.107728</td>
<td>0.043690</td>
<td>0.013672</td>
</tr>
<tr>
<td>Sibship</td>
<td>2.35e-08</td>
<td>Unavailable</td>
<td>Unavailable</td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.762178</td>
<td>0.073201</td>
<td>&lt;1.00e-07</td>
</tr>
<tr>
<td>Covariates</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>transAGE</td>
<td>0.286489</td>
<td>0.026688</td>
<td>&lt;1.00e-07</td>
</tr>
<tr>
<td>APOENUM</td>
<td>0.341442</td>
<td>0.132151</td>
<td>0.009774</td>
</tr>
<tr>
<td>SEX_CODE</td>
<td>-0.122558</td>
<td>0.133860</td>
<td>0.359891</td>
</tr>
<tr>
<td>Transformation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lambda1</td>
<td>0.663163</td>
<td>0.049974</td>
<td>&lt;1.00e-07</td>
</tr>
<tr>
<td>Lambda2</td>
<td>0.000000</td>
<td>Fixed</td>
<td></td>
</tr>
</tbody>
</table>

Final ln likelihood: -19186.715751

---

**From SOLAR on transchange3MS:**

Pedigree: nuclear_ped.csv  
Phenotypes: nuclear_phen.csv  
Trait: transchange3MS  
Individuals: 966

H2r is 0.1582733  p = 0.0313914 (Significant)  
H2r Std. Error: 0.0882923

transAGE  p = 0.0000152 (Significant)  
EDUC  p = 0.0128187 (Significant)  
BMI  p = 0.7422005 (Not Significant)  
SEX  p = 0.2621021 (Not Significant)  
transAGE*SEX  p = 0.0254981 (Significant)  
transAGE^2  p = 0.2780199 (Not Significant)  
transAGE^2*sex  p = 0.5222651 (Not Significant)

The following covariates were removed from final models:  
BMI SEX transAGE^2 transAGE^2*sex
Proportion of Variance Due to All Final Covariates Is 0.1199559

Residual Kurtosis is 0.7578, within normal range