Genome-Wide Association Study of Brain Arteriolosclerosis

Lincoln M.P. Shade1

Yuriko Katsumata1,2

Timothy J. Hohman3

Julie A. Schneider4

Kwangsik Nho5

Andrew Saykin5

Shubhabrata Mukherjee6

Kevin L. Boehme7

John Kauwe8

Peter T. Nelson2,9

David W. Fardo1,2

Alzheimer’s Disease Genetics Consortium

for the Alzheimer’s Disease Neuroimaging Initiative\*

2021-05-20

# Author Affiliations

1Department of Biostatistics, College of Public Health, University of Kentucky, Lexington, KY

2Sanders-Brown Center on Aging and Alzheimer’s Disease Research Center, University of Kentucky, Lexington, KY

3Vanderbilt Memory & Alzheimer’s Center, Department of Neurology, Vanderbilt University Medical Center, Nashville, Tennessee

4Departments of Neurology and Pathology, Rush University Medical Center, Chicago, IL

5Department of Radiology & Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN

6Department of Medicine, University of Washington, Seattle, WA

7ARUP Laboratories, Salt Lake City, UT

8Office of the President, Brigham Young University–Hawaii, Laie, HI

9Pathology and Laboratory Medicine, University of Kentucky, Lexington, KY

\*Data used in preparation of this article were obtained from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: <http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf>

## Corresponding Author

Lincoln Shade

Email: [lincoln.shade@uky.edu](mailto:lincoln.shade@uky.edu)

Phone: 859-893-5798

# Abstract

Brain arteriolosclerosis (B-ASC) is a small-vessel cerebrovascular disease involving sclerotic thickening of arterioles in the brain. The genetics of other contributors of dementia have been studied extensively, but to date no in-depth study has been conducted on genetic risk of autopsy-proven B-ASC. We perform the first genome-wide association study (GWAS) on B-ASC using multiple aged neuropathologic cohorts. Initial Stage 1 GWAS (*n* = 3382) and Stage 2 mega-analysis (*n* = 4569) were performed using data from the two largest cohorts. Replication of top variants and additional Stage 3 mega-analysis were performed incorporating two smaller cohorts: (*n* = 47 and 512). Lead variants in the top two loci in the Stage 1 mega-analysis (rs7902929, *p* = 1.8× 10-7 ; rs2603462, *p =* 4× 10-7) replicated in the first replication cohort (rs7902929, *p* = 0.01174; rs2603462, *p =* 0.01154). The rs2603462 lead variant colocalizes with *ELOVL4* expression in the cerebellum (posterior probability = 92.9%). Our study constitutes the most in-depth and comprehensive study of the genetic risk of brain arteriolosclerosis neuropathologic endophenotype to date.

# Keywords

Please provide 4 to 6 keywords which can be used for indexing purposes.

# Introduction

Brain arteriolosclerosis (B-ASC) is a subtype of cerebral small vessel neuropathologic change characterized by thickening of arteriole walls in the brain. These changes can include hypertrophy or atrophy of vascular smooth muscle and luminal extracellular deposition of elastin and collagen [[1](#ref-ighodaro2017)–[3](#ref-buchman2013)]. B-ASC is commonly found in autopsied elderly individuals, with more than 50% of participants showing some B-ASC pathology in large autopsy studies [[1](#ref-ighodaro2017), [3](#ref-buchman2013), [4](#ref-chou2013) ]. B-ASC is associated with multiple neuropathologies including Alzheimer’s disease (AD), limbic-predominant age-related TDP-43 encephalopathy (LATE), micro-infarcts, and large-vessel infarcts [[1](#ref-ighodaro2017), [5](#ref-arvanitakis2016), [6](#ref-neltner2014) ]. B-ASC is also associated with cognitive decline after adjusting for age and the presence of other neuropathologies [[1](#ref-ighodaro2017), [5](#ref-arvanitakis2016)]. Despite the clinical importance of B-ASC, its risk factors, other than age and sex [[1](#ref-ighodaro2017)], remain largely uncharacterized. Hypertension (HTN) and diabetesmellitus (DM) are established clinical risk factors for arteriolosclerosis in the kidneys, but their association with B-ASC is inconsistent; in an autopsy based cohort stratified by age at death (<80 years vs. ≥80 years), hypertension was significantly associated with B-ASC in the younger group only, and diabetes was not associated with B-ASC in either group [[1](#ref-ighodaro2017), [7](#ref-wu2005)–[9](#ref-cameron2006)].

Genome-wide association studies (GWAS) have provided a powerful resource for investigating genomic risk of complex diseases through analysis of millions of genetic variants with diseases of interest, and have to date identified tens-of-thousands of variants associated with disease [[10](#ref-buniello2019)]. GWAS have successfully been used to identify genetic risk loci for stroke and for imaging-based phenotypes of cerebrovascular disease such as white matter hyperintensities (WMH) and brain infarcts; however, we found no published GWAS for autopsy-proven B-ASC to date [[11](#ref-traylor2016)–[13](#ref-chauhan2019)]. Given the unique and complex structures of brain arterioles and their associated structures, including astrocytes and other components of the blood-brain barrier, there is reason to suspect that the genomic risk for B-ASC may not be wholly shared with other cerebrovascular phenotypes. In one study of Religious Orders Study/Memory and Aging Project (ROSMAP) participants, 167 independent genetic variants previously meeting genome-wide significance threshold (P < 5×10^-8^) in GWAS of stroke or stroke risk factorswere tested for association with B-ASC pathology [[4](#ref-chou2013)]. The authors found six variants nominally associated with B-ASC at the 0.01 < *p* < 0.05 significance level, none of which remained significant after correction for multiple tests [[4](#ref-chou2013)]. Investigating genomic risk factors of B-ASC at the genome-wide level may provide important insights into its pathophysiological development as well as its relationship to neuroimaging and other neuropathological phenotypes.

In the present study, we conduct a GWAS using B-ASC pathology as the endophenotype analyzing neuropathology and genotype data across four aged autopsy cohorts. In Stage 1, we analyzed data from the National Alzheimer’s Coordinating Center (NACC) Neuropathology data set linked to genotype data from the Alzheimer’s Disease Genetics Consortium (ADGC). In Stage 2, we attempted to replicate top variants from the NACC GWAS in ROSMAP and then mega-analyzed both data sets together. In Stage 3, we attempted to replicate the top variants from Stage 2 in two smaller autopsy cohorts consisting of Alzheimer’s Disease Neuroimaging Inititative (ADNI) and Adult Changes in Thought (ACT) participants. To investigate potential biologically functional correlates to disease risk, we then perform colocalization analysis on B-ASC associated variants identified as quantitative trait loci (QTL) using data from the Genetic Tissue Expression (GTEx) Project and gene-based association analyses [[14](#ref-thegeno2013)]. To investigate the possibility of hypertension and diabetes acting as mediators for B-ASC risk variants, we also performed mediation analyses on a subset of NACC participants with clinical data available. An outline of our study design is shown below in **Figure 1**.

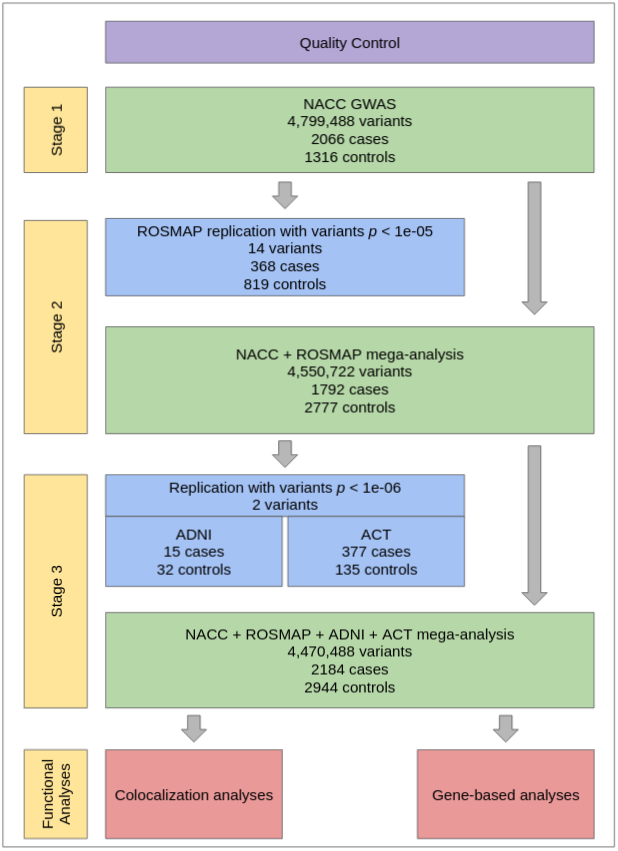


Figure 1: Study Design Outline

# Methods

## Study Participants

In our initial GWAS, we linked neuropathology data from participants in >30 National Institute on Aging-funded Alzheimer’s Disease Research Centers (ADRCs) studies from the NACC Neuropathology Data Set to ADGC genotype data [[15](#ref-welcome), [16](#ref-besser2018)]. Each ADRC has its own recruitment strategies, populations, and study design, and data are collected and aggregated by NACC. We then excluded all participants diagnosed with any of 19 rare neurological conditions (see **Supplementary Table S1** for full exclusion criteria). Similar exclusion criteria were not applied to other studies’ participants due to lack of data availability. A total of 3520 participants had both B-ASC neuropathology and genotype data available and passed initial inclusion criteria.

The ROSMAP study has been previously described in detail and consists of harmonized data from two longitudinal cohorts: The Religious Orders Study (ROS) and the Memory and Aging Project (MAP) [[17](#ref-bennett2018) ]. ROS began in 1994 and recruited older Catholic Priests and Sisters from around the United States. MAP began in 1997 and recruited older adults who at the time had no diagnosis of dementia. A total of 1213 ROSMAP participants had both autopsy and genotype data available.

The ADNI ([adni.loni.usc.edu](http://adni.loni.usc.edu/)) was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early AD. A subset of ADNI participants undergo autopsy and receive neuropathological phenotyping. A total of 60 ADNI participants had both B-ASC and genotype data available.

The ACT study began in 1994 and recruited residents in the greater Seattle area aged 65 years and older without dementia at time of enrollment [[18](#ref-kukull2002)]. The goal of the study has expanded to include three cohorts and continuous enrollment using the same enrollment criteria and has a current total of 4,960 particpants across all three cohorts. A total of 518 ACT participants had both B-ASC and genotype data available.

## Definitions of B-ASC variables used

In the NACC Neuropathology Data Set, B-ASC is graded as an ordinal variable with possible values of 0 (none), 1 (mild), 2 (moderate), or 3 (severe). Grading is performed by trained Neuropathologists at each ADRC and is a global rating, meaning no instruction is given to examine specific brain regions for B-ASC. Both the ADNI and ACT studies follow the NACC Neuropathology Form when grading neuropathological phenotypes. The B-ASC variable in ROSMAP is graded on B-ASC histological changes exclusively in the basal ganglia. Vessel intimal pathology is first graded on a scale from 0 (none) to 6 (severe), and then collapsed to a four-level ordinal variable with the same labels as in NACC [[19](#ref-buchman2011)]. ADNI and ACT autopsy procedures follow the NACC Neuropathology codebook, though B-ASC in ACT is labeled 1-4 rather than 0-3.

## Identifying ethnic outliers

In all cohorts, we performed principal component analysis (PCA) in PLINK v1.9 [[20](#ref-chang2015)] using a pruned subset of independent (linkage disequilibrium (LD) r2 < 0.05) variants from each data set merged to data from the 1000 Genomes Project Phase 3 (1000 Genomes, n = 2504) [[21](#ref-Abecasis2012)]. All variants were checked to ensure that major and minor alleles matched in the study data sets and 1000 Genomes. We then plotted the standardized first and second principal components (PCs) for each participant using the ggplot2 R package in R version 4.0.4 [[22](#ref-rcoreteam2020)]. All participants whose plot positions were located within a Euclidean distance of 0.35 from the mean plot positions of 1000 Genomes participants in the EUR superpopulation were considered to be of European descent and were included in analyses. We then re-ran PCA for the included participants and used included the first five PCs to use as covariates in regression models.

## Quality control (QC) of participants and genotype data

Imputation for ADGC was performed using the Haplotype Reference Consortium (HRC) panel [[23](#ref-mccarthy2016)] (workflow can be found at <https://github.com/KBoehme/ADGC_HRC_MERGE>). The main ROSMAP genotype QC workflow has been previously described [[24](#ref-dumitrescu2020)]. Briefly, imputation was performed for ROSMAP on the Michigan Imputation Server [[25](#ref-das2016)] using the HRC panel. Imputation for ACT was performed on the Michigan Imputation Server using the Trans-Omics for Precision Medicine (TOPMed) reference panel [[26](#ref-fuchsberger2015), [27](#ref-taliun2019)]. We performed additional standard QC procedures on all genotyping data using PLINK v1.9 and KING [[20](#ref-chang2015), [28](#ref-chang)–[31](#ref-marees2018)]. Variants were excluded if they had (1) a minor allele frequency (MAF) less than 5%; (2) a call rate of less than 95%; or (3) a Hardy-Weinberg equilibrium exact test *p* < 10-6 using all participants in a cohort.

Participants were removed if they had a genotype missingness rate above 5% or a heterozygosity rate more than 3 standard deviations above or below the mean of their cohort. All participants were checked for duplicate genotype information across studies. All participants with duplicate information within one cohort were removed, while participants with duplicate information across cohorts were preferentially kept in the data set analyzed first and removed from cohorts used later in the study (*i.e.* NACC > ROSMAP > ADNI and ACT). Of participants with a high degree of relatedness estimated using identical by descent (IBD) indicated by 2nd-degree relation (proportion IBD > 0.18) or closer in PLINK, all but the participant with the highest genotyping rate in each related cluster were removed, with ties broken randomly. In NACC analyses, 4799488 variants and 3382 participants passed QC protocols. In the ROSMAP data set, 4889494 variants and 1187 participants passed QC protocols. Thirteen ADNI participants were identified as NACC duplicates and removed, leaving 47 participants for analysis. A total of 512 ACT participants passed QC measures.

## Statistical analyses

### Single-variant analyses

We first performed single-variant association analyses in the NACC participants using logistic regression in PLINK v1.9 and ordinal regression in R using the MASS package [[22](#ref-rcoreteam2020), [28](#ref-chang)]. To create a dichotomous outcome variable for logistic regression, we counted participants with none or mild B-ASC as controls and those with moderate or severe B-ASC as cases. This decision was made based on previous studies that used these cut points and studies that found that moderate-to-severe B-ASC was associated with worse cognitive functioning [[1](#ref-ighodaro2017), [5](#ref-arvanitakis2016)]. Covariates in the regression models included age at death, sex, ADGC cohort indicators, and the first five PCs of the genetic relatedness matrix. An additive mode of inheritance was assumed in all analyses. We used a Bonferroni-corrected threshold of *p* < 5×10-8 for genome-wide significance and a predetermined “suggestive” threshold of *p* < 1×10-5. Then, using PLINK v1.9, we clumped variants meeting this threshold to create a set of independent variants (defined as LD *r2* < 0.05). For suggestive variants in ordinal regression analyses, we tested the proportional odds assumption using the Brant test in the brant R package [[32](#ref-schlegel2020)]. Finally, we examined single variants previously found to be putatively associated with B-ASC in Chou et al. (2013) to determine if any were validated in NACC at the *p* < 0.05 significance level [[4](#ref-chou2013)].

In the ROSMAP analyses, we first performed single-variant regression analyses with variants identified as suggestively significant in the NACC analyses, using a significance threshold of *p* < 0.05 for these variants. We then performed logistic and ordinal genome-wide single-variant as in the NACC analyses. Covariates included age at death, sex, study (ROS vs. MAP) and the first five PCs. We then performed a mega-analysis on both cohorts, including an indicator variable for ROSMAP participants. We also performed both fixed- and random-effects meta-analyses in PLINK v1.9 for ADGC and ROSMAP logistic GWAS results.

To seek replication of our results, we analyzed variants with NACC + ROSMAP mega-analytic *p* < 1x10-6 separately in the ADNI and ACT data sets. We then added the ADNI and ACT participants to our combined NACC and ROSMAP data set to perform a combined mega-analyss. A total of 4470488 variants were shared between the NACC, ROSMAP, ADNI, and ACT data sets that passed QC measures.

### Gene-based analyses

Following single-variant analyses, we performed gene-based analyses using MAGMA [[33](#ref-leeuw2015)]. First, we annoted variants within 1000 kilo-basepairs (kb) of genes using the Genome Reference Consortium Human Build 37 (GRCh37/hg19) positions available on the MAGMA website (<https://ctg.cncr.nl/software/magma>), producing a total of 18473 genes with at least one annotated variant. We then performed competitive gene-based analyses in MAGMA.

### Colocalization analyses

To investigate potential functional mechanisms of variants driving GWAS signals for B-ASC, we performed colocalization analyses for suggestive variants using the coloc R package and QTL summary statistic data from The Genotype-Tissue Expression (GTEx) project V8 publicly available data (accessible at <https://www.gtexportal.org/home/datasets>) [[14](#ref-thegeno2013), [34](#ref-aguet2017)–[36](#ref-gtexpor)]. QTL are calculated in GTEx by performing single-variant analysis of gene expression for all variants within 1000 kb of the transcription start or end site for each gene. First, we systematically checked to determine if suggestive variants from our Stage 3 mega-analysis (*p* < 1×10-5) were significant expression QTL (eQTL) or splicing QTL (sQTL) across 52 tissues using GTEx summary statistics of participants of European descent. We then performed colocalization analysis on each significant phenotype/tissue combination using default prior probabilities in the coloc package at P1 = P2 = 1x10-4 and P12 = 1x10-5. These parameter choices indicate that each variant used has a 1/10,000 prior probability of being a B-ASC risk variant or a QTL, and a variant that is either a B-ASC risk variant or a QTL has a 1/10 prior probability of being both [[35](#ref-giambartolomei2014)]. Because the coloc package can currently only analyze dichotomous or continuous variables, we did not analyze variants from the ordinal regression analyses. A posterior probability of colocalization (PPH4) of 50% or greater was chosen to indicate evidence for colocalization.

### Mediation analyses

To test the hypotheses that variants associated with B-ASC risk may be mediated by HTN or DM, we performed mediation analyses using R on the subset of NACC participants with clinical variables available [[1](#ref-ighodaro2017), [37](#ref-baron)]. Participants were labeled as DM or HTN cases if they had at least one of 1) self-reported diagnosis, 2) clinician-reported diagnoses, or 3) reported use of DM or HTN medication on their most recent clinical visit prior to death.

### Sensitivity analyses

Given previously identified potential differences in clinical risk factors for B-ASC in participants stratified by age at death, we re-analyzed our analyses in NACC and ROSMAP using only participants with an age at death of eighty or above. To further assess the robustness of our results under different model assumptions, we performed several sensitivity analyses in our NACC data set using the binary B-ASC outcome variable. We first performed single-variant analyses on each ADGC cohort and then meta-analyzed in PLINK v1.9 rather than using fixed-effect cohort indicators in our regression models. For variants meeting our suggestive threshold, we included related participants and performed mixed-effects analyses with a random effect incorporating the kinship matrix estimated from KING in R using the GMMAT, GENESIS, and SNPRelate packages [[38](#ref-chen2020)–[41](#ref-zheng2012)]. To overcome issues with computing PCs with samples with related participants, we used the PC-AiR method [[40](#ref-gogarten)].

# Results

Of the 3382 NACC participants that met inclusion criteria for analysis, 935 (27.6%) had no B-ASC, 1023 (30.2%) had mild B-ASC, 1043 (30.8%) had moderate B-ASC, and 381 (11.3%) had severe B-ASC (see **Table 1**). ROSMAP participants that met inclusion for analysis had comparatively less B-ASC pathology (*p* = 0.002): 414 (34.9%) had no B-ASC, 405 (34.1%) had mild B-ASC, 284 (23.9%) had moderate B-ASC, and 84 (7.1%) had severe B-ASC. NACC participants were also significantly more likely to be male (50% vs.33%, *p* < 0.001) and had younger ages at death on average (mean age at death 82 vs. 89.6, *p* < 0.001) compared to ROSMAP participants. ADNI had the highest proportion of male participants of any cohort at 83 %, while ACT had the lowest percentage at participant with no B-ASC with 7 (1.4%) participants.

| Variable | Labels | NACC | ROSMAP | ADNI | ACT |
| --- | --- | --- | --- | --- | --- |
| B-ASC |  |  |  |  |  |
|  | None | 935 (27.6%) | 414 (34.9%) | 1 (2.1%) | 7 (1.4%) |
|  | Mild | 1023 (30.2%) | 405 (34.1%) | 31 (66%) | 128 (25%) |
|  | Moderate | 1043 (30.8%) | 284 (23.9%) | 13 (27.7%) | 257 (50.2%) |
|  | Severe | 381 (11.3%) | 84 (7.1%) | 2 (4.3%) | 120 (23.4%) |
| Sex |  |  |  |  |  |
|  | Female | 1680 (49.7%) | 800 (67.4%) | 8 (17%) | 286 (55.9%) |
|  | Male | 1702 (50.3%) | 387 (32.6%) | 39 (83%) | 226 (44.1%) |
| Age of Death |  |  |  |  |  |
|  | Mean (SD) | 82 (9.4) | 89.6 (6.5) | 83.3 (7.1) | 89 (6.5) |
|  | Median [Min, Max] | 83 [47, 111] | 90 [66, 108.3] | 84 [59, 97] | 90 [70, 106] |

## Single-variant analyses

In the Stage 1 analyses, 1424 (42.1%) and 368 (31%) of participants had either moderate or severe B-ASC and were counted as cases in NACC and ROSMAP, respectively. In the Stage 1 NACC GWAS (*n* = 3382), one locus on Chromosome 6q14.1 with sentinal variant rs2603462 was genome-wide significantly associated with B-ASC (odds ratio [OR] = 1.45, *p* = 2.5× 10-8). We identified 13 other loci that met our suggestive association threshold of *p* < 1×10-5, described in **Table 2**. Of the 14 NACC loci that met our suggestive threshold, rs7902929 on Chromosome 10q25.1 was validated at *p* < 0.05 level in ROSMAP (NACC OR = 1.57, *p* = 7.8× 10-6; ROSMAP OR = 1.61, *p* = 0.0069). For the other 13 loci, two variants did not pass QC in ROSMAP, two had non-significant ROSMAP ORs <1.00, and 9 had non-significant effect sizes in the same direction as in NACC. Results in the ordinal regressions were broadly similar to the logistic analyses in both cohorts. Of the top 14 loci in the NACC primary logistic regression, five also met our suggestive threshold in the ordinal regression, and ten had *p*-values < 1x10-4. Of the six variants found to be nominally associated with B-ASC by Chou et al. (2013), no variants were validated at the *p* < 0.05 significance level in the NACC cohort. In the secondary analysis using only participants with an age at death of 80 years or older, no variants achieved genome-wide significance (*p* < 5x10-8). In the NACC analysis, we identified 12 independent loci that met the suggestive threshold. No suggestive variants from the NACC analyses were validated in the ROSMAP analyses at the *p* < 0.05 significance threshold (Supplementary Tables S2 and S3).

| Table 1: Cohorts Used in Analyses | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| SNP | CHR | BP | Gene | A1/A2 | NACC OR [95% CI] | NACC P | ROSMAP OR [95% CI] | ROSMAP P |
| rs2603462 | 6 | 81,418,667 | TBD | A/C | 1.45 [1.27-1.66] | 2.5e-08 | 1.05 [0.84-1.33] | 0.65 |
| rs3774902 | 4 | 23,890,782 | TBD | G/A | 1.77 [1.42-2.21] | 4.5e-07 | - | NA |
| rs2418491 | 2 | 59,535,434 | TBD | G/A | 1.33 [1.19-1.49] | 4.8e-07 | 1.02 [0.84-1.23] | 0.84 |
| rs115980554 | 6 | 34,456,632 | TBD | C/T | 1.63 [1.33-2] | 3.4e-06 | 0.95 [0.65-1.37] | 0.77 |
| rs61944465 | 13 | 27,272,704 | TBD | G/A | 1.33 [1.18-1.5] | 3.6e-06 | 1.16 [0.93-1.46] | 0.19 |
| rs4470773 | 5 | 177,507,080 | TBD | C/G | 1.28 [1.15-1.42] | 3.9e-06 | 1.07 [0.89-1.29] | 0.47 |
| rs6574718 | 14 | 26,395,832 | TBD | C/T | 1.28 [1.15-1.42] | 5.3e-06 | - | NA |
| rs387083 | 19 | 35,847,115 | TBD | G/A | 1.28 [1.15-1.42] | 5.4e-06 | 1.02 [0.85-1.22] | 0.86 |
| rs28362345 | 6 | 31,165,836 | TBD | T/C | 1.27 [1.15-1.41] | 6.3e-06 | 0.94 [0.78-1.14] | 0.53 |
| rs2069126 | 20 | 19,749,957 | TBD | G/A | 1.27 [1.14-1.4] | 6.7e-06 | 1.11 [0.92-1.34] | 0.27 |
| rs7902929 | 10 | 107,237,532 | TBD | T/C | 1.57 [1.29-1.92] | 7.8e-06 | 1.61 [1.14-2.26] | 0.0069 |
| rs12700439 | 7 | 23,472,061 | TBD | T/C | 1.26 [1.14-1.4] | 8.0e-06 | 1.07 [0.89-1.29] | 0.47 |
| rs10790707 | 11 | 124,460,485 | TBD | G/A | 1.26 [1.14-1.4] | 8.1e-06 | 1.03 [0.85-1.24] | 0.76 |
| rs9895518 | 17 | 49,645,803 | TBD | A/G | 1.28 [1.15-1.42] | 8.3e-06 | 1.09 [0.9-1.32] | 0.39 |

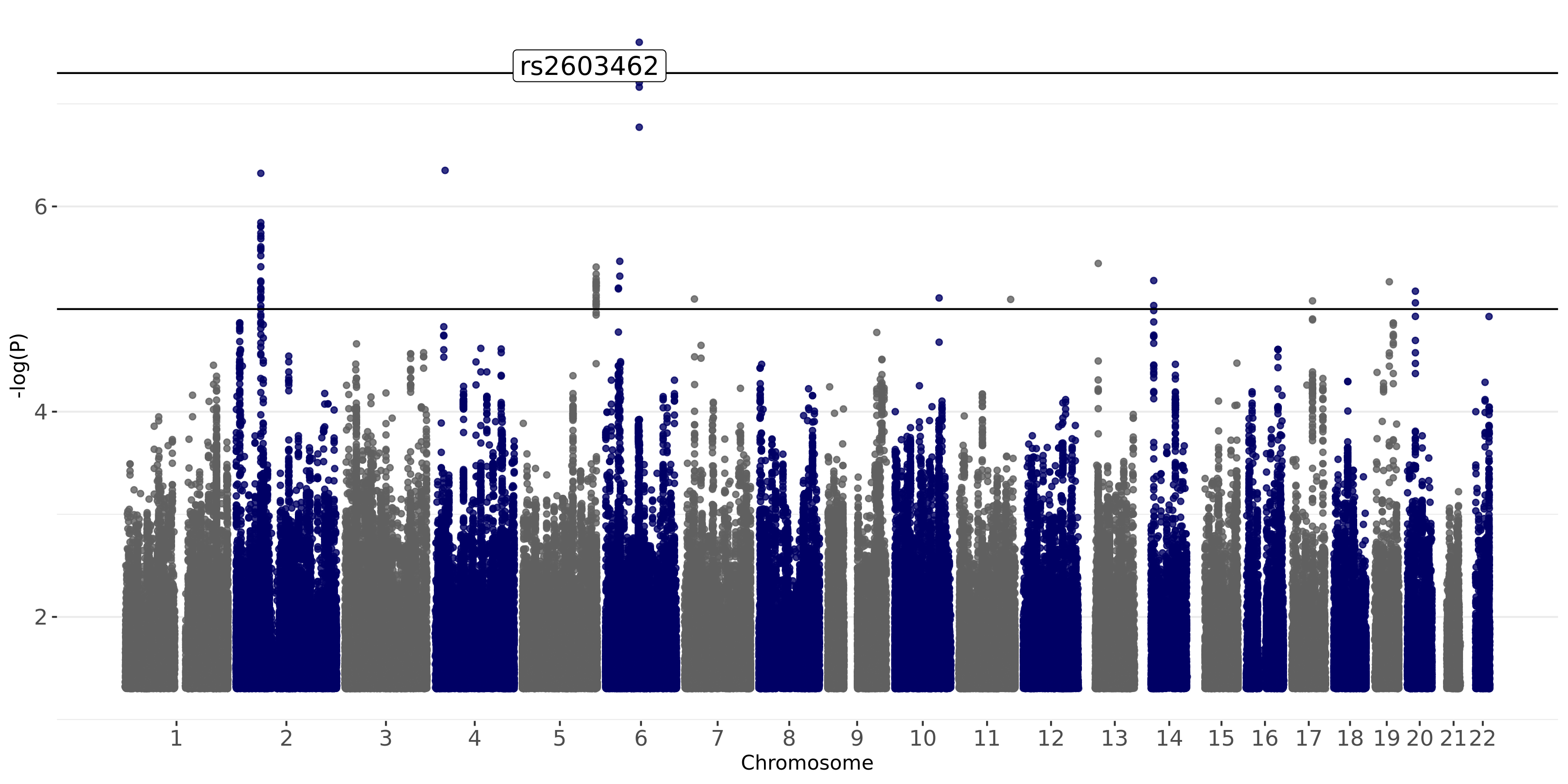


Figure 2: Manhattan Plot of NACC GWAS

In the Stage 2 NACC and ROSMAP mega-analysis (*n* = 4569), 11 loci met the suggestive significance threshold, with rs7902929 (OR = 1.58, *p* = 1.8× 10-7) and rs2603462 (OR = 1.34, *p* = 4× 10-7) producing the smallest *p*-values. Because the ADNI (*n* = 47) and ACT (*n* = 512) sample sizes were much smaller than those NACC and ROSMAP, we limited the replication attempts in these cohorts to variants meeting a more stringent threshold of *p* < 1x10-6, which two variants, rs2603462 and rs7902929, met. In the ADNI neuropathology cohort, effects of both rs2603462 (OR [95% confidence interval (CI)] =4.75 [1.42-15.91], P =0.012) and rs7902929 (OR [95% CI] =26.49 [2.07-338.8], P =0.012 ) were replicated at the *P* < 0.05 significance level (**Table 3**). In the ACT neuropathology cohort, neither variant was replicated, and both variants had non-signifant effects in the opposite direction than in the other cohorts.

| Table 3: Stage 2 Mega-Analysis and Replication Results | | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | Stage 2 Mega-Analysis | | ADNI | | ACT | |
| SNP | CHR | BP | Gene | OR [95% CI] | P | OR [95% CI] | P | OR [95% CI] | P |
| rs61776730 | 1 | 71,057,739 | TBD | 1.46 [1.24-1.72] | 7.3e-06 | - | - | - | - |
| rs2418491 | 2 | 59,535,434 | TBD | 1.24 [1.13-1.37] | 7.6e-06 | - | - | - | - |
| rs13082422 | 3 | 22,338,009 | TBD | 1.22 [1.12-1.34] | 8.1e-06 | - | - | - | - |
| rs3774902 | 4 | 23,890,782 | TBD | 1.55 [1.27-1.88] | 9.7e-06 | - | - | - | - |
| rs7675509 | 4 | 161,518,462 | TBD | 1.24 [1.13-1.36] | 7.2e-06 | - | - | - | - |
| rs10050232 | 4 | 168,333,572 | TBD | 1.25 [1.14-1.38] | 7.0e-06 | - | - | - | - |
| rs6898408 | 5 | 58,406,648 | TBD | 1.23 [1.12-1.35] | 9.3e-06 | - | - | - | - |
| rs4370294 | 5 | 177,510,515 | TBD | 1.24 [1.13-1.35] | 5.3e-06 | - | - | - | - |
| rs2603462 | 6 | 81,418,667 | ELOVL4 | 1.34 [1.2-1.5] | 4.0e-07 | 4.75 [1.42-15.91] | 0.012 | 0.77 [0.52-1.13] | 0.18 |
| rs7902929 | 10 | 107,237,532 | SORCS3 | 1.58 [1.33-1.87] | 1.8e-07 | 26.49 [2.07-338.8] | 0.012 | 0.81 [0.47-1.4] | 0.45 |
| rs61944465 | 13 | 27,272,704 | TBD | 1.29 [1.16-1.44] | 2.3e-06 | - | - | - | - |

In the Stage 3 mega-analysis combining the NACC, ROSMAP, ADNI, and ACT cohorts, no variants met the genome-wide significance threshold. A total of nine loci met the threshold for suggestive significance (**Table 4**), of which five were identified in Stage 2, while four were novel loci. Both rs2603462 and rs7902929 met the suggestive threshold (**Figure 3**).

| Table 4: Stage 3 Mega-Analysis Results | | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| SNP | CHR | BP | Gene | A1/A2 | OR [95% CI] | P |
| rs61776730 | 1 | 71,057,739 | TBD | A/G | 1.44 [1.23-1.69] | 4.6e-06 |
| rs4370294 | 5 | 177,510,515 | TBD | G/A | 1.23 [1.13-1.34] | 3.3e-06 |
| rs2603462 | 6 | 81,418,667 | TBD | C/A | 1.3 [1.17-1.45] | 1.9e-06 |
| rs10738370 | 9 | 14,582,342 | TBD | T/C | 1.25 [1.13-1.38] | 8.9e-06 |
| rs7902929 | 10 | 107,237,532 | TBD | C/T | 1.51 [1.28-1.78] | 7.1e-07 |
| rs76828179 | 12 | 32,594,268 | TBD | A/G | 1.36 [1.19-1.56] | 5.2e-06 |
| rs61944465 | 13 | 27,272,704 | TBD | G/A | 1.29 [1.16-1.42] | 1.1e-06 |
| rs415468 | 19 | 44,456,672 | TBD | C/G | 1.22 [1.11-1.32] | 9.3e-06 |
| rs16981092 | 20 | 19,745,011 | TBD | G/A | 1.22 [1.11-1.32] | 8.5e-06 |

![](data:application/pdf;base64,)

![Figure 3: Regional plots of rs2603462 (top) and rs7902929 (bottom)](data:application/pdf;base64,)

Figure 3: Regional plots of rs2603462 (top) and rs7902929 (bottom)

## Gene-based analyses

In the gene-based analyses, no genes achieved Bonferroni-adjusted significance (*p* < 2.5x10-6). *SORCS1* (*p* = 5.5e-05) achieved the smallest *p*-value of any gene, while the adjacent *SORCS3* gene *p*-value was somewhat larger (*p* = 8.3e-03).

## Colocalization analyses

In the mega-analyses, five suggestive variants were significant QTLs across nine phenotypes and 12 tissues in GTEx, leading us to perform colocalization analyses for a total of 18 phenotype-tissue combinations. Of these, nine met our evidential threshold of PPH4 > 50% for colocalization. The strongest evidence of colocalization was found in *FFAR3* expression in visceral adipose tissue with PPH4 = 98.9%, followed by expression of the long intergenic non-coding RNA (lincRNA) AC058791.1 in the cerebellum and hypothalamus with PPH4 of 94.6% and 91.6%, respectively. *ELOVL4* expression in the cerebellar hemisphere colocalizes with B-ASC risk with a PPH4 of 92.9%. *WASF3* expression colocalizes with B-ASC risk across three brain regions, including the basal ganglia with PPH4 = 86.5%, and preferential splicing of *WASF3* in fibroblasts colocalizes with 64.9% (**Table 5**).

## Mediation analyses

Finally, in the subset of our NACC participants who had clinical data for DM (N = 1725) and HTN (N = 1726) status, none of the suggestive variants were associated with diabetes or hypertension status at the *p* < 0.05 level, and all variant effect sizes on B-ASC in regressions including diabetes or hypertension were within the 95% confidence intervals in the base model excluding them.

## Sensitivity analyses

Performing fixed-effects meta-analysis on the NACC ADGC genotyping cohorts produced nearly identical effect sizes and *p*-values compared to including indicator variables for the non-reference cohorts in our primary analysis. Similarly, including the participants originally excluded due to relatedness did not produce notably different effect sizes or *p*-values for any of the suggestive variants.

# Discussion

We performed the first GWAS of autopsy-proven B-ASC using neuropathology and genotype data from four neuropathology cohorts derived from multiple research centers. We found a significant association between one locus on Chromosome 6 and B-ASC in the Stage 1 NACC GWAS (*p* = 2.5× 10-8). In the Stage 3 mega-analysis of all four cohorts, this locus colocalizes with *ELOVL4* gene expression in GTEx (PPH4 = 92.9%), providing evidence that this locus may affect B-ASC risk through mediating *ELOVL4* expression (see **Figure X**). Another locus on Chromosome 10 near *SORCS3* suggestively associated with B-ASC was validated in the ROSMAP cohort. These variants’ effects were successfully replicated in the ADNI cohort, but failed to replicate in the ACT cohort. We also found suggestive evidence for association between other loci and B-ASC risk in both cohorts and that some of these loci colocalize with gene eQTL and sQTL in GTEx.

*ELOVL4* codes for the elongation of very long chain fatty acids-4 protein, an elongase that synthesizes very long chain saturated and unsaturated fatty acids. Much of the research on *ELOVL4* has focused on its association with mendelian diseases affecting the visual and nervous systems, such as Stargardt-like macular dystrophy and spinocerebellar ataxia [[42](#ref-hopiavuori2019)]. In GTEx, *ELOVL4* is more highly expressed in the brain (median transcripts per kilobase million [TPM] 5.5-43.9) relative to most other tissues. We searched the GWAS Catalog and found two recent studies that found genome-wide significant associations between variants mapped to *ELOVL4* and multiple body weight-related phenotypes, including waist circumference adjusted for body mass index (BMI) (rs76567515, (*p* = 2x10-12), waist-to-hip ratio (rs76567515, *p* = 3x10-10), and waist-to-hip ratio adjusted for BMI (rs1849275, *p* = 1x10-9) [[10](#ref-buniello2019), [43](#ref-zhu2020), [44](#ref-kichaev2019)]. These variants are significant eQTLs for *ELOVL4* in GTEx, but are each located >200 kb from the locus identified in our study and are not in LD with the lead variant (*r2* < 0.05). Additionally, another suggestive variant in NACC, rs387083, is located between the free fatty acid receptor genes *FFAR1* and *FFAR3* on Chromosome nineteen. This locus acts as an eQTL of *FFAR3* in the visceral omentum adipose. These results suggest that fatty acid metabolism and signaling may play a role in B-ASC risk, and that *ELOVL4* and *FFAR3* may potentially be related to B-ASC risk through effects on BMI-related phenotypes.

The variant suggestively associated with B-ASC in the NACC primary analyses subsequently validated in ROSMAP, rs7902929, is located approximately 212 kb from the 3’ end of the gene *SORCS3*. While the genes on which intergenic variants exert their effects are not necessarily the closest genes, *SORCS3* is the only protein-coding gene within a 1 megabase window from rs7902929, increasing the likelihood that the locus is functionally tied to it. *SORCS3* codes for the sortilin-related VPS10 domain-containing receptor 3, a vacuolar protein expressed in the brain [[45](#ref-reitz2013), [46](#ref-wang2020)]. Previous studies using candidate gene designs have provided tentative evidence that genetic variation in *SORCS3* may be associated with AD [[45](#ref-reitz2013), [46](#ref-wang2020)]. In the GWAS Catalog, *SORCS3*-mapped variants are significantly associated with multiple phenotypes, including depressive symptoms (rs1021363, *p* = 1x10-13), self-reported educational attainment (rs11599236, *p* = 1x10-13), and systolic blood pressure (rs191784289, P = 3x10-13) [[44](#ref-kichaev2019), [47](#ref-baselmans2019), [48](#ref-lee2018)].

Additionally, a member of the VPS10 family, *SORCS1*, acheived the smallest *p*-value in the gene-based analysis. We performed a *post-hoc* analysis of this region and identified two independent loci with lead *p*-values just above our suggestive threshold driving this effect: one locus in LD with rs7902929 and one not (*p* = 1.2x10-5). We examined the lead variant in the second locus and identified a consistent effect across all four cohorts that was nominally significant in NACC (*p* = 1x10-4) and ACT (*p* = 0.001), though not in ROSMAP (*p* = 0.25) or ADNI (*p* = 0.07). Previous studies have identified *SORCS1* as a potential risk variant for AD [[45](#ref-reitz2013), [46](#ref-wang2020)]. Therefore, it is possible that the VPS10 family influences B-ASC risk through AD risk, or vice versa, though further work should be done to clarify the possible role of the VPS10 family on both phenotypes.

None of the effects of top loci in the NACC GWAS were mediated by HTN or DM status in NACC. This finding fails to help clarify the somewhat inconsistent evidence of association between these clinical risk factors and B-ASC pathology. However, this analysis was limited by a substantially smaller sample size (n = 1726) than our other analysis and the definitions of the HTN and DM used, as the diagnostic variables in NACC used consisted of a mixture of patient self-reports, physician reporting, and use of HTN or DM medications. Future studies employing causal inference methods such as two-sample Mendelian randomization may be able to provide clearer evidence for or against the mediation of B-ASC genetic risk by clinical risk factors.

The most obvious limitation of our study is the relatively low sample size for a GWAS. NACC, the largest cohort used, had only 3382 participants available for GWAS, and the total sample size of the Stage 3 mega-analysis with all four cohorts was 5128. In contrast, a recent GWAS meta-analysis of AD exceed 90,000 participants [[49](#ref-kunkle2019)], and GWAS of other phenotypes have exceeded 1 million participants [[48](#ref-lee2018)]. Nevertheless, smaller GWAS of other neuropathologic endophenotypes have identified risk loci [[50](#ref-katsumata2017)], demonstrating that smaller sample sizes do not necessarily prevent discovery. Despite sample size limitations, one suggestive locus on Chromosome 10 near *SORCS3* identified in NACC was validated in ROSMAP, and both this locus and a locus near *ELOVL4* were replicated in ADNI. Furthermore, 11 of 13 B-ASC-associated loci in NACC had affect sizes in the same direction in ROSMAP, which suggests that these loci may be associated with B-ASC but suffer from the “winner’s curse” in the NACC analyses. Furthermore, the use of downstream functional analyses can provide additional evidence for risk loci that do not meet the conservative Bonferroni-corrected significance thresholds.

Most genetic loci identified as being associated with B-ASC in NACC were not validated in ROSMAP. This could be potentially explained via the significantly different demographic attributes and study designs between the two cohorts, as ROSMAP participants were simultaneously older at death and had lower risk of B-ASC pathology (**Table 1**), indicating that there may be unaccounted-for genetic confounding and differential selection bias between cohorts. For instance, ROS recruits from Catholic sisters and brothers, who are on average more highly educated than the general population. In contrast, the NACC participants are recruited primarily from clinic-based populations across the ADRCs. Differences in study design are also seen in the ADNI cohort, whose recruitment practices primarily follow those of ADRCs, and in ACT, whose participants are recruited from the aged dementia-free community in the Seattle, WA area.

Heterogeneity in neuropathological grading of B-ASC in each cohort may also contribute to impairing study power and limiting reproducibility between studies. In one study investigating the inter-rater reliability of cerebrovascular neuropathological endophenotypes, B-ASC had the second-lowest reliability (after microinfarcts), with reliability below 0.8 in five of twelve brain regions graded [[51](#ref-skrobot2016)]. This relatively low inter-rater reliability indicates there may be substantial measurement variance in B-ASC between Neuropathologists both within and between study centers. Furthermore, B-ASC is likely graded in different brain regions in different centers and studies. B-ASC is graded globally in NACC, with no particular brain region specified, but is graded in the basal ganglia in ROSMAP [[4](#ref-chou2013)]. Documentation on the brain regions graded for B-ASC in ADNI and ACT is limited.

In closing, we used autopsy-derived endophenotypes linked to genetic data in multiple cohorts to identify genetic loci associated with B-ASC pathology. We found one locus significantly associated with B-ASC that colocalizes with ELOVL4 gene expression in GTEx. This study provides first GWAS of autopsy-verified B-ASC pathology.

# Acknowledgements

This work was supported by the University of Kentucky Center for Clinical and Translational Science TL-1 Fellowship (TLITR001997), the National Center for Advancing Translational Sciences, UL1TR001998, and the Dean of the College of Medicine, University of Kentucky. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH or the University of Kentucky.

Genotyping was performed by Alzheimer’s Disease Genetics Consortium (ADGC), U01 AG032984, RC2AG036528.

We thank the study participants and staff of the Rush Alzheimer’s Disease Center. The ROSMAP studies are supported by NIA grants **xxxx**.

Data collection and sharing for this project was funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer’s Association; Alzheimer’s Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health ([www.fnih.org](https://fnih.org/)). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer’s Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

The NACC database is funded by NIA/NIH Grant U01 AG016976. NACC data are contributed by the NIA-funded ADRCs: P30 AG019610 (PI Eric Reiman, MD), P30 AG013846 (PI Neil Kowall, MD), P50 AG008702 (PI Scott Small, MD), P50 AG025688 (PI Allan Levey, MD, PhD), P50 AG047266 (PI Todd Golde, MD, PhD), P30 AG010133 (PI Andrew Saykin, PsyD), P50 AG005146 (PI Marilyn Albert, PhD), P50 AG005134 (PI Bradley Hyman, MD, PhD), P50 AG016574 (PI Ronald Petersen, MD, PhD), P50 AG005138 (PI Mary Sano, PhD), P30 AG008051 (PI Thomas Wisniewski, MD), P30 AG013854 (PI Robert Vassar, PhD), P30 AG008017 (PI Jeffrey Kaye, MD), P30 AG010161 (PI David Bennett, MD), P50 AG047366 (PI Victor Henderson, MD, MS), P30 AG010129 (PI Charles DeCarli, MD), P50 AG016573 (PI Frank LaFerla, PhD), P50 AG005131 (PI James Brewer, MD, PhD), P50 AG023501 (PI Bruce Miller, MD), P30 AG035982 (PI Russell Swerdlow, MD), P30 AG028383 (PI Linda Van Eldik, PhD), P30 AG053760 (PI Henry Paulson, MD, PhD), P30 AG010124 (PI John Trojanowski, MD, PhD), P50 AG005133 (PI Oscar Lopez, MD), P50 AG005142 (PI Helena Chui, MD), P30 AG012300 (PI Roger Rosenberg, MD), P30 AG049638 (PI Suzanne Craft, PhD), P50 AG005136 (PI Thomas Grabowski, MD), P50 AG033514 (PI Sanjay Asthana, MD, FRCP), P50 AG005681 (PI John Morris, MD), P50 AG047270 (PI Stephen Strittmatter, MD, PhD).

# References

1. Ighodaro ET, Abner EL, Fardo DW, et al (2017) Risk factors and global cognitive status related to brain arteriolosclerosis in elderly individuals. Journal of Cerebral Blood Flow & Metabolism 37:201–216. <https://doi.org/10.1177/0271678X15621574>

2. Grinberg LT, Thal DR (2010) Vascular pathology in the aged human brain. Acta Neuropathologica 119:277–290. <https://doi.org/10.1007/s00401-010-0652-7>

3. Buchman AS, Yu L, Boyle PA, et al (2013) Microvascular brain pathology and late-life motor impairment. Neurology 80:712–718. <https://doi.org/10.1212/WNL.0b013e3182825116>

4. Chou S-Y, Shulman JM, Keenan BT, et al (2013) Genetic susceptibility for ischemic infarction and arteriolosclerosis based on neuropathologic evaluations. Cerebrovascular Diseases (Basel, Switzerland) 36:181–188. <https://doi.org/10.1159/000352054>

5. Arvanitakis Z, Capuano AW, Leurgans SE, et al (2016) Relation of cerebral vessel disease to alzheimer’s disease dementia and cognitive function in older persons: A cross-sectional study. The Lancet Neurology 15:934–943. <https://doi.org/10.1016/S1474-4422(16)30029-1>

6. Neltner JH, Abner EL, Baker S, et al (2014) Arteriolosclerosis that affects multiple brain regions is linked to hippocampal sclerosis of ageing. Brain 137:255–267. <https://doi.org/10.1093/brain/awt318>

7. Wu J, Chen X, Xie Y, et al (2005) Characteristics and risk factors of intrarenal arterial lesions in patients with IgA nephropathy. Nephrology Dialysis Transplantation 20:719–727. <https://doi.org/10.1093/ndt/gfh716>

8. Ikee R, Honda K, Ishioka K, et al (2010) Postprandial hyperglycemia and hyperinsulinemia associated with renal arterio-arteriolosclerosis in chronic kidney disease. Hypertension Research: Official Journal of the Japanese Society of Hypertension 33:499–504. <https://doi.org/10.1038/hr.2010.22>

9. Cameron JS (2006) The discovery of diabetic nephropathy: from small print to centre stage. Journal of Nephrology 19 Suppl 10:S75–87

10. Buniello A, MacArthur JAL, Cerezo M, et al (2019) The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Research 47:D1005–D1012. <https://doi.org/10.1093/nar/gky1120>

11. Traylor M, Zhang CR, Adib-Samii P, et al (2016) Genome-wide meta-analysis of cerebral white matter hyperintensities in patients with stroke. Neurology 86:146–153. <https://doi.org/10.1212/WNL.0000000000002263>

12. Malik R, Chauhan G, Traylor M, et al (2018) Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nature Genetics 50:524–537. <https://doi.org/10.1038/s41588-018-0058-3>

13. Chauhan G, Adams HHH, Satizabal CL, et al (2019) Genetic and lifestyle risk factors for MRI-defined brain infarcts in a population-based setting. Neurology 92:e486–e503. <https://doi.org/10.1212/WNL.0000000000006851>

14. (2013) The genotype-tissue expression (GTEx) project. Nature genetics 45:580–585. <https://doi.org/10.1038/ng.2653>

15. Welcome to alzheimer’s disease genetics consortium | ADGC

16. Besser LM, Kukull WA, Teylan MA, et al (2018) The Revised National Alzheimer’s Coordinating Center’s Neuropathology Form-Available Data and New Analyses. Journal of Neuropathology and Experimental Neurology 77:717–726. <https://doi.org/10.1093/jnen/nly049>

17. Bennett DA, Buchman AS, Boyle PA, et al (2018) Religious orders study and rush memory and aging project. Journal of Alzheimer’s disease : JAD 64:S161–S189. <https://doi.org/10.3233/JAD-179939>

18. Kukull WA, Higdon R, Bowen JD, et al (2002) Dementia and Alzheimer Disease Incidence: A Prospective Cohort Study. Archives of Neurology 59:1737. <https://doi.org/10.1001/archneur.59.11.1737>

19. Buchman AS, Leurgans SE, Nag S, et al (2011) Cerebrovascular disease pathology and parkinsonian signs in old age. Stroke; a journal of cerebral circulation 42:3183–3189. <https://doi.org/10.1161/STROKEAHA.111.623462>

20. Chang CC, Chow CC, Tellier LC, et al (2015) Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience 4: <https://doi.org/10.1186/s13742-015-0047-8>

21. (2012) An integrated map of genetic variation from 1,092 human genomes. Nature 491:56–65. <https://doi.org/10.1038/nature11632>

22. R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

23. McCarthy S, Das S, Kretzschmar W, et al (2016) A reference panel of 64,976 haplotypes for genotype imputation. Nature Genetics 48:1279–1283. <https://doi.org/10.1038/ng.3643>

24. Dumitrescu L, Mahoney ER, Mukherjee S, et al (2020) Genetic variants and functional pathways associated with resilience to alzheimer’s disease. Brain 143:2561–2575. <https://doi.org/10.1093/brain/awaa209>

25. Das S, Forer L, Schönherr S, et al (2016) Next-generation genotype imputation service and methods. Nature Genetics 48:1284–1287. <https://doi.org/10.1038/ng.3656>

26. Fuchsberger C, Abecasis GR, Hinds DA (2015) minimac2: Faster genotype imputation. Bioinformatics 31:782–784. <https://doi.org/10.1093/bioinformatics/btu704>

27. Taliun D, Harris DN, Kessler MD, et al (2019) Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. bioRxiv 563866. <https://doi.org/10.1101/563866>

28. Chang CC, Purcell S Plink 1.9

29. Purcell S, Neale B, Todd-Brown K, et al (2007) PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. The American Journal of Human Genetics 81:559–575. <https://doi.org/10.1086/519795>

30. Manichaikul A, Mychaleckyj JC, Rich SS, et al (2010) Robust relationship inference in genome-wide association studies. Bioinformatics 26:2867–2873. <https://doi.org/10.1093/bioinformatics/btq559>

31. Marees AT, de Kluiver H, Stringer S, et al (2018) A tutorial on conducting genome-wide association studies: Quality control and statistical analysis. International Journal of Methods in Psychiatric Research 27:e1608. <https://doi.org/10.1002/mpr.1608>

32. Schlegel B, Steenbergen M (2020) Brant: Test for parallel regression assumption

33. Leeuw CA de, Mooij JM, Heskes T, Posthuma D (2015) MAGMA: Generalized Gene-Set Analysis of GWAS Data. PLOS Computational Biology 11:e1004219. <https://doi.org/10.1371/journal.pcbi.1004219>

34. Aguet F, Brown AA, Castel SE, et al (2017) Genetic effects on gene expression across human tissues. Nature 550:204–213. <https://doi.org/10.1038/nature24277>

35. Giambartolomei C, Vukcevic D, Schadt EE, et al (2014) Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. PLOS Genetics 10:e1004383. <https://doi.org/10.1371/journal.pgen.1004383>

36. GTEx portal

37. Baron RM, Kenny DA The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. 10

38. Chen H, Conomos MP (2020) GMMAT: Generalized linear mixed model association tests

39. Conomos MP, Gogarten SM, Brown L, et al (2019) GENESIS: GENetic EStimation and inference in structured samples (GENESIS): Statistical methods for analyzing genetic data from samples with population structure and/or relatedness

40. Gogarten SM, Sofer T, Chen H, et al Genetic association testing using the GENESIS R/Bioconductor package. Bioinformatics. <https://doi.org/10.1093/bioinformatics/btz567>

41. Zheng X, Levine D, Shen J, et al (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics 28:3326–3328. <https://doi.org/10.1093/bioinformatics/bts606>

42. Hopiavuori BR, Anderson RE, Agbaga M-P (2019) ELOVL4: Very long-chain fatty acids serve an eclectic role in mammalian health and function. Progress in Retinal and Eye Research 69:137–158. <https://doi.org/10.1016/j.preteyeres.2018.10.004>

43. Zhu Z, Guo Y, Shi H, et al (2020) Shared genetic and experimental links between obesity-related traits and asthma subtypes in UK Biobank. The Journal of Allergy and Clinical Immunology 145:537–549. <https://doi.org/10.1016/j.jaci.2019.09.035>

44. Kichaev G, Bhatia G, Loh P-R, et al (2019) Leveraging Polygenic Functional Enrichment to Improve GWAS Power. American Journal of Human Genetics 104:65–75. <https://doi.org/10.1016/j.ajhg.2018.11.008>

45. Reitz C, Tosto G, Vardarajan B, et al (2013) Independent and epistatic effects of variants in VPS10-d receptors on alzheimer disease risk and processing of the amyloid precursor protein (APP). Translational Psychiatry 3:e256. <https://doi.org/10.1038/tp.2013.13>

46. Wang M, Wang S, Li Y, et al (2020) Integrated analysis and network pharmacology approaches to explore key genes of xingnaojing for treatment of alzheimer’s disease. Brain and Behavior 10: <https://doi.org/10.1002/brb3.1610>

47. Baselmans BML, Jansen R, Ip HF, et al (2019) Multivariate genome-wide analyses of the well-being spectrum. Nature Genetics 51:445–451. <https://doi.org/10.1038/s41588-018-0320-8>

48. Lee JJ, Wedow R, Okbay A, et al (2018) Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nature Genetics 50:1112–1121. <https://doi.org/10.1038/s41588-018-0147-3>

49. Kunkle BW, Grenier-Boley B, Sims R, et al (2019) Genetic meta-analysis of diagnosed Alzheimer’s disease identifies new risk loci and implicates A?, tau, immunity and lipid processing. Nature Genetics 51:414. <https://doi.org/10.1038/s41588-019-0358-2>

50. Katsumata Y, Nelson PT, Ellingson SR, Fardo DW (2017) Gene-based association study of genes linked to hippocampal sclerosis of aging neuropathology: GRN, TMEM106B, ABCC9, and KCNMB2. Neurobiology of Aging 53:193.e17–193.e25. <https://doi.org/10.1016/j.neurobiolaging.2017.01.003>

51. Skrobot OA, Attems J, Esiri M, et al (2016) Vascular cognitive impairment neuropathology guidelines (VCING): The contribution of cerebrovascular pathology to cognitive impairment. Brain 139:2957–2969. <https://doi.org/10.1093/brain/aww214>

# Supplementary Material