Multi-Omics of the Iron Homeostasis Pathway in Patient Outcomes after Aneurysmal Subarachnoid Hemorrhage

by

Lacey W. Heinsberg

Bachelor of Science in Civil Engineering, Southern Illinois University Carbondale, 2009

Bachelor of Science in Nursing, Goldfarb School of Nursing at Barnes-Jewish College, 2012

Submitted to the Graduate Faculty of the

School of Nursing in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

SCHOOL OF NURSING

This dissertation was presented

by

Lacey W. Heinsberg

It was defended on

May 27, 2020

and approved by

Sheila A. Alexander, PhD, RN, FCCM, Associate Professor, School of Nursing

Elizabeth A. Crago, PhD, RN, Assistant Professor, School of Nursing

Ryan L. Minster, PhD, MSIS, Assistant Professor, Graduate School of Public Health

Samuel M. Poloyac, PharmD, PhD, Professor, School of Pharmacy

Daniel E. Weeks, PhD, Professor, Graduate School of Public Health

Dissertation Advisor: Yvette P. Conley, PhD, FAAN, Professor, School of Nursing

Multi-Omics of the Iron Homeostasis Pathway in Patient Outcomes after Aneurysmal Subarachnoid Hemorrhage

Lacey W. Heinsberg, PhD, RN

University of Pittsburgh, 2020

Patient outcomes after aneurysmal subarachnoid hemorrhage (aSAH) are variable and healthcare providers are often unable to predict those who will do poorly and are in need of more intensive nursing management. This ancillary, longitudinal, observational, candidate gene association study investigated the hypothesis that epigenetic and genetic variability of genes in the iron homeostasis pathway may account for a proportion of variability in patient outcomes post-aSAH. This study was conducted using a two-tier design (targeted discovery and replication) and capitalized on an existing cohort of aSAH participants with genome-wide DNA methylation, genotype, and patient outcome data as well as stored biosamples. This study used group-based trajectory analysis (Aim 1), binary logistic regression (Aims 1 and 2), and Bayesian statistical methods (Aim 2) to assess the relationship between inferred DNA methylation trajectory groups (Aim 1) and tagging SNPs (Aim 2) from iron homeostasis candidate genes (n = 39) and acute (cerebral vasospasm [CV] and delayed cerebral ischemia [DCI]) and long-term (Glasgow Outcome Scale [GOS] and death at 3 and 12-months) patient outcomes post-aSAH. In Aim 1, we identified three DNA methylation sites in three candidate genes (Amyloid Precursor Protein [*APP*], STEAP3 metalloreductase [*STEAP3*], and Tumor Necrosis Factor [*TNF*]) with inferred trajectory groups significantly associated with patient outcomes in our discovery sample. Specifically, associations were identified between cg08866780 (*APP*) and GOS at 12 months ($p = 0.0001$) and death at 3 and 12 months (both $p=0.0002$); cg25713625 (*STEAP3*) and GOS at 3 and 12 months ($p = 0.00005$ and $p = 0.0005$,

respectively) and death at 3 and 12 months ($p = 0.0013$ and $p = 0.0015$, respectively); and cg08553327 (*TNF*) with GOS at 3 months ($p = 0.00001$). Replication data are pending. For Aim 2, we identified two SNPs in two candidate genes (Ceruloplasmin [*CP*] and Cubilin [*CUBN*]) that were associated with patient outcomes in both the discovery and replication samples. Specifically, rs10904850 (*CUBN*) was significantly associated with DCI (mega-analysis combining targeted discovery and replication samples, $p = 0.004$) and rs17838831 (*CP*) with GOS at 3 and 12 months (mega-analysis $p = 0.0004$ and $p = 0.057$, respectively).

Table of Contents

List of Tables

List of Figures

Preface

To begin, I would like to sincerely thank my dissertation chair, Dr. Yvette Conley, dissertation committee members, Drs. Daniel Weeks, Sheila Alexander, Elizabeth Crago, Ryan Minster, and Samuel Poloyac, and School of Nursing laboratory manager, Sandra Deslouches. There are no words to communicate the ways in which the time, patience, and mentorship you have so generously given me have contributed to my academic and personal growth. I am forever grateful. I would also like to acknowledge the financial support that made this dissertation study and completion of my doctoral studies possible, including funds from the National Institute of Nursing Research of the National Institutes of Health (F31NR017311, T32NR009759, R01NR004339, and R01NR013610), Nightingale Awards of Pennsylvania, Center for Jonas and Veterans Healthcare, the University of Pittsburgh School of Nursing Jayne F Wiggins Memorial Award and Leslie A. Hoffman Endowed Research Award, Sigma Theta Tau (Eta Chapter), the International Society of Nurses in Genetics, and the American Nurses Foundation (including support from the Eastern Nursing Research Society and Eleanor C. Lambertson Scholar Fund). I am grateful that you chose to invest in my research. Additionally, to the participants of this study and their families, I thank you for all you have given, and I hope to someday pay you back by improving outcomes for people who experience subarachnoid hemorrhage. Finally, I would like to thank my husband, Kurt, and son, Henrik, for bringing laughter, joy, and lightness to my world. You enrich my life beyond words.

1.0 Dissertation Proposal

Section 1.0 of this document summarizes the approved dissertation proposal finalized at the comprehensive examination and overview.

1.1 Specific Aims

Aneurysmal subarachnoid hemorrhage (aSAH) is a devastating injury and substantial public health problem.[1] While accounting for only a small percentage of strokes, aSAH is one of the leading contributors to loss of productive life years in America.[2] Among survivors, recovery is variable and the cumulative effect of aSAH is astounding with over 50% of survivors having long-term physical and neuropsychological outcome deficits.[3,4] Unfortunately, knowledge of modifiable factors associated with poor outcomes remains elusive. Evidence suggests that secondary complications occurring during the acute phase window following aSAH are important moderators of poor outcomes, however, the pathophysiology of these complications is largely unclear.[5] Preclinical studies have identified metabolic mechanisms of iron homeostasis as significant in that they contribute to secondary brain injury (similar to complications observed in humans such as cerebral vasospasm [CV] and delayed cerebral ischemia [DCI]),[6–9] representing a plausible biological pathway for targeted therapeutic intervention. We posit that genetic and epigenetic variability in the iron homeostasis pathway are important players in the vascular pathology that leads to poor outcomes following aSAH.

The long-term goal of this research is to identify biologic targets for therapeutic intervention in order to prevent or mitigate secondary complications following aSAH and ultimately improve outcomes within this population. To lay the groundwork for this research trajectory, the overarching purpose of the proposed study is to characterize the relationship between epigenetic (DNA methylation) and genetic (polymorphisms) variability in the iron homeostasis pathway and acute and long-term outcomes following aSAH. We hypothesize that the magnitude of iron released post-aSAH may saturate the mechanisms responsible for normal iron homeostasis and that genomic differences may alter response to this saturation, subsequently contributing to the variability in patient outcomes observed in our cohort. This study is strengthened by its use of data from a large prospective study conducted by a seasoned team of genomic experts which includes existing genome-wide methylomic and genomic data, as well as stored serial biospecimens, linked to extensive phenotype data. This study design allows for examination of hypotheses prior to utilization of any new funds or resources as well as replication of significant findings. The proposed study has the potential to illuminate important contributions to aSAH recovery and impact care through the following aims:

Aim 1: Characterize the relationship between temporal patterns of cerebrospinal fluid (CSF) DNA methylation of candidate genes related to iron homeostasis and variability of acute and long-term outcomes following aSAH. For each methylation site within candidate genes, participants will be assigned to trajectory groups based on their site-specific DNA methylation and potential relationships with patient outcomes will be examined using a two-tiered approach. Tier 1: Using existing genome-wide CSF DNA methylation data (n=260), trajectory groups will be inferred and analyzed for a relationship with acute outcomes (CV and DCI) and

long-term outcomes (Glasgow Outcome Scale [GOS] and mortality). Tier 2: Significant findings will be replicated in an independent sample (up to n=401, data generated by the candidate).

Hypothesis 1: *DNA methylation trajectory group assignments for methylation sites within genes related to iron homeostasis will be associated with variability of acute and long-term outcomes following aSAH*.

Aim 2: Examine the relationship between polymorphisms in candidate genes related to iron homeostasis and variability of acute and long-term outcomes following aSAH. Polymorphisms in candidate genes related to iron homeostasis will be analyzed using a two-tiered approach. Tier 1: Using existing genome-wide genotype data (n=245), candidate gene polymorphisms will be analyzed for a relationship with acute outcomes (CV and DCI) and long-term outcomes (GOS and mortality). Tier 2: Significant findings will be replicated in an independent sample (up to n=416, data generated by the candidate).

Hypothesis 2: *Variability in genes related to iron homeostasis will be associated with variability of acute and long-term outcomes following aSAH*.

1.2 Background and Significance

Aneurysmal subarachnoid hemorrhage (aSAH) affects an estimated 30,000 people in the United States each year.[10] While the incidence of aSAH seems relatively small, the effects are devastating causing tremendous burden not only for survivors and their families, but also on our healthcare system. Moreover, the overall mortality rate after aSAH is estimated to be between 25% and 50%[5,11] and when aSAH is not immediately fatal, over half of survivors suffer from long-term physical and neuropsychological outcome deficits.[4] The culmination of prolonged intensive care

unit and hospital stays, cost of treatment and extended rehabilitation, and dependence on caregivers, in addition to difficulty returning to work for survivors[12], the overall economic burden of aSAH is huge and increasing annually.[13,14]

Unfortunately, healthcare providers are often unable to predict which patients are at risk for poor outcomes. Data suggest that secondary complications such as cerebral vasospasm (CV) and delayed cerebral ischemia (DCI) contribute to the grim statistics associated with aSAH.[5] Although these secondary complications occurring during the acute phase window represent an opportunity for intervention to prevent or mitigate poor long-term outcomes, the causes of CV and DCI remain largely unknown and studies to develop therapeutic interventions are hampered by the inability to predict subpopulations of patients at risk. Research addressing the pathophysiology of the development of CV and DCI in humans is needed.

In preclinical models, heme and catabolized heme products (e.g., iron) within the subarachnoid space have been shown to be toxic to nearby tissues. Preclinical models have also illuminated biological mechanisms that contribute to a unique type of cell death called "ferroptosis" which is nonapoptotic in form but is dependent upon excessive intracellular iron levels. Ferroptosis has been shown to result in secondary ischemic brain injury (similar to DCI) in animals following aSAH.[6,7] Additional studies demonstrate that iron chelators effectively reduce neuronal cell death associated with this complication in various preclinical models of neurologic injury.[7,15–17] Unfortunately, the role of variability of candidate genes related to iron homeostasis and heme metabolism in aSAH in humans has received less attention in this regard which precludes advances in clinical care.

Moreover, despite the strong relationship between CV and DCI and poor outcomes, recent clinical trials targeting these complications have failed to improve patient outcomes after aSAH,

which underscores the disconnect of our understanding of CV and DCI pathology.[18] It is important to know if variability in the iron homeostasis- and heme-related candidate genes impacts long-term outcomes of death and functional status (Glasgow Outcome Score [GOS]) directly (or indirectly via an unmeasured acute complication of vascular or microvascular dysfunction). In an effort to address these gaps in knowledge, we are proposing to evaluate the relationship between genetic and epigenetic variation of iron homeostasis- and heme-related candidate genes and acute and long-term patient outcomes after aSAH.

Virtually all living organisms are dependent on iron for many biological mechanisms including hemoglobin formation, oxygen transport, cell signaling, host defense, energy metabolism, neuronal development, myelination, neurotransmitter production, and DNA synthesis.[19–21] Despite rapid iron utilization and turnover, plasma iron concentrations are generally stable under normal physiologic conditions, indicating that iron delivery and recycling is homeostatically controlled.[20] In aSAH, blood accumulates in the subarachnoid space where heme is catabolized by heme-oxygenase (inducted in response to cellular stress) into carbon monoxide, biliverdin, and non–protein-bound iron.[22] Under normal physiologic conditions, iron is tightly bound to carrier proteins as ferric iron and is recycled within the body.[23] However, because of physiologic changes associated with aSAH, including acidic brain tissue pH, hypoxia, and an influx of catecholamines in the extracellular fluid, iron is liberated to the less stable, non-protein bound ferrous form.[24] This non-protein bound iron is capable of accepting and donating electrons readily and has the potential to damage nearby tissues via the formation of hydroxyl radicals and oxidative stress.[8,25,26] In addition to preclinical studies demonstrating these metabolic mechanisms contribute to secondary brain injury in animals,[6,7] and the promise of iron chelators as a therapeutic intervention,[7,15–17,27] intriguing data from our team demonstrates that variability in genes related to

iron homeostasis are associated with poor outcomes and mortality post-aSAH.[28] Additional evidence from protein biomarker studies of human subjects suggests a relationship between free iron and unfavorable outcomes following aSAH, as well as a protective role of high CSF levels of important iron homeostasis proteins, ceruloplasmin[8] and ferritin.[29,30]

DNA methylation is a primary influence over gene expression largely influencing adult brain function, plasticity, and injury recovery and is thought to change rapidly during the cell cycle.[31] Pilot work conducted in support of the proposed study demonstrates a relationship between inferred DNA methylation trajectory groups of an iron homeostasis candidate gene and outcomes following aSAH[32] (Appendix C). Additional research demonstrates that hypomethylation of a candidate region of the genome is associated with increased mortality after ischemic stroke.[33] Similarly, global DNA hypomethylation occurred post-injury in rats with traumatic brain injury when compared to uninjured controls.[34] Unfortunately, the potentially alterable mechanisms underlying differential DNA methylation after aSAH is poorly understood. To date, no human studies have explored the relationship between DNA methylation trajectories of multiple candidate genes related to iron homeostasis and heme metabolism and patient outcomes following aSAH.

Human genetic polymorphisms in the proposed iron homeostasis candidate genes have been identified. Because of this, patients may not uniformly respond to therapeutics targeting dysregulated iron homeostasis. Pilot work in support of the proposed study demonstrates a relationship between genetic variation in an iron homeostasis candidate gene and patient outcomes following aSAH[32] (Appendix C). Additional data from our team supports exploration of genetic variability, indicating polymorphisms of candidate genes are significantly associated with patient outcomes following aSAH[35] and that CSF biomarker concentrations change over time and are associated with patient outcome variability.[36,37] To date, no human studies have explored the

relationship between genetic variation of multiple candidate genes related to iron homeostasis and heme metabolism and patient outcomes following aSAH.

1.3 Innovation

Despite preclinical studies suggesting its significance, current care standards do not acknowledge iron homeostasis or potential dysregulation as important following aSAH given the lack of research in humans. The proposed study employs an innovative multi-omic model to examine the interplay of iron homeostasis post-aSAH utilizing epigenetic (Aim 1) and genetic (Aim 2) approaches. The proposed study innovatively extends the understanding of iron homeostasis following aSAH on several fronts.

(1) The aSAH research community has no data related to the role of genetic and epigenetic variability in the iron homeostasis pathway in persons with aSAH, although there is great potential for these data to inform our understanding of pathology after aSAH.

(2) This project will explore a unique set of data related to dynamic DNA methylation during the acute phase recovery period in persons with aSAH. The impact of DNA methylation trajectories in the CNS may offer important data to understand disrupted physiological processes after aSAH.

(3) The use of this established aSAH cohort allows us to shed light on mechanisms in humans with associated clinical data, which lends itself more readily to clinical translation (vs. an animal model).

(4) Information found in this study may apply to other types of neurological injuries, such as ischemic stroke and traumatic brain injury, thereby extending its applicability and impact.

1.4 Design and Methodology

1.4.1 Study Overview

The proposed study is designed to characterize the relationship between epigenetic (DNA methylation) and genetic (polymorphisms) variability of the iron homeostasis pathway and patient outcomes after aSAH. The proposed study capitalizes on existing, prospective longitudinal data collected through the parent project including clinical and biological data from the acute outcome phase (inpatient stay from day of insult up to 14 days post-injury) and long-term outcome phase (interview at 3 and 12 months). Patient outcomes of interest in the proposed study include acute outcomes of CV and DCI and long-term outcomes of GOS and mortality. A schematic representation of the proposed aims can be found in Figure 1.

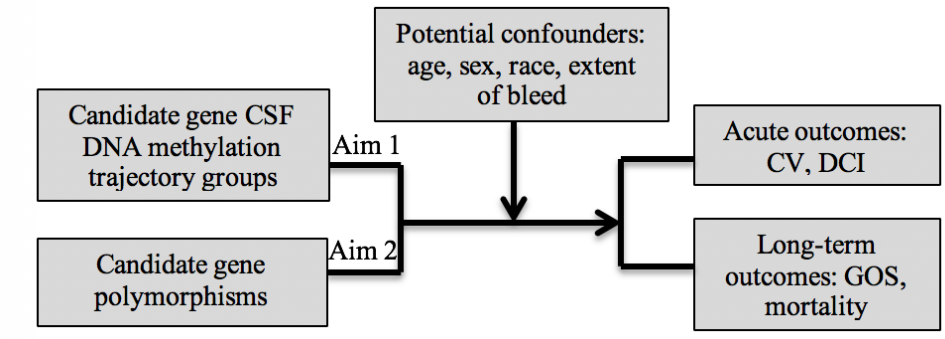


Figure 1 Schematic Representation of Proposed Study Aims

1.4.2 Sample, Setting, and Parent Project Description

This study utilizes data and biosamples collected through the parent project. Participants were prospectively recruited as part of NIH-funded studies [R01NR004339 and R01NR013610] approved by the University of Pittsburgh Institutional Review Board. Informed consent was

obtained from the study participant or their legal proxy upon recruitment. The cohort consists of 661 aSAH patients recruited from a Neurovascular Intensive Care Unit at UPMC Presbyterian in Pittsburgh, PA, USA. Patients were included in the parent project if they (1) were aged 18 years and older; (2) were newly diagnosed (≤5 days) with aSAH verified with cerebral angiogram; (3) had a Hunt and Hess grade ≥ 2 and/or Fisher grade ≥ 3; (4) were able to read/speak English; and (5) had no history of debilitating neurological disorders. Additional inclusion criteria for Aim 1 of the proposed study includes ventriculostomy insertion as part of clinical management (to supply CSF samples). The parent project collected a variety of data including demographic, confounding, acute and long-term outcomes, and methylomic and genomic data as well as stored CSF biosamples.

1.4.3 Sample Size and Justification

The sample sizes for tier 1 of the proposed study are set based on existing genome wide DNA methylation and genotype data available through the parent project. There are 661 participants enrolled in our aSAH study (encompassing both tier 1 and tier 2), however, sample sizes for tier 2 (replication in an independent test sample) will vary based on biosample and outcome data availability. Power calculations below were completed using the Power Analysis and Sample Size software version 16 (PASS, NCSS statistical software, LLC, Kaysville, Utah, USA).

Aim 1: The sample size for Aim 1 (tier 1) analysis, n=260 with up to five serial DNA methylation measurements, is set based on existing data. Tier 2 analyses will include a sample size of up to n=401 with up to five serial DNA methylation measurements. An illustrative power estimate for tier 1 demonstrates that if our outcome of interest, development of DCI, has a 44% chance of occurring[36], then we would have at least 80% power to detect an effect of methylation

on development of DCI with an odds ratio (OR) as small as 1.40 in the population when using a binary logistic regression analysis at a two-tailed significance level of 0.05 when there is a 0% correlation between DNA methylation and covariates. Although global DNA methylation has been shown to be associated with age (correlation as high as 0.97),[38] our pilot work shows this is not the case for CpG sites within an iron homeostasis candidate gene[32] (Appendix C). To provide a more conservative estimate of the detectable effect size, however, if we assume a 70% correlation between DNA methylation and a covariate, we will have at least 80% power to detect an effect of methylation on development of DCI with an OR as small as 1.84 in the population when using a binary logistic regression analysis at a two-tailed significance level of 0.05.

Aim 2: Similarly, the sample size for Aim 2 (tier 1) analysis of n=245 is set based on existing data. Tier 2 analyses will include a sample size of up to n=416. An illustrative power estimate for tier 1 demonstrates that a sample size of 245 achieves at least 80% power to detect an OR for DCI and a particular polymorphism as small as 1.43 in the population assuming a baseline prevalence of DCI of 44% when using binary logistic regression at a two-tailed significance level of 0.05 assuming no correlation between polymorphisms and covariates.

1.4.4 Data Collection

1.4.4.1 Candidate Genes

Candidate genes for the proposed study (n=39) were selected based on their known biological roles in iron homeostasis. Genes of interest in the proposed study are listed in Table 1. The known biological functions for each gene are presented in Appendix A.

Table 1 Iron Homeostasis Candidate Genes

ACO1	*FXN*	*PCBP1*
ACO2	*GSTP1*	*PGRMC1*
APP	*GLRX5*	*SCL46A1*
CALR	*HAMP*	*SLC11A1*
CD163	*HEPH*	*SLC11A2*
CP	*HFE*	*SLC25A37*
CUBN	*HJV (HFE2)*	*SLC40A1 (FP)*
CYBRD1	*HMOX1*	*SLC48A1*
FECH	*HMOX2*	*STEAP3*
FLVCR1	*HP*	*TF*
FTH1	*HPX*	*TFR2*
FTL	*LRP1*	*TFRC*
FTMT	*IREB2*	*TNF*

1.4.4.2 CSF DNA Methylation Data Collection (Aim 1)

Genome-wide methylome data (existing data [R01NR013610]): For Aim 1 (tier 1) data exploration, the proposed study capitalizes on existing genome-wide methylomic data (n=260 at up to 5 time points) collected with the Infinium Human Methylation450 BeadChip from Illumina using serial CSF biosamples. The HumanMethylation450 DNA Analysis BeadChip allows

interrogation of >450,000 CpG sites including CpG islands/shores/shelves associated with genes as well as outside of coding regions of the genome, dense coverage of promoter/5'/3' regions of genes including dense coverage of the promoter regions of mRNA genes. Bisulphite conversion of DNA was carried out; converted DNA samples were amplified then fragmented, precipitated, denatured and applied to the Infinium arrays. The BeadChip was then placed into a flow-through chamber and single base extension was performed on bound primers with labeled nucleotides. Staining steps were performed to attach florescent dyes to the labeled nucleotides and the final array was scanned using an Illumina iSCAN and the raw data analyzed using Genome Studio. The decision to use the Illumina Infinium HumanMethylation450 platform was based on the accuracy and reproducibility of bisulphate-based methods, the exceptional coverage and resolution of the data, its non-reliance on methylated DNA capture methods, and on budget and chip availability. For each candidate gene, CpG sites within the gene region +/- 2kb will be extracted from the genome-wide DNA methylation data for tier 1 analysis.

DNA methylation replication data (to be generated by the candidate): For Aim 1 (tier 2), DNA methylation status of the top methylation site identified in tier 1 analysis will be replicated in an independent test sample (up to n=401 at up to 5 time points) using stored serial CSF biosamples. While the final decision on a data collection platform will depend on the sites moved forward for replication (not all sites are amenable to all methylation data collection platforms) we propose at this point to use pyrosequencing. Bisulfite conversion of the DNA followed by pyrosequencing of the associated regions will be conducted for high priority regions requiring validation. Epitect Bisulfite Kits (Qiagen Corp) will be used to convert unmethylated cytosines to uracils. Two internal controls (1 fully methylated and 1 fully unmethylated) will also be converted and used to normalize the data and assist in data interpretation. PyroMark PCR Kit (Qiagen Corp)

will be used to generate the specific overlapping fragments for sequencing and perform the sequencing reaction for evaluation by pyrosequencing. Samples with incomplete bisulfite conversion or poor sequencing success will be eliminated from analyses. Technical replicates will be included.

1.4.4.3 Genotype Data Collection (Aim 2)

Genome-wide polymorphism data collection (existing data [R01NR013610]): For Aim 2 (tier 1) data exploration, the proposed study capitalizes on existing genome-wide polymorphism data (n=245 Caucasian participants) collected with Affymetrix Gene Chip Assay SNP 6.0 (Affymetrix, Santa Clara, CA, USA) using 500 ng of genomic DNA extracted from peripheral blood. The decision to use the Affymetrix SNP 6.0 was based on the accuracy of the data in addition to the exceptional coverage (906,600 single nucleotide polymorphisms [SNPs]). For each candidate gene, tagging SNPs within the gene region +/- 1 kb will be extracted from the genome-wide polymorphism data for tier 1 analysis. If no SNPs exist within a +/- 1 kb window, this will be expanded to +/- 3 kb window. Given that allele frequencies differ based on ancestry, the parent project collected genome-wide data for participants who self-reported as Caucasian only.

Polymorphism replication data collection (to be generated by the candidate): Genotype data of significant candidate genes from tier 1 analysis will be replicated in an independent test sample (up to n=416 Caucasian participants). The genotyping platform for tier 2 data collection will be selected based on the number of sites nominated (defined as any association that hits our level of significance [p-value <0.05]) by tier 1 analysis. For medium to high throughput (>24 SNPs nominated), the iPLEX MassArray platform will be used to allow for multiplex SNP genotyping with proven accuracy (>99.7% concordance rate). If low throughput (<24 SNPs nominated), TaqMan allelic discrimination assay with ABI Prism 7000 Sequence

Detection System (Applied Bioscience, Carlsbad, CA, USA) will be utilized. If genotype data cannot be collected for a SNP nominated by tier 1 analysis, another highly informative SNP in high linkage disequilibrium will be selected. As with tier 1, only Caucasian participants will be included in an effort to control for population substructure.

1.4.4.4 Participant Outcome Data

The proposed study capitalizes on existing patient phenotype data [R01NR004339]. As part of the parent project, participants were followed daily during hospitalization (days 0-14) and assessed at 3 and 12 months after aSAH by a trained neuropsychological technician or a registered nurse trained in neuropsychological testing. These time points were chosen to assess a proximal recovery time point (3 months) and a time point when neurological recovery is likely to have stabilized (12 months).

Acute phase outcomes selected for analysis in the proposed study include CV and DCI.

(1) Cerebral vasospasm (CV) was defined as the radiographic presence of cerebral vessel narrowing based on cerebral angiogram of ≥25% evaluated by a neurosurgeon and will be treated as binary (yes vs. no).

(2) Delayed cerebral ischemia (DCI) was defined as the presence of neurological deterioration, after exclusion of non-ischemic causes, that is accompanied by evidence of abnormal cerebral blood flow. DCI will be treated as binary (yes vs. no). Neurological deterioration measures include (a) increase of 2 or more points for National Institutes of Health Stroke Scale (NIHSS); (b) decline in level of consciousness on the Glasgow Coma Scale; (c) persistent (>1 hour) new focal neurological deficit; and/or (d) decline in pupil reactivity. Cerebral blood flow measures include: (a) cerebral angiography performed as ordered by the treating physician as clinical deterioration is present; or (b) transcranial Doppler

(sensitivity reported between 84%-85.7% and specificity between 89-100%[39,40]) performed daily for 14 days on all subjects enrolled in the parent project.

Long-term phase outcomes selected for analysis in the proposed study include the Glasgow Outcome Scale and mortality.

(1) The Glasgow Outcome Scale (GOS) which asks subjects to indicate their level of functioning. The scale ranges from 1 (death) to 5 (good recovery). Validity has been well established in patients with neurological insult.[41] GOS will be dichotomized as good (scores of 4 or 5) and poor (scores of 1, 2, or 3).

(2) Mortality data were updated through medical records/Social Security Death Index and will be treated as binary (yes vs. no).

Covariate/confounding variables: As part of the parent project, potential covariate/confounding data were collected from the medical record. These data include demographic variables (e.g., age, sex, race), clinical variables (e.g., severity of injury as measured by the clinical grading scales Fisher grade, Hunt and Hess (HH) score, and World Federation of Neurosurgical Societies [WFNS] grade), treatment variables (e.g., intervention [clip vs. coil], medication), and social variables (e.g., marital status, smoking status).

1.4.5 Scientific Rigor and Transparency

An important aspect of the proposed study design includes imbedded replication to allow for validation of findings adding to the study's robustness. Furthermore, extensive QC was (and will be) conducted on all project arms and disclosed. DNA methylation data QC (Aim 1): QC for tier 1 included consideration of the distribution of samples and phenotypes across arrays; evaluation of bisulfite conversion efficiency (samples with low bisulphate conversion efficiency or with

<95% CpG coverage/sample were excluded); inclusion of technical replicates and methylation controls; and use of data cleaning pipelines that addressed chip, column, row and batch effects. For tier 2, distribution of samples will be taken into consideration including distribution of duplicates amongst the batches and ensuring that all samples from a subject are run in the same batch. Technical replicates will be used to assess reliability of data. Data will incorporate a span of methylation standards ranging from fully methylated to fully unmethylated control samples within each batch of data collection and use the bisulphate conversion efficiency measures. Samples with low bisulphate conversion efficiency or with <95% CpG coverage/sample will be excluded. Genotype data QC (Aim 2): Samples will be checked for consistency and integrity of genotyping data by inclusion of duplicate controls on each plate for internal and plate-to-plate consistency, using genotype call rate criteria of >95%, comparing allele frequencies to frequencies in existing databases, and performing checks for Hardy-Weinberg Equilibrium (HWE) consistency. SNPs that do not meet QC standards will be eliminated.

1.4.6 Analyses

1.4.6.1 Descriptive and Preliminary Analyses

Statistical analyses will be conducted using R Statistical Software (Version 3.5.0, Vienna, Austria), SAS (Version 9.4, SAS Institute Inc., Cary, NC, USA), and PLINK (Version 1.9). Standard descriptive statistics will be computed in R for all independent, dependent, and potentially confounding/covariate data given the variable's level of measurement (e.g., means and standard deviations for continuous type normally distributed variables, frequency counts and percentages for categorical variables). Group comparisons will be performed using t-tests to evaluate equality of means for continuous type variables and Pearson's chi-square test of independence or Fisher's

exact test for equality of proportions for categorical variables. Preliminary analyses will be conducted to identify potential confounders/covariates. Known predictors of outcomes after aSAH include degree of hemorrhage (as measured by the clinical grading scale, Fisher grade), age, race, and sex.[42] Other potentially relevant confounders include smoking status, body mass index, comorbidities, education, and race. Using PLINK, Hardy Weinberg Equilibrium (HWE) will be calculated for each SNP. Raw data for SNPs in violation of HWE will be examined for data accuracy.

1.4.6.2 Data Screening Procedures

Data screening procedures will be performed in R given the variable's level of measurement. Data will be screened for accuracy using frequency tables for categorical variables (e.g., sex, race) and summary statistics (mean, standard deviation, minimum, and maximum) for continuous variables (e.g., age, DNA methylation). To identify outliers, data will be examined in frequency tables for categorical variables and in histograms and sina plots for continuous variables. If any potential outliers are detected, data will be inspected for accuracy and the potential influence will be evaluated in a sensitivity analysis. When possible, outliers will be retained in the analysis when accurate and full transparency will be used during project reporting and dissemination. For DNA methylation data, any DNA methylation value labeled as an extreme outlier (a DNA methylation value above or below three times the interquartile range) will be replaced with the maximum or minimum observed DNA methylation value below the extreme outlier threshold for values on each day. Data will be screened for independence using scatterplots, multicollinearity among independent variables using collinearity diagnostics (tolerance, variance inflation factors, and condition indices with variance decomposition proportions), and normality using histograms with

a normal curve overlay and Kolmogorov-Smirnov test of normality. The pattern of missing data will be assessed and described. Missing data will be extracted from the medical record if available.

1.4.6.3 Analyses for Aim 1

Overview: Data analyses for Aim 1 will be completed in two steps. First, DNA methylation trajectory groups will be inferred, and participants will be assigned a trajectory group class. Next, the relationship between group assignments and patient outcomes will be evaluated. For DNA methylation analyses, M-values will be used.[43]

Group-based trajectory analysis: Group-based trajectory analysis (GBTA) for methylation data will be conducted in SAS using PROC TRAJ assuming a censored normal model (which helps correct for right censored data within longitudinal studies). For individual CpG sites within candidate gene methylation data, the trend or change pattern over time will be examined in GBTA to identify the distinct trajectory patterns of site-specific methylation across time following aSAH. In GBTA, models with a varying number of groups and polynomial orders (group trajectory shapes) are compared to find the model that best fits the longitudinal data.[44] Given the subjectivity required in traditional GBTA modeling and the large number of candidate genes and CpG sites to be analyzed in the proposed study, the model selection process will be largely automated for 39 possible models with a maximum of three groups and comprehensive combinations of polynomial orders of 0 (intercept only), 1 (linear), and 2 (quadratic).

First, out of the 39 possible models, a best fitting 'preliminary' model will be identified based on the maximum Bayesian Information Criterion (BIC). In general, BIC measures improvement in model fit gained with the estimation of more parameters such as an increased number of groups and/or more complex trajectory shapes, but also applies a penalty for model complexity. Typically, larger BIC values indicate a better model fit. Using GBTA, two BICs will

be computed based on the number of participants and the total number of methylation observations over time with the true BIC lying within these values.[44] In cases where the BICs do not agree on the best preliminary model, we will assign a 'simplicity score' to assist in identifying the more parsimonious model. Following selection of a preliminary model for each CpG site, we will perform a secondary evaluation of model adequacy using several traditional diagnostics including (1) an average posterior probability (AvePP) >0.7, (2) odds of correct classification (OCC) >5, (3) estimated group membership (π) >5%, (4) reasonably close estimated group membership (π) versus the assigned group proportion (P*), and (5) a relatively narrow 95% confidence interval for the estimated group probability (π).[45] If preliminary models fail secondary evaluation, the 'next best' fitting model will be selected and secondary evaluation will be performed again. If preliminary models for a CpG site fail a second time, we will conclude that methylation trajectory groups cannot not be inferred with high accuracy for that site and it will be excluded from further analysis. A more detailed explanation of GBTA methodology including protocol and flow charts depicting best model selection can be found within the publication summarizing the pilot work for this study[32] (Appendix C).

Cell type heterogeneity: Of particular importance to the proposed study is cell type heterogeneity (CTH). CTH is an important consideration in methylation studies as differential methylation between cell types may impact overall methylation levels and confound results. In cases where genome-wide data are collected, there are methods to control for CTH.[46] However, if only CpG site-specific DNA methylation data are collected, as would be the case with a biomarker used clinically, controlling for CTH is not possible. In an effort to explore the potential clinical utility of methylation values unadjusted for CTH (as they would likely be in the hospital setting) as well as effects of CTH, the aforementioned procedures will be conducted without adjustment

for CTH ('unadjusted models') and with adjustment for CTH ('CTH-adjusted models') for all CpG sites in tier 1. However, only unadjusted analyses will be performed in tier 2 as genome-wide data will not be available in the replication sample. Using the parent study genome-wide data, CTH data will be computed in R statistical software using Houseman's reference-free method.[47] In the CTH-adjusted models, CTH data will be included as a time-varying covariate.

Trajectory group - patient outcome associations: Binary logistic regression will be performed in R to determine the relationship between inferred trajectory groups for each methylation site and patient outcomes while controlling for age, sex, race, and Fisher grade. Given the exploratory nature of this study, relatively small sample size, and correlation between outcomes, correction for multiple testing will not be made in tier 1. However, in the tier 2 replication analyses, permutation testing, rather than strict Bonferroni corrections, will be used to correct for testing correlated outcomes. This will be done by shuffling the independent variables 10,000 times, performing binary logistic regression for each outcome, and recording the minimum p-value across the outcome-specific results. In this distribution of 10,000 simulated null minimum p-values, the value that occurs at the 95^{th} percentile is taken to be the empirical significance threshold T – that is, to test at the 0.05 significance level while correcting for testing six outcomes, we would have to observe a minimum p-value less than T to consider the results significant.

Potential post hoc analysis: For trajectory groups with significant associations with patient outcomes, determine if there are statistical differences between the frequencies of intervention type (clip vs. coil) for methylation groups.

Interpretation of results: Pathway analysis (e.g., String-DB, Ingenuity Pathway Analysis) may be conducted to evaluate hyper- and hypo-methylated genes in relationship to one another to aid in the interpretation of results.

1.4.6.4 Analyses for Aim 2

Genetic association analyses will be performed in PLINK using a binary logistic regression model adjusted for age, sex, and Fisher grade. Given the large number of SNPs to be examined, only additive models (treating SNPs as ordinal based on variant allele dosage) will be considered. The OR and 95% confidence interval will be calculated for each SNP. If outliers exist within the SNPs and we are unable to perform an additive test, SNPs will be discarded for the primary analysis and models will potentially be collapsed and examined in post hoc analysis. The Approximate Bayes Factor will be used to compute Bayesian False Discovery Probabilities (BFDP) for each SNP.[48,49] The BFDP will then be used to prioritize SNPs for replication in tier 2. In order to aid in the interpretation of results, using the tier 1 genome-wide data, principal components analysis (PCA) will be conducted using PLINK to determine the role of genetic ancestry membership compared with self-reported race. In tier 2, however, only self-reported race will be considered as a covariate given genome-wide data will not be available. As described above, multiple testing will not be made in tier 1. However, in the tier 2 replication analyses, permutation testing will be used to compute an empirical significance threshold as described.

1.4.7 Limitations

Although this dissertation study has many strengths, there are several important limitations that should be acknowledged.

(1) This study will not address all potential regulatory mechanisms of candidate genes and will not measure iron or candidate gene protein levels.

(2) By capitalizing on data and samples from an existing cohort, we are limited to existing outcome and covariate data previously collected.

(3) The population examined in the proposed study is limited to primarily European descent, although this is consistent with Pittsburgh, PA demographics.

(4) The majority of subjects from this study are female, although this is consistent with aSAH demographics.

(5) Not all aSAH insults require the collection of CSF post-aSAH, which limits the applicability of the results of Aim 1.

(6) All aSAH patients receive a standard treatment protocol, however, there is potential for treatment variability. Although we have data on treatment and interventions, it is unlikely we will be able to control for this variability due to the overwhelming loss of power.

(7) Genome-wide data may not be representative of important CpG sites or tagging SNPs for candidate genes.

(8) The power of Aim 2 is relatively low as most genetic effects for complex diseases have OR <1.4.

(9) Given the exploratory nature of this study, there will be a large number of tests conducted which limits the interpretability of the results.

1.4.8 Future Directions

Knowledge garnered from this dissertation study will serve as a foundation for the candidate's future research, which has the potential to impact nursing science by offering insights into the etiology and biological mechanisms of patient outcomes after aSAH. Future potential directions of this work include investigating the relationship between iron levels and protein biomarkers of sites nominated from the analyses of Aims 1 and 2 and patient outcomes (with priority given to concordant genes) to inform evaluation of additional phenotype-influencing mechanisms.

Additional potential future directions include enlargement of the aSAH cohort to increase power; replication of the proposed study in additional ancestries; examining the role of iron homeostasis in the context of endogenous environment (e.g. estrogen; testosterone) post-aSAH; exploration of the utility of alternate biosamples for findings related to DNA methylation (e.g. blood); development of therapeutic interventions to improve outcomes after aSAH; and exploring variability of the iron homeostasis pathway in other neurologic injury populations (e.g., traumatic brain injury). Importantly, the range of competencies gained through completion of this study, particularly the combination of genomic knowledge and bench and computational skills, will prepare the candidate to lead a multidisciplinary team in the future.

1.5 Hazardous Material and Procedures

Working with human subject biosamples involves some risk; to reduce risk the study will be conducted following training in approved labs and using personal protective equipment and safer sharps. The candidate has completed, and will continue to complete annually, blood borne pathogen safety training and chemical hygiene training in addition to ongoing laboratory safety training. Additionally, the laboratory where the candidate will conduct data collection is fully equipped to handle biohazards safely and is OSHA compliant. The University of Pittsburgh Environmental Health and Safety office conducts an annual inspection of all laboratories on campus. The University of Pittsburgh is also in full compliance with disposal of biohazardous and chemical wastes.

1.6 Research Participant Risk and Protections

This study has current IRB approval (STUDY19100368). A copy of the most recent IRB approval memorandum and modification is located in Appendix J. All study participants included in this project have identical recruitment and informed consent processes as well as data collection protocols. All participants were recruited through the parent project, which is made up of two NIH funded projects [R01NR004339 and R01NR013610]. The proposed study will utilize the samples and data collected from the parent project and does not itself recruit subjects or involve contact with subjects.

1.6.1 Human Subjects Involvement, Characteristics, and Design

Participants that will be included in this study have completed two separate informed consents: one to participate in the study entitled "Determining Genetic and Biomarker Predictors of DCI and Long Term Outcomes after aSAH (NR004339)", which will be referred to as the "NR004339" study and one to allow for collection and utilization of their genomic material in connection with the data collected through "NR004339". These subjects were admitted to the neurological intensive care unit of UPMC-Presbyterian Hospital as a consequence of aSAH. Inclusion criteria for the parent project includes: (1) aged 18-75 years; (2) newly diagnosed (≤5 days) aSAH verified with cerebral angiogram; (3) Hunt and Hess grade greater than or equal to 2 and/or Fisher grade greater than or equal to 3; (4) able to read/speak English; and (5) no history of neurological disorders. Additional inclusion criteria for Aim 1 of the proposed research is: (1) ventriculostomy insertion as part of clinical management.

1.6.2 Inclusion of Special Classes

Women of child-bearing age or pregnant women will not be excluded from the study.

1.6.3 Sources of Materials

This project will use anonymous specimens and data. The samples collected for this study are coded with the study's unique identification number while at the clinical site and the candidate and laboratory technicians are only provided samples with this unique identification number. The biological samples that will be used for this project are CSF, blood, and DNA. The candidate will not have access to any of the linkage information to identify these anonymous samples, nor will she need this information.

1.6.4 Potential Risks

All data collected for this study will be collected using de-identified samples and all demographic and phenotype data provided for analyses will be provided using these de-identified codes. Specimens for processing arrive at the parent project primary investigator's laboratory (Conley lab) already de-identified. Using anonymous samples and databases reduces the risk to the subjects. Breach of confidentiality is a theoretical risk for this study. Confidentiality of the genetic information gathered on each subject will be facilitated by the fact that the candidate will not know any personal identifiers for the subjects for the proposed study. Even though the data will be anonymous and reported in aggregate only, the data files will be protected by passwords and all hard copies of the genetic data generated will remain within the laboratory, which is always locked

and not accessible to unauthorized persons. All material and data used in this study will be used exclusively for research purposes.

1.6.5 Recruitment and Informed Consent

Participants that will be included in this study were recruited and informed consent was obtained through efforts of the NR004339 study. The genetic portion of the study utilizes an additional, separate consent form that covers information about investigating genes and gene products that may inform us about recovery after aSAH as well as linking the genetic data to the phenotypic data in an anonymous manner. The University of Pittsburgh Institutional Review Board approves the consent forms and the protocols for recruitment and data/specimen collection annually.

1.6.6 Protections Against Risk

All of the samples and data that will be utilized for the proposed study will be provided in a de-identified manner through the efforts of the parent project. The proposed study strictly utilizes samples and data from the parent project and does not at all require any additional subject contact. Therefore, the proposed study poses no additional health or safety risks to the subjects. Confidentiality of the genetic information generated by this project, a theoretical risk, is a very minimal risk given that all of the samples and the databases are anonymous, and each sample is associated with only a unique code. In addition, every precaution will be taken to minimize exposure of the data to persons outside of this project by using passwords for all computer files and keeping all hard copies of data within the laboratory, which are always locked, and

unauthorized personnel not admitted. All data generated from this project will be reported as aggregate data only.

1.6.7 Potential Benefits of the Proposed Research

There are no known direct benefits to participants. The results of the study may lead to future understanding of genetic and other biological factors involved in complications after aSAH. Such knowledge could translate into improved intervention and therapy for aSAH survivors in addition to a better understanding and intervention potential for patients suffering from other conditions involving neuronal injury from cerebral hemorrhage. Given the minimal risk to these anonymous patients, the potential benefit of this project far outweighs the risks.

1.6.8 Importance of the Knowledge to be Gained

The complications that occur in aSAH survivors occur at great personal, family, and social cost. Such complications might be reduced in future aSAH survivors if more were known about the biological factors involved in onset and severity of these complications. The data generated from this project will answer questions regarding epigenetic and genetic variation in the iron homeostasis pathway and the impact on outcomes in persons following aSAH.

1.6.9 Data Safety Monitoring Plan

This is not a clinical trial; however, a data safety monitoring committee to review the collection and integrity of data is in place for the NR004339 study and the activities for this project will be incorporated into their efforts.

2.0 Summary of Study

2.1.1 Overview

The purpose of this dissertation study was to examine the relationships between epigenetic (Aim 1) and genetic (Aim 2) variability of genes from the iron homeostasis pathway and patient outcomes after aneurysmal subarachnoid hemorrhage (aSAH). Two data-based papers directly related to the aims of this dissertation have been published in the peer-reviewed journal *Neurocritical Care* and are provided in Appendices C and D, respectively. The first article, entitled "Genetic Variability and Trajectories of DNA Methylation May Support a Role for *HAMP* in Patient Outcomes After Aneurysmal Subarachnoid Hemorrhage" was a pilot study exploring associations between genetic and epigenetic variability of one iron homeostasis candidate gene, *HAMP*, and patient outcomes after aSAH[32] (Appendix C). The second article, entitled "Genetic Variability in the Iron Homeostasis Pathway and Patient Outcomes after Aneurysmal Subarachnoid Hemorrhage" summarizes the results of Aim 2 of this dissertation study[50] (Appendix D).

The results for Aim 1 of this dissertation are presented in the data-based manuscript, "Associations Between DNA Methylation Trajectories in Iron Homeostasis Candidate Genes and Patient Outcomes after Aneurysmal Subarachnoid Hemorrhage," included in Section 3.0 of this document.

In addition to the two data-based publications directly related to the aims of this study and the third data-based manuscript presented in Section 3.0, an additional five articles related to nursing research and omics were written during the course of PhD training and are provided in

Appendices E, F, G, H, and I. The first article, published in *Biological Research for Nurses* and entitled "Evaluation of *APOE* Genotype and Ability to Perform Activities of Daily Living Following Aneurysmal Subarachnoid Hemorrhage" summarizes a data-based project completed through an apprenticeship during doctoral training[51] (Appendix E). This study was in the aSAH population and provided early training in statistical analysis and genetics. The second article, published in *Biological Research for Nurses* and entitled "Symptom Science: Advocating for Inclusion of Functional Genetic Polymorphisms" was a review paper that identified functional polymorphisms associated with common symptoms[52] (Appendix F). The conclusion of this paper provided a parsimonious list of polymorphisms for suggested use as biological common data elements in symptom studies across nursing research. The third article, published in the *Annual Review of Nursing Research* and entitled "Omics for Nurse Scientists Conducting Environmental Health Research" was a review paper that provided an overview of omics methodologies for nurse scientists conducting environmental health research, exemplar applications of these techniques in the literature, future directions for nursing research, and funding opportunities that demonstrate the growing need and interest of environmental health research[53] (Appendix G). The fourth article, published in the *Journal of Nursing Measurement* and entitled "Psychometric Properties of the Patient Assessment of Own Functioning Inventory Following Aneurysmal Subarachnoid Hemorrhage" was an exploratory factor analysis evaluating the underlying factor structure of a self-assessment tool after aSAH[54] (Appendix H). The fifth article, currently under review with *Frontiers in Genetics* and entitled "Characterization of Cerebrospinal Fluid DNA Methylation Age During the Acute Recovery Period Following Aneurysmal Subarachnoid Hemorrhage" compares the performance of three epigenetic clocks in cerebrospinal fluid (CSF) after aSAH (Appendix I).

2.2 Proposal Changes

Several changes were made to the approved dissertation proposal. These changes and associated rationales are provided below.

2.2.1 Aim 1

2.2.1.1 Replication Data

The approved dissertation proposal included replication of findings for both Aim 1 and Aim 2. While replication data were generated, analyzed, and published for Aim 2[50], this was not possible for Aim 1 due to delays at the Center for Inherited Disease Research (CIDR) laboratory (Johns Hopkins University, Baltimore, MD) related to the COVID-19 pandemic. While DNA was extracted from biosamples and sent to CIDR in November of 2019 for replication of findings for our top three hits from Aim 1, these data have not yet been returned. Therefore, replication analyses have not been included in this dissertation document but will be analyzed and published in the future.

2.2.1.2 ComBat Sensitivity Analysis

An important consideration in omic studies is systematic technical variation. One such source of technical variation is confounding of data due to experimental batch effects (i.e., differences in DNA methylation values due to between plate differences). This is a particular concern in DNA methylation studies because a batch effect could be interpreted as a biologically important finding.[55] The first step in reducing batch effects is careful experimental design and distribution of biosamples on the plates. However, even with careful and thoughtful experimental design, batch

effects may still exist. A slight deviation was taken in this study to explore the potential utility of ComBat correction[56] for batch effects. In a sensitivity analysis for Aim 1, ComBat correction[56] for batch was applied to the genome-wide data and an Aim 1 sensitivity analysis for three of our candidate genes was performed. The results of those analyses were quite discordant from the original Aim 1 analyses. After careful deliberation and evaluation of the data, it was decided that the use of ComBat resulted in an overcorrection for batch and was not required. It should be noted that batch correction was made as part of the quality control (QC) pipeline as implemented in the 'preprocessfunnorm' function from the minfi package.[57,58] This method of functional normalization is an extension of quantile normalization often used in DNA methylation studies. While functional normalization does not remove effects due to biological phenomena (e.g., cell-type heterogeneity [CTH]), it does use internal probes to remove effects of technical variation (e.g., batch effects) and performs quite well compared to other methods of batch correction including ComBat.[57]

2.2.1.3 Preprocessing Method Sensitivity Analysis

To accommodate advancements in DNA methylation array technology, many QC pipelines have been developed utilizing different preprocessing methods to clean and prepare data for analysis. One such preprocessing method that is an important step in DNA methylation data analysis is data normalization. This is critical in processing Illumina DNA methylation data as different probe designs use different chemistry and possess different beta distributions.[57] As described above, the genome-wide DNA methylation data available for Aim 1, tier 1 of this study were normalized using functional normalization as implemented in the preprocessFunnorm function.[57] However, we identified a similar method known as 'funtooNorm' that is implemented in the funtooNorm package.[59] Importantly, funtooNorm was designed to better handle data gathered across time

points and allows for interactions between tissue-types.[59] Because the aSAH DNA methylation data were generated across time in CSF, a tissue in which we expect (1) methylation profiles to differ as a function of time post-injury and (2) an array of dynamic cell-types, we wanted to understand how these preprocessing methods performed in our data. In an effort to explore this, a tier 1 sensitivity analysis was performed using the funtooNorm normalized data. The results of this sensitivity analyses were quite concordant and are discussed in the data-based manuscript in Section 3.0.

2.2.2 Aim 2

2.2.2.1 Tier 2

In the original dissertation proposal, it was stated that correction for multiple testing in tier 2 of Aim 2 would be made by calculating an empirical significance threshold using permutation analysis. However, because the Bayesian methods used in tier 1 were carried forward for tier 2 analysis, correction for multiple testing using permutation was not required. Specifically, because the Approximate Bayes Factor (ABF) is computed using the maximum likelihood estimate and standard error from a logistic regression analysis (which in turn accounts for sample sizes, minor allele frequency, and frequency of individuals within each genotype) as well as the prior odds on the null for each SNP, the Bayesian False Discovery Probability (BFDP) can be interpreted as a probability of false discovery regardless of power, sample size, or how many SNPs were tested.[50] The methods, results, and interpretation of data associated with Aim 2, including this change to the original proposal, have been peer reviewed and published.[50] A copy of this publication is included in Appendix D.

2.3 Study Strengths and Limitations

This hypothesis-driven dissertation was the first study to formally examine epigenetic and genetic variability of candidate genes from the iron homeostasis pathway in patient outcomes after aSAH in humans. The availability of existing genome-wide genetic and DNA methylation data from which to extract our candidate gene data for tier 1 analyses offered a unique opportunity to explore the multi-omic features of patient outcomes without using additional resources or funds. Moreover, the DNA methylation data used in Aim 1 were quite unique. In addition to being longitudinal in nature spanning 14 days post-aSAH, these data were also generated from CSF, a tissue which we believe offers unique insight into the local environment of the central nervous system. By analyzing these unique data using trajectory analysis, we were able to identify dynamic changes in DNA methylation after aSAH. Similarly, because genome-wide data existed for our tier 1 analyses, we were able to perform additional sensitivity analyses that we could not have performed in a traditional candidate gene study (e.g., effects of CTH and preprocessing methods in Aim 1 and effects of ancestry in Aim 2).

A notable part of this dissertation study was the development of a novel group-based trajectory analysis (GBTA) protocol which allowed us to largely automate GBTA for a large number of methylation sites. Additional strengths of this dissertation included the rigorous quality control of omic data, the use of Bayesian statistics to aid in issues of multiple testing and interpretation of findings, and imbedded replication (though Aim 1 data have not been returned for analysis). These strengths, paired with examination of important and unique phenotypes of patient recovery post-aSAH, provided a rigorous and novel study with important nursing implications as summarized below.

Despite the many strengths of this study, there are some important limitations that should be acknowledged. First, this study was a candidate gene association study. Some experts argue that there is little need for candidate gene approaches in the era of big data and that selecting genes *a priori* introduces bias. However, while genome-wide studies offer a great deal of value, when considering the relative impact of individual genes in complex phenotypes such as aSAH outcomes, it is possible that important genetic contributors will not rise to significance in genome-wide analyses. In other words, while iron homeostasis may be important to phenotypes selected for a given study, candidate genes may not appear significant in a genome-wide analysis even in the presence of real effects. While genome-wide studies are better able to detect novel genes of larger impact, they fall short in identifying contributions from multiple genes of smaller impact that may be quite important to the biology of recovery post-aSAH. In addition, genome-wide methods for GBTA of DNA methylation data have not yet been developed and, for now, must be confined to candidate genes. To combat the inherent limitations associated with the nature of this study, we imbedded replication in the study design. However, as discussed above, replication data for Aim 1 are pending so at this time we are unable to determine if our findings can be replicated.

In addition, for the discovery analyses of this study, we extracted genotype and DNA methylation data for our genes within the gene transcript region ± a designated window of 2,000 base pairs and 1,000 base pairs for Aims 1 and 2, respectively. While this region was selected based on public databases and the advice of experts in genomic regulatory mechanisms, for DNA methylation in particular, this window is not be wide enough to capture all CpG sites important to epigenetic regulation of our candidate genes. Broadening this genomic region is an important future direction, particularly for genes that have been observed to be responsive to nuclear

receptors (e.g., *HAMP*).[32] For these genes, it is possible that long-range interactions may be centered on our loci of interest and were therefore not captured in this study.

Next, while the sample sizes for this study are quite large compared with other aSAH patient outcomes studies, we were underpowered for Aim 2 and could have missed small but real effects. Moreover, candidate genes were selected based on their known biological roles in iron homeostasis, however, given the dynamic nature of the science and complexity of iron homeostasis, this list likely does not capture all genes involved in this pathway. Additionally, our sample size was primarily limited to participants who self-reported their race as White impeding generalizability of findings. Given the role of race and ancestry in omic mechanisms, the scientific premise of this study should be explored in more detail in other ancestries in the future.

Specific to Aim 1, important limitations related to our analytical decisions should be understood. While there are several types of trajectory analyses in the literature, we chose to use GBTA to model longitudinal changes at DNA methylation sites. Hierarchical modeling and latent curve analysis are methods of modeling longitudinal changes that are quite common in the literature, but GBTA has receive recent attention as an alternative and there are some important differences in these methods. If applied to our DNA methylation data, hierarchical modeling and latent curve analysis would estimate the sample average DNA methylation trajectory of our sample and use covariates to explain variability about this average. In contrast, group-based trajectory modeling performed in Aim 1 assumes that the sample is composed of distinct groups, each with a different underlying DNA methylation trajectory. This method allows us to estimate not only how covariates impact trajectory shapes, but also group membership. GBTA allowed us to ask and answer questions specific to our research questions and hypotheses but other types of longitudinal analyses could be valuable in uncovering important signals in our data. Moreover, during

traditional GBTA, researchers iteratively evaluate and refine models. This was not feasible in a dataset this size so we created an objective protocol accompanied by R and SAS scripts to help automate GBTA as described in detail as part of the pilot work for this study[32] (Appendix C). Importantly, this objective protocol used posterior model QC criteria and when a CpG failed model QC twice, it was eliminated resulting in a sizable reduction of CpG sites formally examined for associations with patient outcomes (specifically, only 183 of our 573 candidate CpG sites were actually carried forward for testing with patient outcomes). While this resulted in more rigorous modeling, an unfortunate consequence was the exclusion of many CpG sites within each gene from patient outcome association testing.

Specific to Aim 2, while our tier 1 genome-wide data QC pipeline resulted in more accurate and rigorous analysis, it reduced our tier 1 sample size (specifically, a reduction of 244 to 197 participants). Likewise, because we limited our tier 1 analysis to SNPs available in the discovery data (i.e., genome-wide genotype data), we were unable to fully evaluate variability in every gene. Likewise, in a few cases, no SNP data existed for a gene and we were therefore unable to evaluate any variability for that given gene. Therefore, it is possible that the SNPs or genes not examined in tier 1 had important associations with outcomes. Additionally, as described[50], we were missing replication data for some associations flagged as noteworthy in tier 1 so were unable to determine if all associations could be replicated. Therefore, while two genes rose to the top in our analyses (Ceruloplasmin [*CP*] and Cubilin [*CUBN*]), we can't necessarily eliminate the other candidate genes as plausible targets for future investigation in aSAH recovery research.[50] In addition, while both genetic and epigenetic variability of our candidate genes were examined for associations with patient outcomes, we did not identify any genes that overlapped as potentially important between the Aims of this study, though we have discussed several possible reasons for this in the limitations

above. Finally, the scope of this study did not include functional investigation of genes flagged as noteworthy nor did it include addition of other omic mechanisms (e.g., gene expression, protein quantification).

2.4 Future Studies and Implications for Nursing

This study provided important doctoral training and data for future research. Future work specific to Aim 1 should include analyzing the replication data currently pending from CIDR to determine if our top hits can be replicated. If our findings replicate, additional functional investigation of the top hits should be conducted to better understand the biology of aSAH recovery and potential avenues for clinical translation. Future work specific to Aim 2 should first attempt replication of the top hits for which we were lacking replication data as well as investigation of genes for which tier 1 data did not exist. Additionally, in an effort to translate our findings clinically, genes identified as noteworthy in our Aim 2 analyses (*CP* and *CUBN*) should be examined functionally after aSAH. Additional future directions specific to this study include investigating our findings in more diverse ancestries and exploring other omic mechanisms of these candidate genes (e.g., gene expression). Important future directions in aSAH patient outcomes research should include examining the dynamic nature of DNA methylation post-injury from a genome-wide point-of-view (versus the candidate gene approach applied in this dissertation). The availability of serial CSF DNA methylation is a unique resource that could provide valuable insight into the biology of aSAH recovery including the response to injury and development of acute complications.

This dissertation study has great potential for clinical translation. Specifically, if our findings for Aim 1 replicate in an independent test sample, it is possible that the trajectory groups

could be used to identify DNA methylation thresholds for poor outcomes. In other words, a range of DNA methylation values could be determined from the trajectory groups which could be used to flag future patients for more intensive clinical management. Beyond this, the findings of this study may provide a broader insight and value beyond the potential for identifying clinical biomarkers. Our analytical choices have allowed us to uncover dynamic changes in the acute recovery period which may provide insight into pathology of the development of poor outcomes and be potentially relevant therapeutic targets in the future.

During the conduct of this dissertation work, significant insight was gained into the importance of training nurses in omics, statistical genetics, and bioinformatics. Currently, there is not a sufficiently trained cadre of nurse scientists proficient in these areas. Given the increasing size and complexity of omic data, in conjunction with omic-specific QC and analysis considerations (e.g., Hardy-Weinberg Equilibrium, effects of ancestry and relatedness, downstream effects of preprocessing pipelines), it is recommended that nursing PhD students studying omics should receive, at a minimum, introductory training in both statistical genetics and bioinformatics in addition to required training in biostatistics. Nurses can provide important insight into the development of statistical methods for omic data, particularly longitudinal omic data collected in unique phenotypes (e.g., symptoms, patient outcomes, end-of-life, wellness across the lifespan), and should be integral team members in data analysis.

Overall, this dissertation work has provided a foundation for a line of research to identify clinically relevant biomarkers and improve outcomes and lives after aSAH. Moreover, the research skills developed during the conduct of this dissertation study will allow for application of the skills and knowledge gained in additional populations of chronic and complex conditions. Nurses belong at the forefront of symptoms, patient outcomes, end-of-life, and wellness research as we coordinate

and conduct biomarker testing, create patient-centered care-plans post-injury or illness, and deliver patient healthcare and wellness education across the lifespan and omics can bring important insights to these types of studies. Likewise, the nursing lens is very much needed in omic studies, including how we approach the study and interpret the data. The value of omics in nursing science and practice cannot be overstated and is an important area of training and research.

3.0 Data-Based Manuscript: Associations Between DNA Methylation Trajectories in Iron Homeostasis Candidate Genes and Patient Outcomes after Aneurysmal Subarachnoid Hemorrhage

3.1 Abstract

Background/purpose: Preclinical models have identified a biological mechanism that contributes to a unique type of cell death called "ferroptosis" which is nonapoptotic in form but is dependent upon excessive intracellular iron levels. Following hemorrhagic brain injury such as aneurysmal subarachnoid hemorrhage (aSAH), the brain is particularly susceptible to cell death by ferroptosis. Animal models of subarachnoid hemorrhage have found that therapeutic intervention targeting the iron homeostasis pathway shows promise for mitigating the effects of ferroptosis and improving recovery, but very little work has been conducted following aSAH in humans. Therefore, the exploration of trajectories of DNA methylation in iron homeostasis candidate genes is an important area of patient outcomes research that may provide insight into the biological mechanisms related to ferroptosis and associated with poor outcomes post-aSAH. The purpose of this study was to examine the relationship between trajectories of DNA methylation in candidate genes related to iron homeostasis (n=39) and acute and long term outcomes after aSAH. **Methods:** This study was a longitudinal, observational study of participants with aSAH that included two samples (targeted discovery [n=260 at up to five time points] and replication [n=123 at up to five time points]). Serial cerebrospinal fluid (CSF) samples were used to generate longitudinal DNA methylation data over 14 days post-aSAH. Acute outcome data were collected for cerebral vasospasm (CV) and delayed cerebral ischemia (DCI) over 14 days post-aSAH and long-term outcome data of functional status

measured using the Glasgow Outcome Scale (GOS, poor = 1-3) and death were collected at 3 and 12 months post-aSAH. Group-based trajectory analysis was used to infer distinct DNA methylation trajectory groups for CpG sites within our candidate genes and binary logistic regression was used to determine the relationship between inferred trajectory group assignment and acute and long-term patient outcomes. **Results:** In the discovery sample, 637 CpG sites were available in 36 candidate genes for analysis. DNA methylation trajectories with two or more groups were inferred at 183 CpG sites in 33 candidate genes and carried forward for association testing with patient outcomes. Our top three hits included cg08866780 in the Amyloid Precursor Protein gene (*APP*), cg25713625 in the STEAP3 Metalloreductase gene (*STEAP3*), and cg08553327 in the Tumor Necrosis Factor gene (*TNF*). Specifically, significant associations were identified between cg08866780 (*APP*) and GOS at 12 months ($p = 0.0001$) and death at 3 and 12 months (both $p = 0.0002$) and a suggestive association with GOS at 3 months ($p = 0.01$); significant associations between cg25713625 (*STEAP3*) and GOS at 3 and 12 months ($p = 0.00005$ and $p = 0.0005$, respectively) and death at 3 and 12 months ($p = 0.0013$ and $p = 0.0015$, respectively); and a significant association between cg08553327 (*TNF*) with GOS at 3 months ($p = 0.00001$) and suggestive associations with GOS at 12 months ($p = 0.02$) and death at 3 and 12 months ($p = 0.009$ and $p = 0.004$, respectively). Biosamples from an independent test sample were sent for pyrosequencing to replicate these findings, however, those data were not returned for inclusion at the time of publication of this dissertation documents. **Conclusion:** Among the genes examined, CpG sites in *APP, STEAP3,* and *TNF* rose to the top as candidates for replication. If these findings can be replicated, efforts to translate findings into clinical practice should include additional research examining the functional role of these genes after aSAH.

3.2 Background

Aneurysmal subarachnoid hemorrhage (aSAH) is a devastating type of stroke with a high degree of variability in recovery between patients. The 30 day aSAH fatality rate falls between 25% and 50%[5,11] and a large percentage of survivors never fully recover or function independently.[3,4] Of those who survive, a range of symptoms and functional disability is observed which impacts cognitive function and mental health, ability to perform activities of daily living and return-to-work, and ability to maintain healthy relationships.[4] Moreover, the lifetime costs associated with hospital stays, medication, rehabilitation, loss of productive life years, and dependence on caregivers results in an overall enormous economic burden.[12–14]

Despite advances in healthcare over recent years, outcomes for many patients after aSAH remain poor. Although there are several predictors of poor outcomes after aSAH (e.g., severity of hemorrhage, clinical presentation, age), we continue to observe a high degree of variability between outcomes in patients despite similar characteristics and injuries (e.g., two patients with the same age and similar size of hemorrhage/aneurysm location). It is widely accepted that secondary complications occurring during the acute recovery period such as cerebral vasospasm (CV) and delayed cerebral ischemia (DCI) are important predictors of poor long-term outcomes but we still don't fully understand the pathology of these complications and are therefore unable to predict or prevent their occurrence.[5] Moreover, despite the strong relationship between CV and DCI and poor long-term outcomes, clinical trials targeting these complications have failed to improve patient outcomes after aSAH, which underscores the disconnect in our understanding of CV and DCI pathology.[18] Research addressing the pathophysiology of the development of these acute complications in humans is needed to improve both acute and long-term outcomes following aSAH.

Multiple forms of cell death have been identified and studied as important contributors to early brain injury such as DCI after aSAH. Preclinical models have identified a biological mechanism that contributes to a unique type of cell death called "ferroptosis" which is nonapoptotic in form but is dependent upon excessive intracellular iron levels. Excessive iron contributes to ferroptosis through the accumulation of lethal, lipid-based reactive oxygen species.[60,61] Specifically, increased levels of iron can produce superoxides which cause oxidative degradation of lipids that make up cell membranes resulting in damage to cell membranes, disruption of cell function, and cell death by ferroptosis.[61] Normally, iron is tightly bound to carrier proteins as ferric iron and is homeostatically controlled within the body resulting in generally stable plasma iron concentrations.[23] However, after aSAH, heme within the subarachnoid space is broken down into carbon monoxide, biliverdin, and a non–protein-bound ferrous iron (i.e., free iron).[24] Importantly, free iron is capable of accepting and donating electrons readily and has high potential for producing superoxides and ferroptosis as discussed above. Free iron and related homeostatic mechanisms have been shown to be important to outcomes following neurological injury. For example, free iron and cell death by ferroptosis have been shown to result in secondary ischemic brain injury (similar to DCI) in animals following experimental subarachnoid hemorrhage[6,7] and iron chelators have been found to effectively reduce neuronal cell death associated with ischemic injury in various preclinical models of neurologic injury.[7,15–17] In addition, evidence from protein biomarker studies of human research participants suggests a relationship between free iron and unfavorable outcomes following aSAH, as well as a protective role of high cerebrospinal fluid (CSF) levels of proteins involved in iron homeostasis including ceruloplasmin[8] and ferritin.[29,30] Finally, in our own work, we have identified associations between

genetic variability of iron homeostasis candidate genes and poor long-term outcomes following aSAH.[50]

Given the research described above, we believe that DNA methylation of genes in the iron homeostasis pathway may also be an important factor in patient outcomes after aSAH. DNA methylation plays an important role in gene expression and has a substantial impact on adult brain function, plasticity, and injury recovery and is thought to change rapidly during the cell cycle.[31] Importantly, DNA methylation is potentially alterable and this dynamic, biological mechanism could be an important biomarker of patient outcomes after aSAH and therapeutic target for intervention to improve patient outcomes. While little DNA methylation research has been performed in the aSAH population specifically, blood DNA hypomethylation of a candidate region of the genome has been associated with increased mortality after ischemic stroke[33] and global DNA hypomethylation in damaged brain tissue has been observed post-injury in rats after traumatic brain injury.[34] Existing studies of DNA methylation after neurologic injury have been largely cross-sectional and reported in blood and have not examined tissues more proximal to the central nervous system such as CSF. DNA methylation is tissue- and time-sensitive and examining DNA methylation in CSF over the acute phase period following aSAH may uncover important evidence about the pathology of aSAH recovery and poor outcomes. Notably, as part of our pilot work for this study, we developed a method to examine DNA methylation over time following aSAH and identified suggestive relationships between DNA methylation trajectory groups of an iron homeostasis candidate gene, hepcidin (*HAMP*), known as the 'master' iron regulator, and patient outcomes following aSAH.[32]

The identification of DNA methylation sites that vary over time and are associated with outcomes after aSAH may provide insight into biological mechanisms of poor outcomes or may

themselves be important biomarkers for identifying subpopulations of patients that require more intensive nursing management to facilitate optimal stroke care delivery and improve patient outcomes. In addition to understanding associations with acute outcomes, it is important to know if variability in DNA methylation from iron homeostasis candidate genes is associated with long-term outcomes either directly or indirectly via an unmeasured acute complication of vascular or microvascular dysfunction. Therefore, the purpose of this study was to investigate associations between DNA methylation trajectory groups of CpG sites in candidate genes related to iron homeostasis and acute and long-term patient outcomes following aSAH.

3.3 Methods

3.3.1 Study Design

This study was a longitudinal, observational study that capitalized on an existing cohort of aSAH participants with extensive phenotype data, stored biosamples, and genome-wide DNA methylation data for a subset of participants. This study assessed the relationship between candidate gene DNA methylation trajectories and patient outcomes acutely (days 0 to 14 post-aSAH) and in the long-term (at 3 and 12 months post-aSAH) using a two phase design (targeted discovery and replication). As described in Section 2.0 of this document, the targeted discovery analyses for this study were performed, top hits were prioritized for replication, and DNA from an independent (i.e., non-overlapping) test sample were sent to the Center for Inherited Disease Research (CIDR) laboratory (Johns Hopkins University, Baltimore, MD) for pyrosequencing to replicate the presented findings. Because of a temporary shutdown in the CIDR laboratory related

to COVID-19, replication data have not yet been returned. Therefore, the replication data collection, analysis, and results have not been included as part of this dissertation document. Before this manuscript is submitted to a peer-reviewed journal for publication, the methodology and results will be expanded to include details of replication.

3.3.2 Setting and Sample

This study adhered to all ethical considerations and was approved by the Institutional Review Board of the University of Pittsburgh. Following informed consent, participants were prospectively recruited from UPMC Presbyterian neurovascular intensive care unit in Pittsburgh, Pennsylvania between 2000 and 2013 if they (1) were aged 18 years and older; (2) were newly diagnosed (≤5 days) with aSAH verified with cerebral angiogram; (3) had a Hunt and Hess grade ≥ 2 and/or Fisher grade ≥ 3; (4) were able to read/speak English; and (5) had no history of debilitating neurological disorders. Additional inclusion criteria for this ancillary study included ventriculostomy insertion as part of clinical management to supply CSF samples and availability of serial CSF samples across 14 days post-aSAH. The parent project collected a variety of data including demographic, confounding, acute and long-term outcomes, and genome-wide DNA methylation data as well as stored CSF biosamples.

3.3.3 DNA Methylation Data

3.3.3.1 DNA Methylation Data Collection (Tier 1 Targeted Discovery)

This study capitalized on existing, longitudinal, genome-wide DNA methylation data collected for 273 participants at up to five time points over 14 days following aSAH as previously described (https://www.biorxiv.org/content/10.1101/2020.03.24.005264v1.full).[62] DNA methylation data were collected from DNA extracted from bagged CSF collected as standard treatment and changed daily for 14 days following aSAH by trained study staff using sterile procedures. The collected CSF samples were centrifuged, and the cellular pellet and supernatant were stored at -80º until DNA extraction. DNA was extracted from the cellular pellet using the Qiamp DNA extraction kit from Qiagen Corp (Qiagen, Valencia, CA, USA) and bisulfite conversion was performed. Genome-wide DNA methylation data were generated and scanned using the Infinium Human Methylation450 Beadchip and Illumina iSCAN (Illumina, Incorporated, San Diego, CA, USA) at CIDR. As part of our laboratory quality control (QC) procedures, samples from a participant were run on the same chip and within the same batch and, beyond that, samples were randomly distributed across the batches. In addition, technical replicates and DNA methylation control samples were included to assess the reliability of data. Raw genome-wide DNA methylation data were analyzed using Genome Studio Software (Illumina, Incorporated, San Diego, CA, USA).

3.3.3.2 DNA Methylation Data Quality Control

Our data cleaning and QC pipeline was implemented using R packages minfi[58] and ENmix.[63] This pipeline included removal of poorly performing and outlying samples based on bisulphite control intensities and detection p-values and background and dye bias correction to remove non-specific

signals from the DNA methylation data.[63,64] To further reduce technical variation related to Infinium 450K platform chemistry (i.e. differences in Type I and Type II probes) and batch effects (i.e., chip, row, and column effects), functional normalization was performed using the 'preprocessfunnorm' function from the minfi package.[57,58] Cell-type heterogeneity (CTH) data were generated as estimates of five cell types using Houseman's reference-free method.[47] Additional details of our QC pipeline are summarized elsewhere.[62]

3.3.3.3 Candidate Gene DNA Methylation Data Extraction

Candidate genes (n=39) were selected based on their known biological roles in iron homeostasis and are presented in Appendix Table 2 and described in detail as part of our previous, related work.[32,50] DNA methylation data for our candidate genes were extracted from the gene transcript region ± 2,000 base pairs upstream and downstream. Data for two genes (*HEPH* and *PGRMC1*) were not available in the cleaned genome-wide data and data for the hepcidin gene (*HAMP*) were analyzed in the pilot work to develop the protocols for this study.[32]

3.3.4 Patient Outcomes

Acute outcome measures included cerebral vasospasm (CV) and delayed cerebral ischemia (DCI) occurring within a 14 day window following aSAH. For this study, CV was defined as ≥25% cerebral vessel narrowing measured by a neurosurgeon via cerebral angiogram[32] and DCI was defined as the co-occurrence of neurological deterioration (e.g., an increase of ≥2 points on the National Institutes of Health Stroke Scale, new and persistent neurological deficit) and abnormal cerebral blood flow measured using cerebral angiogram or transcranial Doppler.[32]

Long-term outcome measures for this study included the Glasgow Outcome Scale (GOS) and death at 3 and 12 months following aSAH. These data were collected by trained study staff via in person or telephone interviews. The GOS is an indication of participants' ability to function on a scale of 1 (death) to 5 (good recovery) and has established validity in people with neurological injury.[65] Death data were obtained from the medical record, caregiver/family report, or the Social Security Death Index. In cases where participants were unable to participate in the interview, a caregiver or family member knowledgeable about the participants ability to function was interviewed. All study staff involved in outcome data collection were blinded to DNA methylation status.

Finally, this study capitalized on existing confounding/covariate data extracted from the medical record or collected as part of a larger research study. These data included demographic data such as age, sex, and race; social data such as marital status and education; and clinical data such as treatment/intervention and severity of injury as measured using the Fisher grade.

3.3.5 Statistical Analysis

3.3.5.1 Descriptive Statistical Evaluation

All statistical analyses were conducted using R[66] and SAS (SAS Institute Incorporated, Cary, NC, USA). Acute outcomes of CV and DCI were treated as binary (occurrence during the 14 day study window versus no occurrence during the 14 day study window). In the long-term, GOS was dichotomized as good (scores of 4 to 5) or poor (scores of 1 to 3) and death was treated as binary (yes versus no) at the outcome timepoint of interest. For DNA methylation data, M values were used and participants with only one DNA methylation data observation were removed as this precluded trajectory analysis.

Standard descriptive statistics were computed in R for all independent, dependent, and potentially confounding/covariate data given the variable's level of measurement (e.g., means and standard deviations for continuous type normally distributed variables, frequency counts and percentages for categorical variables). Group comparisons were performed using t-tests to evaluate equality of means for continuous type variables and Pearson's chi-square test of independence or Fisher's exact test for equality of proportions for categorical variables. Preliminary analyses were conducted to identify potential confounders/covariates. Known predictors of outcomes after aSAH include degree of hemorrhage (as measured by the clinical grading scale, Fisher grade), age, race, and sex.[42] Other potentially relevant confounders included alternative measures of severity of injury or degree of hemorrhage (e.g., Hunt and Hess [HH] Score, World Federation of Neurosurgical Societies [WFNS] Grade), smoking status, body mass index, comorbidities, education, and race.

Data screening procedures were performed in R given the variable's level of measurement. Data were screened for accuracy using frequency tables for categorical variables (e.g., sex, race) and summary statistics (mean, standard deviation, minimum, and maximum) for continuous variables (e.g., age, DNA methylation). To identify outliers, data were examined in frequency tables for categorical variables and in histograms and sina plots for continuous variables. If potential outliers were detected, data were inspected for accuracy and the influence of data were evaluated in a sensitivity analysis. For DNA methylation data, any DNA methylation value labeled as an extreme outlier (a DNA methylation value above or below three times the interquartile range) was replaced with the maximum or minimum observed DNA methylation value below the extreme outlier threshold for values on each day. Data were screened for independence using scatterplots, multicollinearity using tolerance, variance inflation factors, and condition indices with variance

decomposition proportions, and normality using histograms with a normal curve overlay and Kolmogorov-Smirnov test of normality. Missing data were assessed and filled in from the medical record when possible.

3.3.5.2 Group-based Trajectory Analyses

Group-based trajectory analysis (GBTA) for DNA methylation data was conducted in SAS using Proc TRAJ assuming a censored normal model, which helps correct for right censored data within longitudinal studies. For individual CpG sites within our candidate gene DNA methylation data, the change pattern over time was examined to infer distinct trajectory patterns/groups of site-specific methylation across time over the acute recovery period post-aSAH. In GBTA, models with a varying number of groups and polynomial orders (group trajectory shapes) are compared to find the model that best fits the longitudinal data.[44] As part of this modeling process, participants are assigned a posterior probability of assignment for each trajectory group in a model; each participant is then assigned to whichever group they have the highest probability of being in.[44] Given the subjectivity required in traditional GBTA modeling and the large number of candidate genes and CpG sites analyzed in this study, the model selection process was largely automated for 39 possible models with a maximum of three groups and comprehensive combinations of polynomial orders of 0 (intercept only), 1 (linear), and 2 (quadratic).

Our GBTA automated protocol has been described in detail as part of our pilot work.[32] First, out of the 39 possible models, a best fitting 'preliminary' model was identified based on the maximum Bayesian Information Criterion (BIC). In general, BIC measures improvement in model fit gained with the estimation of more parameters such as an increased number of groups and/or more complex trajectory shapes, but also applies a penalty for model complexity. In SAS, a larger BIC value indicates a better model fit. Using GBTA in SAS, two BIC values were computed based

on (1) the number of participants and (2) the total number of methylation observations over time. The true BIC falls somewhere within these values.[44] In cases where the BICs did not agree on the best preliminary model, we assigned a 'simplicity score' to assist in identifying the more parsimonious model which was carried forward. Following selection of a preliminary model for each CpG site, we performed a secondary evaluation of model adequacy (i.e., posterior model QC) using several traditional diagnostics including (1) an average posterior probability (AvePP) >0.7, (2) odds of correct classification (OCC) >5, (3) estimated group membership (π) >5%, (4) reasonably close estimated group membership (π) versus the assigned group proportion (P*), and (5) a relatively narrow 95% confidence interval for the estimated group probability (π).[45] For CpG sites which preliminary models failed secondary evaluation, the 'next best' fitting model was selected and secondary evaluation was performed again. If preliminary models for a CpG site failed a second time, we concluded that DNA methylation trajectory groups could not be inferred with high accuracy for that site and it was excluded from further analysis.[32]

3.3.5.3 Cell-type Heterogeneity

Cell-type heterogeneity (CTH) can be an important confounder in DNA methylation studies as cell-type differences within biosamples can impact the overall DNA methylation level. If genome-wide DNA methylation data are collected, it is possible to control for CTH.[67] Given the expense of generating longitudinal, genome-wide DNA methylation data it is currently not practical to generate genome-wide data for a clinical biomarker. Therefore, it might be ideal to identify a biomarker that is robust to confounding cell-type. To examine the potential clinical utility of DNA methylation trajectories unadjusted for CTH as well as evaluate potential confounding by cell-type, we implemented our GBTA protocol twice for all CpG sites to compute DNA methylation trajectories both unadjusted for CTH and adjusted for CTH.[32] As part of the data cleaning and QC

pipeline discussed above, CTH data were generated using Houseman's reference-free method.[47] These data include percentages of five cell-types for each biosample and were controlled for as time-varying covariates during GBTA in our CTH-adjusted models.

3.3.5.4 Patient Outcome Association Analyses

Finally, binary logistic regression was performed in R to determine the relationship between inferred trajectory groups for each DNA methylation site and patient outcomes while controlling for age, sex, race, and Fisher grade. A likelihood ratio test was used to produce a global p-value of the overall model fit by comparing the full model (including the CpG site) with a restricted model (omitting the CpG site). Given the correlation between patient outcomes in this study, permutation testing, rather than strict Bonferroni corrections, was used to correct for testing correlated outcomes.[32]

3.3.5.5 Sensitivity to Functional Normalization

A range of preprocessing methods exists to clean and prepare DNA methylation data for analysis. One such preprocessing method that is an important step in DNA methylation data analysis is known as data normalization. This is critical in processing Illumina DNA methylation data as different probe designs use different chemistry and possess different beta distributions. As described above, the targeted discovery data available for this study were normalized using functional normalization as implemented in the preprocessFunnorm function.[57] However, we later identified a similar method known as 'funtooNorm' which was designed to better handle data gathered across time points and allows for interactions between tissue-types.[59] Because the aSAH DNA methylation data were generated across time in CSF, a tissue in which we expect (1) DNA methylation profiles to differ as a function of time post-injury and (2) an array of dynamic cell-

types, we wanted to understand how our choice of functional normalization impacted our results. In an effort to explore this question, we repeated the aforementioned procedures in a sensitivity analysis using funtooNorm normalized data.

3.3.5.6 Prioritization of Top Hits

Following gene-specific data screening and analysis, top hits for replication were prioritized based on the strength of the identified associations, consistency of results after adjustment for CTH, and presence of hotspots (i.e., multiple CpG sites near each other associated with outcomes) as outlined in Figure 1. As discussed above, DNA was extracted from an independent test sample of aSAH participants and sent for pyrosequencing at CIDR. However, those replication data have not yet been returned for analysis. Prior to publication of the paper based on Section 3.0 of this document in a peer reviewed journal, these replication data will be analyzed, and the methods and results will be expanded to include the replication data in this manuscript. Given the breadth of this study and pending replication data, the results below focus on our top three hits with pending replication data. However, detailed results for all 36 candidate genes analyzed are presented in the Supplementary Material.

3.4 Results

3.4.1 Descriptive Statistical Evaluation

Longitudinal genome-wide DNA methylation data from days zero through fourteen following aSAH were available for 273 participants. As part of data screening, there were very few

observations available on days 0 and 14, so these time points were dropped from our analyses. Next, 13 participants with only one DNA methylation measurement were removed. Our final sample size consisted of 260 participants and the sample characteristics are presented in Appendix Table 3. Our sample had a mean age of 53.1 +/- 11.0 years, was 68.8% female, and 86.5% self-reported their race as White. Outcome distribution for our sample is presented in Appendix Table 4 and included CV in 54.2%, DCI in 49.2%, poor GOS at 3 and 12 months in 33.2% and 26.0%, respectively, and death at 3 and 12 months in 16.8 and 21.6%, respectively. Participants had between 2 and 5 DNA methylation observations with an average of 3.2 observations. The number of DNA methylation observations per day ranged from 32 on day 12 to 103 on day 5.

In bivariate analyses, older age was associated with DCI (p=0.04), poor GOS at 3 and 12 months (p=0.03 and p=0.04, respectively), and death at 3 and 12 months (p=0.02 and p=0.04). Non-White race was associated with poor GOS at 3 months (p=0.04) and higher Fisher grade was associated with the occurrence of CV (p=0.01) and DCI (p=0.03) as well as poor GOS and death measured at 3 and 12 months (all p<0.001). No associations between sex and patient outcomes or intervention (surgical vs. coil embolization) and patient outcomes were observed. However, sex was included in all logistic regression models given the importance of estrogen response elements in iron homeostasis.[68]

From the genome-wide data, DNA methylation data for 36 candidate genes were available and analyzed as part of this study (Appendix Table 2). A total of 637 CpG sites passed genome-wide QC procedures and were included. During data screening, DNA methylation observations that were identified as extreme outliers >3 times the IQR were pulled in as described above; gene-specific outlier screening and score adjustment is presented in Appendix Table 5. Of our 637 CpG sites, our objective protocol eliminated 412 sites as trajectory groups could not be inferred with

high accuracy for those sites, primarily due to an OCC< 5. At an additional 42 sites, we inferred only one trajectory group. After elimination of these sites from our analysis, 183 CpG sites in 33 of our candidate genes were carried forward for patient outcome association testing as summarized in the Supplementary Material.

3.4.2 Prioritization of Top Hits

Our budget limited the number of CpG sites that could be sent for replication with funds allowing for replication data collection at three sites. Following gene-specific data screening and analysis, top hits for replication were prioritized based on several criterion including the strength of the identified associations, consistency of results after adjustment for CTH, and presence of hotspots (i.e., multiple CpG sites near each other associated with outcomes) as outlined in Appendix Table 6, Appendix Figure 1, and Appendix Table 7. Three CpG sites in three candidate genes were flagged as noteworthy of replication based on this criteria. These sites included cg08866780 in the Amyloid Precursor Protein (*APP*), cg25713625 in STEAP3 Metalloreductase (*STEAP3*), and cg08553327 in Tumor Necrosis Factor (*TNF*). A summary of results for all associations with a *p*-value < 0.05 are presented in Appendix Table 6 and detailed results for these three CpG sites are presented below and detailed results for all 36 candidate genes analyzed are presented in the Supplementary Material.

3.4.3 Group-based Trajectory and Patient Outcome Association Analyses

Both unadjusted and CTH-adjusted GBTA was performed and trajectory plots for our top three hits in *APP*, *STEAP3*, and *TNF* are presented in Appendix Figure 2 and participant characteristics

by trajectory group are presented in Appendix Table 8. It should be noted here that the trajectory plots unadjusted for CTH are not directly comparable to the trajectory plots adjusted for CTH as group membership changes (e.g., for *APP* in our plot unadjusted for CTH, Group 1 has 63 participants while in our CTH-adjusted plot, Group 1 has 67 participants as shown in Appendix Table 8). Binary logistic regression results exploring associations of inferred DNA methylation trajectory groups with patient outcomes are presented for our CpG sites of interest in *APP*, *STEAP3*, and *TNF* in Appendix Tables 9, 10, and 11, respectively.

For cg08866780 (*APP*), in the unadjusted model we inferred three flat trajectory groups including a low DNA methylation group (Group 1, 24.2%), intermediate DNA methylation group (Group 2, 53.5%), and high DNA methylation group (Group 3, 22.3%). Interestingly, even after correcting for CTH, we observed very similar trajectory patterns at this CpG site (Appendix Figure 2). Participant characteristics by trajectory group are presented in Appendix Table 8. For the model unadjusted for CTH, characteristics were quite similar across the groups with the exception of intervention. In the low DNA methylation group (Group 1), 73% of participants received a coil embolization while only roughly 55% received a coil embolization in Groups 2 and 3. Results of binary logistic regression are presented in Appendix Table 9 and observed outcome distributions for the trajectory groups unadjusted for CTH are presented in Appendix Table 12. In association analyses, significant associations were identified with favorable GOS at 12 months (p=0.0001) and survival at 3 and 12 months (p=0.001 and p=0.0006, respectively) and a suggestive association was identified with GOS at 3 months (p=0.006). Overall, we observed that participants in the intermediate DNA methylation group (Group 2) had a decreased odds of poor outcomes compared with low and high DNA methylation groups (Groups 1 and 2, respectively) as summarized in Appendix Table 7. These results were consistent even after correction for CTH with suggestive

associations identified between inferred trajectory groups at cg08866780 and favorable GOS at 3 and 12 months (p=0.004 and p=0.01, respectively) and survival at 3 and 12 months (p=0.02 and p=0.01, respectively). No associations between inferred trajectory groups at cg08866780 and acute outcomes of CV or DCI were observed.

For cg25713625 (*STEAP3*), in the unadjusted model, we again inferred three trajectory groups including a low DNA methylation group (Group 1, 11.1%), intermediate DNA methylation group (Group 2, 63.8%), and high DNA methylation group (25.0%). While mean DNA methylation in the low and high DNA methylation groups were steady over time with no change, the mean DNA methylation level of the intermediate DNA methylation group steadily decreased over time. After controlling for CTH, we observed very similar trajectories, but all three groups were flat with no change over time (Appendix Figure 2). Participant characteristics by trajectory group are presented in Appendix Table 8. For the model unadjusted for CTH, we observed that the low DNA methylation group (Group1) had a lower percentage of female participants (44.8% in Group 1 vs. 71.1% and 73.8% in Groups 2 and 3, respectively), a lower percentage of White participants (69% in Group 1 vs. 88.6% and 89.2% in Groups 2 and 3, respectively), and larger bleeds as evidenced by Fisher Grade distributions (Fisher Grade 2 of 20.7% in Group 1 vs. 32.5% and 27.7% in Groups 2 and 3, respectively). Results of binary logistic regression are presented in Appendix Table 10 and observed outcome distributions for the trajectory groups unadjusted for CTH are presented in Appendix Table 12. In association analyses, significant associations were identified with poor GOS at 3 and 12 months (p=0.0006 and p=0.002, respectively) and suggestive associations were identified with death at 3 and 12 months (p=0.009 and p=0.004, respectively). Overall, we observed that the high DNA methylation group (Group 3) had an increased odds of poor outcomes compared with the low DNA methylation group (Group 1) as summarized in

Appendix Table 7. After correction for CTH, suggestive associations were again observed with GOS at 3 and 12 months (p=0.03 and p=0.02, respectively). No associations between inferred trajectory groups at cg25713625 and acute outcomes of CV or DCI were observed.

Finally, for cg08553327 (*TNF*), in the unadjusted model, we again inferred three trajectory groups including a low DNA methylation group (Group 1, 67.3%), intermediate DNA methylation group (Group 2, 26.9%), and high DNA methylation group (Group 3, 5.8%). In contrast to the CpG sites discussed above, DNA methylation at this site was quite dynamic after aSAH. All three trajectory groups started and ended with roughly the same DNA methylation values, but different patterns were observed over the 13 days post-aSAH as shown in Appendix Figure 2. The low DNA methylation group (Group 1) decreased and then increased slightly; the intermediate DNA methylation group (Group 2) steadily decreased over time; and the high DNA methylation group (Group 3) steadily increased and then decreased below baseline. As shown in the Supplementary Material and summarized in Appendix Table 7, in *TNF*, there was a hotspot of several CpG sites that were significantly associated with patient outcomes after aSAH and had similar trajectory shapes. CpG site cg08553327 was specifically chosen for replication because it has the smallest p-values in the gene and was amenable to pyrosequencing at CIDR. While the CTH-adjusted trajectory model at cg08553327 did not pass posterior model QC, two sites within the hotspot did and resulted in three steadily decreasing trajectory groups. At these sites, after correction for CTH, suggestive associations persisted between inferred trajectory groups and GOS at 3 months as shown in the Supplementary Material. Participant characteristics by trajectory group are presented in Appendix Table 8. At cg08553327, we observed that the high DNA methylation group (Group 3) had a higher percentage of female participants (80% in Group 3 vs. 68% and 68.6% in Groups 1 and 2, respectively), a higher percentage of White participants (93.3% in Group 3 vs. 86.9% and

84.3% in Groups 1 and 2, respectively), and smaller bleeds as evidenced by Fisher Grade distributions (Fisher Grade 2 of 40% in Group 3 vs. 30% and 29.1% in Groups 1 and 2, respectively). Results of binary logistic regression are presented in Appendix Table 11 and observed outcome distributions for the trajectory groups unadjusted for CTH are presented in Appendix Table 12. In association analyses, significant associations were identified with poor GOS at 3 months (p=0.0003) and suggestive associations were identified with poor GOS at 12 months (p=0.02) and death at 3 and 12 months (p=0.009 and p=0.004, respectively). Overall, we observed that participants in both the intermediate (Group 2) and high DNA methylation (Group 3) groups had an increased odds of poor outcomes as summarized in Appendix Table 7. As mentioned above, the trajectory model corrected for CTH did not pass posterior model QC for comparison. No associations between inferred trajectory groups at cg08553327 and acute outcomes of CV or DCI were observed.

3.4.4 Sensitivity Analysis of Functional Normalization

In a sensitivity analysis of *APP*, *STEAP3*, and *TNF* using the funtooNorm functional normalized data (versus the preprocessFunnorm functional normalized data), our results were largely concordant as shown in the Supplementary Material. For cg08866780 (*APP*), we observed nearly identical associations with GOS and death at 3 and 12 months, although none of the observed associations met the empirical significance threshold computed in permutation testing. Likewise, for cg08553327, we also continued to observe associations with GOS and death at 3 and 12 months and our association with GOS at 3 months remained significant after correction for multiple testing. In contrast, however, cg25713625 (*STEAP3*) did not pass posterior model QC in our GBTA so we were unable to carry it forward for evaluation with patient outcomes. Of the CpG

sites that did pass posterior model QC though, the results were very similar with the original analysis (i.e., preprocessFunnorm data).

3.5 Discussion

This study was the first to examine associations between DNA methylation trajectories of iron homeostasis candidate genes and patient outcomes after aSAH. The candidate genes for this study were selected based on characteristics and mechanisms that were relevant to our scientific premise and phenotypes of interest during recovery from aSAH. Of the candidates examined, associations from three genes stand out in our results during recovery from aSAH.

3.5.1 cg08866780 (*APP*)

Specifically, one CpG site in *APP*, cg08866780 (UCSC Genome Broswer GRCh37/hg 19, chr21: 27543523), was significantly associated with GOS at 12 months and death at 3 and 12 months and suggestively associated with GOS at 3 months (Appendix Table 9). Amyloid Precursor Protein, the protein encoded for by *APP*, is highly expressed in the human brain and ubiquitously expressed across many human tissues.[69,70] Although this gene has been extensively studied since its identification, its physiological function remains largely unknown and *APP* has been referred to as a 'biochemical enigma'.[70] However, it is broadly accepted that *APP* plays a role in nervous system development and recovery, synaptic plasticity, calcium metabolism, learning, memory, and as it is most well-known, giving rise to the amyloid-β peptide and contributing to the pathogenesis of Alzheimer's disease.[71] Specific to iron homeostasis, the role of *APP* in managing iron levels is

feedback-regulated as iron influx is an important driver in translational expression of neuronal *APP* via an iron responsive element (IRE) and, in turn, *APP* plays a role in iron efflux by stabilizing the heavy subunit of ferritin and ferroportin-1 at the binding site.[69]

The CpG site in *APP* that rose to the top in the current study, cg08866780, was recently identified as a differentially methylated CpG site in Braak staging, a method used to classify the degree of Alzheimer's disease.[71] Specifically, cg08866780 was identified as being hypomethylated, independent of chronological aging, during Braak stage progressions.[71] It should be noted that an association between this CpG site and Braak staging was not identified in historical studies; rather, the signal became apparent only in neurons and glia cells after purifying cell-types (i.e., cell sorting).[71] This CpG site is located in the *APP* promoter region and overlaps with a known CTCF binding region. CTCF is a regulatory protein with roles as an insulator protein and regulator of chromatin structure and transcription.[72] Hypomethylation of this region of *APP* results in enhanced binding of CTCF and an increase in *APP* transcription.[73] In Alzheimer's disease, an increase in *APP* production results in an overproduction and accumulation of Amyloid-β and Alzheimer's disease progression.[71] While the pathophysiology of Alzheimer's disease progression and recovery after aSAH is very different, recent literature has uncovered the role of iron in amyloid plaque formation in an Alzheimer's disease mouse model,[74] underscoring the interplay between iron homeostasis and lipid metabolism which is discussed in more detail below. While *APP* excess is thought to have detrimental effects in the brain, it has also been suggested that *APP* depletion can be harmful. For example, in a study of non-small cell lung cancer cells, *APP* depletion in lung cells caused G(0)/G(1) phase cell-cycle arrest contributing to necrotic cell death.[75]

In our sample, participants that fell into the intermediate DNA methylation trajectory group (Group 2) had lower odds of poor outcomes compared with the low DNA methylation group (Group 1) while there was no difference observed between the high DNA methylation group (Group 3) and the low DNA methylation group (Group 1). We confirmed that the intermediate DNA methylation group (Group 2) also did better compared with the high DNA methylation group (Group 3) suggesting that high or low DNA methylation at cg08866780 (*APP*) may be detrimental during recovery. Based on these observations, in conjunction with the feedback-regulated role for *APP* in iron homeostasis, we could not determine a direction of effect in relationship to ferroptotic activity for this site. It should be noted that the trajectory groups identified were flat with no change over time (Appendix Figure 2). If we can assume these trajectories can be extended backwards in time before the aSAH insult, then perhaps the *APP* mechanism at play is not related to an influx of free iron or aSAH insult itself, but rather a preexisting physiologic state (e.g., genetics). It should be noted that if we hadn't used GBTA to classify participants based on DNA methylation over time, we would have missed this association. In other words, in cross-sectional analyses of the associations between continuous DNA methylation at this site (ignoring trajectory groups) and patient outcomes, there is no association as the high and low DNA methylation groups (Groups 3 and 1, respectively) did poorly compared with the intermediate DNA methylation group.

Of the 21 CpG sites available post-QC in this gene, we were able to infer trajectory groups at 8 CpG sites and significant or suggestive associations with patient outcomes were observed in 4 CpG sites with similar directions of effect (Appendix Table 6 and Supplementary Material). An interesting finding in this study was that even after correction for CTH, the inferred DNA methylation trajectory groups looked nearly identical to the unadjusted trajectory groups (Appendix Figure 2), indicating that CSF CTH might not be an important confounder of DNA

methylation at this site after aSAH. This finding contrasts with the finding discussed above related to differential DNA methylation at this CpG site in Braak staging as this CpG site only rose to the top in purified glia and neuron cells.[71]

3.5.2 cg25713625 (*STEAP3*)

Next, one CpG site in *STEAP3*, cg25713625 (UCSC Genome Broswer GRCh37/hg 19, chr2: 120022835), was significantly associated with GOS at 3 and 12 months and death at 3 and 12 months (Appendix Table 10). Metalloreductase STEAP3, also known as six-transmembrane epithelial antigen of prostate 3, the protein encoded by the *STEAP3* gene, plays a primary role in iron transport and homeostasis by converting the stable but insoluble ferric form of iron to a less stable but soluble ferrous form.[76] It is known that *STEAP3* deficiency leads to impaired iron homeostasis and presents clinically as microcytic anemia with iron overload.[76] In addition, knockdown of *STEAP3* has been shown to inhibit cell progression suggesting this protein may also play a role in downstream responses to p53, including promoting apoptosis.[77] *STEAP3* is also thought to be important in immunity. Iron deficiency confers resistance to risk of infection and *STEAP3*, the only STEAP family member to be highly expressed in macrophages, has been shown to be down-regulated by lipopolysaccharide administered as a surrogate for infection in rats.[76] In addition to its importance in iron homeostasis and immunity, *STEAP3* has been identified as a novel pathway for increased oxidative stress in mature red blood cells leading to changes in lipid oxidation, which we have discussed in more detail below.[78] The CpG site of interest, cg25713625, falls in the 3' untranslated region of *STEAP3*. To our knowledge, there is no significant literature about cg25713625, specifically. Interestingly though, *STEAP3* has been identified as a differentially expressed gene in the aneurysm wall compared with a control tissue of superficial

temporal artery. Specifically, authors observed a 2.5 fold change in gene expression with higher expression of *STEAP3* in aneurysm tissue.[79]

In our sample, at cg25713625, participants who were assigned to the highest DNA methylation trajectory group (Group 3) had the worst long-term outcomes compared with participants assigned to the lowest DNA methylation trajectory group (Group 1) with odds of poor outcomes between 11.7 and 19.1 (Appendix Table 10). If we assume that an increased level of STEAP3 reduces the theoretical amount of ferroptosis, we observed the expected direction of effect with this site. As above with cg08866780 (*APP*), we observed little change over time in the trajectories inferred at this CpG site (Appendix Figure 2). While the highest and lowest DNA methylation trajectory groups (Groups 3 and 1, respectively) did not change over time, the middle group did steadily decrease over time. After correcting for CTH, however, the trajectory groups look nearly identical with the exception that the middle trajectory group flattened (Appendix Figure 2). After correcting for CTH, similar associations were observed with only suggestive or trending p-values (Appendix Table 10). In cross-sectional analyses of the associations between continuous DNA methylation at this site (ignoring trajectory groups) and patient outcomes, our observed associations persisted with higher DNA methylation being associated with an increased odds of poor long-term outcomes.

In *STEAP3*, of the 29 CpG sites we evaluated using GBTA, only three actually passed posterior model QC and were tested for associations with patient outcomes and it is possible that we missed some important variability in this gene. One other site in *STEAP3* (cg25508118) had suggestive associations with GOS and death at 12 months as shown in Appendix Table 6 and the Supplementary Material.

3.5.3 cg08553327 (*TNF*)

We observed a 'hotspot' of associations between several CpG sites in *TNF* and long-term outcomes of GOS and death at 3 and 12 months, with the strongest association with GOS at 3 months. Of the CpG sites located within that hotspot, inferred trajectory groups and group percentages looked nearly identical as shown in the Supplementary Material. From this hotspot, we selected cg08553327 as a candidate for replication because it had the smallest p-values amongst CpG sites and was amenable to pyrosequencing at CIDR. In *TNF*, cg08553327 (UCSC Genome Broswer GRCh37/hg 19, chr6: 31543647) was significantly associated with GOS at 3 months and suggestively associated with GOS at 12 months and death at 3 and 12 months (Appendix Table 11). Tumor Necrosis Factor, the protein encoded by *TNF* is a pro-inflammatory cytokine most commonly recognized for its role in inflammation.[80] Iron homeostasis and inflammation are intimately tied as iron deficiency confers resistance to risk of infection and improves inflammatory conditions (e.g., anemia of inflammation in chronic disease). *TNF* plays an important role in iron homeostasis and host defense through regulating inflammation and hypoferremia.[81] Specifically, increased inflammation causes upregulation of hepcidin, the 'master' iron homeostasis protein, which triggers a series of events that lead to systemic iron sequestration and prevention of iron absorption. This relationship is feedback-regulated as it has also been shown that iron administration increases TNF expression while iron chelators decrease TNF expression.[81]

Our CpG of interest, cg08553327, is located in exon 1 of *TNF* and has been inversely correlated with aging in a study of delirium and aging,[82] although there was no significant difference between mean age in the three trajectory groups inferred at this site in the current study. This CpG site specifically has also been found to be differentially methylated in two separate

studies of fiber intake.[83,84] Moreover, after subarachnoid hemorrhage specifically, anti-TNF antibody (i.e., TNF inhibitor) administered pre-injury in rats has been shown to reduce apoptosis.[85]

In our sample, participants assigned to the high DNA methylation group (Group 3) had 13.1 times the odds of poor GOS at 3 months compared with the low DNA methylation group (Group 1). If we assume that an increased level of TNF increases the theoretical amount of ferroptosis, we did not observe the expected direction of effect with this site. Assuming that higher DNA methylation leads to lower expression, this finding is inconsistent with literature suggesting increased *TNF* is associated with poor outcomes and that administration of a *TNF* inhibitor has been shown to improve outcomes.[85] It is possible that what we are observing is an acute response to aSAH, a difference in behavior between CSF and blood, or confounding by CTH. It should also be noted that higher DNA methylation does not always correlated with alterations in gene expression and *TNF* has many complex mechanisms regulating its expression.[86] The trajectory model for our site of interest in *TNF* did not pass posterior model QC once we corrected for CTH, preventing a direct comparison here. Replication of this finding is critical to interpretation. If we are able to replicate this finding, it is apparent that timing of biospecimen collection could impact findings for this CpG site specifically. Notably, all DNA methylation trajectory groups started and ended with roughly the same DNA methylation values so if we had only performed a cross-sectional study we may not have observed any differential DNA methylation at this site depending on our time of sampling. Our post hoc cross-sectional analysis confirms this. Specifically, in daily cross-sectional analyses, associations between continuous DNA methylation at this site (ignoring trajectory groups) and patient outcomes were only observed on days five to seven. Our observed associations persisted with higher DNA methylation at this site being associated with an increased odds of poor long-term outcomes.

3.5.4 Lipid Metabolism and Endothelial Dysfunction

In addition to iron overload as a primary mechanism contributing to ferroptosis, accumulation of lethal, lipid-based reactive oxygen species have also been found to be important.[60,61] Specifically, elevated free iron post-aSAH has great potential to cause oxidative degradation of lipids that make up cell membranes. The interplay between iron homeostasis and lipid peroxidation after subarachnoid hemorrhage was recently examined in a study of animals.[60] In this study, researchers demonstrated that an increase in iron contributed to an increase in lipid peroxides and that administration of ferrostatin-1, a lipophilic antioxidant which protects cell membranes from lipid oxidation, decreased free iron and improved lipid peroxidation.[60] This intervention successfully prevented both ferroptosis and early brain injury, but had no effect on apoptosis.[60] In addition, the role of lipid peroxidation metabolites (e.g., 20-HETE) has been associated with poor outcomes after aSAH in humans, specifically.[36] Despite this literature, there is little work exploring how the accumulation of these lipid-based ROS are induced by iron in humans after aSAH.

In the current study, CpG sites in *APP*, *STEAP3*, and *TNF* rose to the top as candidates for replication. Interestingly, these genes all play a role not only in iron homeostasis but have also been associated with lipid peroxidation in the literature. For example, as mentioned above, in cases of *APP* overexpression, increased lipid peroxidation has been shown to precede amyloid plaque formation in a mouse model of Alzheimer's disease.[87] Likewise, increased levels of *STEAP3*, though essential to normal iron homeostasis, may result in degradation of the cellular membrane through lipid peroxidation, leading to failure of hemolysis and clearance of red blood cells.[78] Finally, *TNF* has been shown to alter lipid metabolism and stimulate production of eicosanoids and ROS that potentiate CNS injuries.[88] The potential of improving iron homeostasis following aSAH may be essential in preventing cell oxidative damage and warrants future investigation.

3.5.5 Strengths and Limitations

There are many strengths to this study. In particular, the DNA methylation data used in this study were quite unique. In addition to being longitudinal in nature spanning the acute recovery period post-aSAH, these data were also generated from CSF, a tissue which we believe offers a unique insight into the local environmental of the central nervous system. In this study, we were able to identify dynamic changes in DNA methylation following aSAH (though these signals will need to be validated in our pending replication data). Similarly, because genome-wide data existed in our discovery sample, this allowed us to perform additional sensitivity analyses that we could not have performed in a traditional candidate gene study (i.e., examining the effects of CTH and preprocessing methods). Finally, this study analyzed a relatively large sample of aSAH participants compared with other patient outcome studies after aSAH and included imbedded replication and rigorous data QC and analysis.

Despite these strengths, there are some important limitations that should be discussed. First, this study was primarily made up of participants who self-report their race as White. Although the distribution of race in this study is largely consistent with Southwestern PA demographics, this limits the generalizability of our findings in other races, ethnicities, and ancestries. These factors should be examined in larger and more diverse samples in the future. Next, in addition to GBTA, there are several types of trajectory analyses in the literature such as hierarchical modeling and latent curve analysis. If applied to our DNA methylation data, these alternative methods could be valuable in uncovering important signals in our data that may not be apparent using GBTA. In addition, programs to automate GBTA for large datasets and sites of interest are not available. Because of this, we created an objective protocol accompanied by R and SAS scripts to help us automate our analyses as described in detail as part of the pilot work for this study.[32] Given the

nature of our automated protocol, a large portion of CpG sites were eliminated and not carried forward for testing with patient outcomes because we couldn't reliably identify trajectory groups at those CpG sites.

Furthermore, an array of preprocessing methodology is available in the DNA methylation literature. Our targeted discovery data were analyzed, and top hits were prioritized, using the preprocessFunnorm functional normalization discussed above. It is possible, however, that another functional normalization method such as funtooNorm is more appropriate for longitudinal CSF DNA methylation data. A strength of this study, however, was the concordance between our results using the preprocessFunnorm data and funtooNorm data as shown in the Supplementary Material and discussed above. However, if funtooNorm data had been used from the start, it is possible that different CpG sites may have risen to the top for replication. Likewise, our tier 1 analyses were limited to existing genome-wide DNA methylation data. As part of our genome-wide DNA methylation data QC, CpG sites located on the X chromosome were removed. Therefore, data for two of our candidate genes (*HEPH* and *HJV*) were not available for analysis. It is possible that variability in these genes is important to outcomes after aSAH.

Finally, our replication data have not yet been returned from CIDR, so we are unable to validate our findings. However, our paper based on this dissertation chapter will be expanded to include the results of replication analyses once those data are returned. A similar limitation that we cannot address, however, is the inability to collect replication data for all CpG sites that met correction for multiple testing. While we attempted to be as objective as possible in our prioritization of results for replication, true and important signals may exist for other CpG sites. These sites should be examined in the future.

3.6 Conclusion

The results of this study support a role for methylation of CpG sites in *APP*, *STEAP3*, and *TNF* in recovery following aSAH and may offer potential as a biomarker or therapeutic target to improve outcomes. With respect to ferroptotic activity, we observed the hypothesized direction of effect in our CpG site of interest in *STEAP3*, the opposite of the hypothesized direction of effect in our site of interest in *TNF*, and due to the feedback-regulated role of *APP* in iron homeostasis, the direction of effect could not be determined. The results of this study suggest ferroptosis may be an important area of investigation in outcomes after aSAH. If our findings can be replicated, exploration of the functional role of methylation of these CpG sites and genes should be examined in an effort to translate findings clinically.

Appendix A Iron Homeostasis Candidate Genes

Appendix Table 1 Iron Homeostasis Candidate Genes, Functions, and Known Associated Phenotype Information

Gene	Name	Function as related to iron homeostasis	Primary expression	Expression in brain (Mean RPKM)	Posited effect on plasma free iron if gene expression upregulated	Known associated phenotype (Genome Home Reference Database)[1]
ACO1	Iron regulatory protein 1	Aconitase family; posttranscriptional regulation of iron metabolism genes[2]; if cellular iron levels are high, ACO1 binds to 4S-FS cluster and catalyzes the conversion of citrate to isocitrate; if cellular levels are low, binds to iron-responsive elements (structures in the UTR of ferritin and transferrin receptor mRNA); 'moonlighting protein' because it performs multiple distinct functions; multiple transcript variants[3]; dissociate from IRE upon increase in intracellular iron[3,4]	Fat	7.496[3]	//	//
ACO2	Iron regulatory protein 2	Aconitase family; iron sensing regulator of mitochondrial respiration and erythropoiesis[5]; catalyzes the conversion of citrate to isocitrate in TCA cycle;[6] protein levels of ACO2 were not significantly affected by iron restriction indicating posttranslational mechanism rather than IRE-mediated control[7]; dissociate from IRE upon increase in intracellular iron[4]	Heart	61.318[6]	//	//
APP	Amyloid Beta Precursor Protein	Facilitates iron export by stabilizing heavy subunit of ferritin and ferroportin-1 at the hepcidin-binding site[4]	Brain	395.222[8]	Increase in plasma iron (decrease in tissue/cellular iron)	Alzheimer disease; hereditary cerebral amyloid angiopathy (loss of cognitive function)
CALR	Calreticulin	Upregulated in response to iron overload and acts as oxidative damage protection[9]	Thyroid	98.571[10]	//	19p13.13 deletion syndrome (delayed intellectual and physical function, seizures, abnormalities); Essential thrombocythemia (increased number of platelets); Primary myelofibrosis (scar tissue of bone marrow)
CD163	Hemoglobin scavenger receptor	Hemoglobin-haptoglobin complex clearance[11]	Appendix	8.819[12]	//	//
CP	Ceruloplasmin	Ferroxidase (catalyzes oxidation of Fe^{2+} to Fe^{3+}); copper transport[13]; decreased plasma ceruloplasmin causes hypoferremia even when total body iron stores are normal;	Liver	4.313[14]	Increase in plasma iron (decrease in tissue iron)	Aceruloplasminemia (hypoferremia with tissue iron loading)
CUBN	Cubilin	Glomerulus transferrin and hemoglobin reabsorption[15]; megalin-cubilin complex takes up non-protein bound iron and reduces and reabsorbs as ferrous iron Fe^{2+} which in turns causes and increase in cellular iron and a decrease in plasma iron[16]	Kidney	NA[17]	Decrease in plasma iron (increase tissue/cellular iron)	Imerslund-Grasbeck syndrome (low vitamin B12 and megaloblastic anemia)
CYBRD1	Duodenal cytochrome b	Ferrireductase (catalyzes reduction of Fe^{3+} to Fe^{2+})[18]; because there is no regulated excretion of iron, CYBRD1 plays a role in ascorbate-dependent reduction of non-heme iron in the gut (i.e., non-absorption); if ascorbate is elevated, non-heme iron absorption is promoted[19]	Thyroid	10.688[20]	//	--

FECH	Ferrochelatase	Protein is localized to the mitochondrion; catalyzes ferrous iron into protoporphyrin in mitochondrial heme synthesis[2,21]	Bone marrow	3.931[22]	//	Protoporphyria (abnormal heme production and toxic levels of porphyrins)
FLVCR1	Feline leukemia virus subgroup C receptor	Heme exporter requiring the presence of HPX; required to maintain free heme balance; protective against heme toxicity[23]	Small intestine	1.615[24]	//	Posterior column ataxia with retinitis pigmentosa (PCARP) (neurodegenerative syndrome with vision loss)
FTH1	Ferritin heavy	Encodes for the heavy subunit of ferritin; variation in ferritin subunit composition may impact homeostasis;[25] Ferroxidase activity and ferritin stability[26]	Colon/ Appendix	671.024[25]	//	Hemochromatosis Type 5 (Iron loading)
FTL	Ferritin light	Promotes ion nucleation and increases ferritin stability[26]	Kidney	636.433[27]	//	Neuroferritinopathy (brain iron overload); hyperferritinemia (cataracts)
FTMT	Mitochondrial ferritin	Iron-storage protein in mitochondria that is thought to protect against oxidative damage from excess free iron[28]	Testis/brain[28]	--	//	--
FXN	Frataxin	Mitochondrial iron transport regulation[29]	Adrenal/ Bone Marrow/ Liver	0.773[30]	//	Friedreich Ataxia (mitochondrial iron overload)
GSTP1	Glutathione S Transferase	Function unclear; strong correlation with iron levels; antioxidant[31,32]	Esophagus	42.628[33]	//	--
GLRX5	Glutaredoxin 5	Biogenesis of iron-sulfur clusters required for iron homeostasis[34]	Bone marrow	16.646[33]	//	Sideroblastic anemia 3 (iron deposits in erythroblast mitochondria); spasticity
HAMP	Hepcidin	Master iron regulation, antimicrobial; inhibits iron export by binding to ferroportin (iron export channel)[35]	Liver	2.727[36]	Decrease in plasma iron (increase in intracellular iron)	Juvenile Hemochromatosis Type 2B (iron overload)
HEPH	Hephaestin	Intestinal iron transport and absorption;[37] in mice, defects can lead to microcytic anemia[38]	Colon/small intestine/ duodenum	6.732[38]	//	
HFE	Human hemochromatosis protein	Iron absorption regulation via transferrin/transferrin receptor interaction;[39] regulates production of hepcidin;[40] defects in this gene may lead to over absorption of dietary iron and subsequent tissue iron deposition[39]	Thyroid/gall bladder	0.775[40]	//	Hemochromatosis Type 1 (iron overload); Porphyria (abnormal heme production); X-linked sideroblastic anemia (iron overload)
HJV (HFE2)	Hemojuvelin BMP Co-Receptor	Modulator of hepcidin expression as part of the BMP-SMAD signaling pathway[41]	Liver	NA[42]	Decrease in plasma iron (increase in intracellular iron)	Juvenile hemochromatosis Type 2A (iron overload)
HMOX1	Heme-oxygenase 1	Heme clearance and iron recycling[11,43]	Spleen	55.204[44]	//	Heme oxygenase 1 deficiency (impaired stress hematopoiesis leading to erythrocyte fragmentation and iron deposition in tissues); chronic obstructive

						pulmonary disease (COPD)
HMOX2	Heme-oxygenase 2	Heme clearance[43]	Testis	16.097[45]	//	--
HP	Haptoglobin	Clearance of free hemoglobin;[11,46] complex is endocytosed and degraded through lysosomal pathway and iron is released for recycling[47]	Liver	NA[48]	//	Anhaptoglobinemia (absence of haptoglobin)
HPX	Hemopexin	Heme scavenger; oxidative damage protection;[49] complex is endocytosed and degraded through lysosomal pathway and iron is released for recycling[47]	Liver	NA[50]	//	--
IREB2	Iron responsive element binding	Sense intracellular iron levels and respond through upregulation/downregulation of iron storage proteins via IRE;[51] when intracellular iron levels are low, IREB2 is upregulated and binds to IRE and allows influx of iron into the cell that may be from bloodstream or transported from dietary absorption	Thyroid	4.535[52]	//	--
LRP1	LDL receptor related protein	Heme-hemopexin scavenger; interacts with network of iron homeostasis genes; involved in lipid homeostasis and associated with amyloid beta precursor protein[53]	Fat	22.379[54]	//	Keratosis pilaris (KPA) (atrophy and scarring of facial skin)
PCBP1	Poly(RC) Binding Protein 1	Iron chaperone in delivery of iron to ferritin[55]	--	--	Decrease in plasma iron (increase in tissue iron)	--
PGRMC1	Progesterone receptor membrane	Heme chaperone and involved in lipid homeostasis[56]	Liver	91.919[57]	//	--
SLC46A1	Heme carrying protein 1	Intestinal heme transport[58]	Duodenum	7.338[59]	//	Hereditary folate malabsorption (inability to absorb folate from food)
SLC11A1	Solute Carrier Family 11 Member 1	Equips macrophages for efficient iron recycling[60]	Lung	1.438[61]	//	Juvenile idiopathic arthritis; susceptibility to ulcer and mycobacterium tuberculosis
SLC11A2	Divalent metal transporter 1	Iron and other divalent metal transport and intestinal iron uptake[62]	Duodenum/thyroid	9.017[63]	//	Iron deficiency anemia
SLC25A37	Solute Carrier Family 25 Member 37 (Mitoferrin 1)	Iron import for synthesis of mitochondrial heme[64]	Bone marrow	2.908[65]	//	
SLC40A1	Solute Carrier Family 40 Member 1 (Ferroportin)	Iron exporter[66]	Spleen	6.659[67]	Decrease in plasma iron (increase in tissue iron)	Hemochromatosis Type 4 (hypoferremia with tissue iron loading); African iron overload (iron overload)
SLC48A1	Solute carrier family 48 member 1	Heme transporter[23]	Brain	18.512[68]	//	--
STEAP3	STEAP3 Metalloreductase	Ferrireductase; iron transport facilitator[69]	Liver	1.736[70]	//	Hypochromic microcytic anemia with iron overload (iron deposition in tissues)
TF	Transferrin	Iron transport and delivery[71]; regulation of transferrin bound iron uptake from ferritin[72]; pH dependent uptake of transferrin bound iron[72]	Liver	95.394[73]	//	Atransferrinemia (anemia refractory to iron therapy)
TFRC	Transferrin receptor 1	Regulated in response to intracellular iron levels; receptor to import iron into the cell[72,74]	Bone marrow	12.737[75]	//	Immunodeficiency 46 (immunodeficiency

						disorder with chronic diarrhea, recurrent infections, etc.)
TFR2	Transferrin receptor 2	Regulated in response to intracellular iron levels; bind transferrin[72,74]	Liver	0.814[75]	//	Hemochromatosis Type 3 (iron overload)
TNF	Tumor necrosis factor	Iron sequestration;[76] promotes hypoferremia independent of hepcidin as part of the inflammatory response (down-regulates mucosal absorption of iron and results in systemic iron deficiency)[77]	Bone marrow	1.329[78]	Decrease in plasma iron (increase in intracellular iron)	Juvenile idiopathic arthritis; alopecia areata; idiopathic inflammatory myopathy; narcolepsy; Guillain-Barre syndrome (immune system and nerve disorder); asthma; migraines; susceptibility to malaria; Psoriatic arthritis

// = Responds to iron levels and does not directly act on iron levels; manages protein-bound iron (not free iron); or not enough information to definitively say; RPKM = Reads Per Kilobase Million; -- = information not available.

Appendix A Table 1 References

1. Genetics Home Reference. https://ghr.nlm.nih.gov/. Accessed May 14, 2019.
2. Anderson CP, Shen M, Eisenstein RS, Leibold EA. Mammalian iron metabolism and its control by iron regulatory proteins. *Biochim Biophys Acta - Mol Cell Res*. 2012;1823(9):1468-1483. doi:10.1016/j.bbamcr.2012.05.010.
3. NCBI. ACO1 aconitase 1 [Homo sapiens (human)] Gene ID: 48. NCBI Gene. https://www.ncbi.nlm.nih.gov/gene/48. Accessed May 9, 2019.
4. Rogers JT, Venkataramani V, Washburn C, et al. A role for amyloid precursor protein translation to restore iron homeostasis and ameliorate lead (Pb) neurotoxicity. *J Neurochem*. 2016;138(3):479-494. doi:10.1111/jnc.13671.
5. Gunawardena ND, Schrott V, Richardson C, et al. Aconitase: An Iron Sensing Regulator of Mitochondrial Oxidative Metabolism and Erythropoiesis. *Blood*. 2016;128(22). http://www.bloodjournal.org/content/128/22/74. Accessed May 9, 2019.
6. NCBI. ACO2 aconitase 2 [Homo sapiens (human)] Gene ID: 50. https://www.ncbi.nlm.nih.gov/gene/50. Accessed May 9, 2019.
7. Bullock GC, Delehanty LL, Talbot AL, et al. Iron control of erythroid development by a novel aconitase-associated regulatory pathway. *Blood*. 2010;116(1):97-108. doi:10.1182/blood-2009-10-251496.
8. NCBI. APP amyloid beta precursor protein [Homo sapiens (human)] Gene id: 351. https://www.ncbi.nlm.nih.gov/gene?Db=gene&Cmd=DetailsSearch&Term=351. Accessed May 9, 2019.
9. Pinto JP, Ramos P, de Almeida SF, et al. Protective role of calreticulin in HFE hemochromatosis. *Free Radic Biol Med*. 2008;44(1):99-108. doi:10.1016/j.freeradbiomed.2007.09.014.
10. NCBI. CALR calreticulin [Homo sapiens (human)] Gene ID: 811. https://www.ncbi.nlm.nih.gov/gene/811. Accessed May 9, 2019.
11. Thomsen JH, Etzerodt A, Svendsen P, Moestrup SK. The haptoglobin-cd163-heme oxygenase-1 pathway for hemoglobin scavenging. *Oxid Med Cell Longev*. 2013;2013:523652. doi:10.1155/2013/523652.
12. NCBI. CD163 Hemoglobin scavenger receptor molecule [Homo sapiens (human)] Gene ID: 9332. https://www.ncbi.nlm.nih.gov/gene/?term=9332%5Buid%5D. Accessed May 9, 2019.
13. Cherukuri S, Potla R, Sarkar J, Nurko S, Harris ZL, Fox PL. Unexpected role of ceruloplasmin in intestinal iron absorption. *Cell Metab*. 2005;2(5):309-319. doi:10.1016/j.cmet.2005.10.003.
14. NCBI. CP ceruloplasmin [Homo sapiens (human)] Gene ID: 1356. https://www.ncbi.nlm.nih.gov/gene/?term=1356%5Buid%5D. Accessed May 9, 2019.
15. Verroust PJ, Christensen EI. Megalin and cubilin--the story of two multipurpose receptors unfolds. *Nephrol Dial Transplant*. 2002;17(11):1867-1871. doi:10.1093/NDT/17.11.1867.
16. Weiss A, Spektor L, Cohen LA, et al. Orchestrated regulation of iron trafficking proteins in the kidney during iron overload facilitates systemic iron retention. *PLoS One*. 2018;13(10):e0204471.

doi:10.1371/journal.pone.0204471.

17. NCBI. CUBN cubilin [Homo sapiens (human)] Gene ID: 8029. https://www.ncbi.nlm.nih.gov/gene/?term=8029%5Buid%5D. Accessed May 9, 2019.

18. Choi J, Masaratana P, Latunde-Dada GO, Arno M, Simpson RJ, McKie AT. Duodenal reductase activity and spleen iron stores are reduced and erythropoiesis is abnormal in Dcytb knockout mice exposed to hypoxic conditions. *J Nutr*. 2012;142(11):1929-1934. doi:10.3945/jn.112.160358.

19. Lane DJR, Bae DH, Merlot AM, Sahni S, Richardson DR. Duodenal cytochrome b (DCYTB) in Iron metabolism: An update on function and regulation. *Nutrients*. 2015;7(4):2274-2296. doi:10.3390/nu7042274.

20. NCBI. CYBRD1 cytochrome b reductase 1 [Homo sapiens (human)] Gene ID: 79901. https://www.ncbi.nlm.nih.gov/gene/?term=79901%5Buid%5D. Accessed May 9, 2019.

21. Ferreira GC, Franco R, Lloyd SG, Moura I, Moura JJG, Huynh BH. Structure and function of ferrochelatase. *J Bioenerg Biomembr*. 1995;27(2):221-229. doi:10.1007/BF02110037.

22. NCBI. FECH ferrochelatase [Homo sapiens (human)] Gene ID: 2235. https://www.ncbi.nlm.nih.gov/gene/?term=2235%5Buid%5D. Accessed May 9, 2019.

23. Khan AA, Quigley JG. Heme and FLVCR-related transporter families SLC48 and SLC49. *Mol Aspects Med*. 2013;34(2-3):669-682. doi:10.1016/j.mam.2012.07.013.

24. NCBI. FLVCR1 feline leukemia virus subgroup C cellular receptor 1 [Homo sapiens (human)] Gene ID: 28982. https://www.ncbi.nlm.nih.gov/gene/28982. Accessed May 9, 2019.

25. NCBI. FTH1 ferritin heavy chain 1 [Homo sapiens (human)] Gene ID: 2495. https://www.ncbi.nlm.nih.gov/gene/2495. Accessed May 9, 2019.

26. Li W, Garringer HJ, Goodwin CB, et al. Systemic and cerebral iron homeostasis in ferritin knock-out Mice. Pantopoulos K, ed. *PLoS One*. 2015;10(1):e0117435. doi:10.1371/journal.pone.0117435.

27. NCBI. FTL ferritin light chain [Homo sapiens (human)] Gene ID: 2512. https://www.ncbi.nlm.nih.gov/gene/?term=2512%5Buid%5D. Accessed May 9, 2019.

28. Gao G, Chang YZ. Mitochondrial ferritin in the regulation of brain iron homeostasis and neurodegenerative diseases. *Front Pharmacol*. 2014;5 FEB:19. doi:10.3389/fphar.2014.00019.

29. Lane DJR, Richardson DR. Frataxin, a molecule of mystery: trading stability for function in its iron-binding site. *Biochem J*. 2010;426(2).

30. NCBI. FXN frataxin [Homo sapiens (human)] Gene ID: 2395. https://www.ncbi.nlm.nih.gov/gene/?term=2395%5Buid%5D. Accessed May 11, 2019.

31. Wu K-H, Chang J-G, Ho Y-J, Wu S-F, Peng C-T. Glutathione S-transferase M1 gene polymorphisms are associated with cardiac iron deposition in patients with beta-thalassemia major. *Hemoglobin*. 2006;30(2):251-256. doi:10.1080/03630260600642575.

32. Sharma V, Kumar B, Saxena R. Glutathione S-transferase gene deletions and their effect on iron status in HbE/β thalassemia patients. *Ann Hematol*. 2010;89(4):411-414. doi:10.1007/s00277-009-0847-y.

33. NCBI. GSTP1 glutathione S-transferase pi 1 [Homo sapiens (human)] Gene ID: 2950. https://www.ncbi.nlm.nih.gov/gene/?term=2950%5Buid%5D. Accessed May 11, 2019.

34. Berndt C, Lillig CH. Glutathione, Glutaredoxins, and Iron. *Antioxid Redox Signal*. 2017;27(15):ars.2017.7132. doi:10.1089/ars.2017.7132.

35. Ganz T. Systemic iron homeostasis. *Physiol Rev*. 2013;93(4):1721-1741. doi:10.1152/physrev.00008.2013.

36. NCBI. HAMP hepcidin antimicrobial peptide [Homo sapiens (human)] Gene ID: 57817. https://www.ncbi.nlm.nih.gov/gene/?term=57817%5Buid%5D. Accessed May 11, 2019.

37. Anderson GJ, Frazer DM, McKie AT, Vulpe CD. The ceruloplasmin homolog hephaestin and the control of intestinal iron absorption. *Blood Cells Mol Dis*. 2002;29(3):367-375. doi:10.1006/BCMD.2002.0576.

38. NCBI. HEPH hephaestin [Homo sapiens (human)] Gene ID: 9843. https://www.ncbi.nlm.nih.gov/gene/?term=9843%5Buid%5D. Accessed May 11, 2019.

39. Siah CW, Ombiga J, Adams LA, Trinder D, Olynyk JK. Normal iron metabolism and the pathophysiology of iron overload disorders. *Clin Biochem Rev*. 2006;27(1):5-16. http://www.ncbi.nlm.nih.gov/pubmed/16886043. Accessed July 23, 2016.

40. NCBI. HFE homeostatic iron regulator [Homo sapiens (human)] Gene ID: 3077. https://www.ncbi.nlm.nih.gov/gene/3077. Accessed May 11, 2019.

41. Core AB, Canali S, Babitt JL. Hemojuvelin and bone morphogenetic protein (BMP) signaling in iron homeostasis. *Front Pharmacol*. 2014;5 MAY(May):1-9. doi:10.3389/fphar.2014.00104.

42. NCBI. HJV hemojuvelin BMP co-receptor [Homo sapiens (human)] Gene ID: 148738.

43. Ayer A, Zarjou A, Agarwal A, Stocker R. Heme Oxygenases in Cardiovascular Health and Disease. *Physiol Rev*. 2016;96(4):1449-1508. doi:10.1152/physrev.00003.2016.

44. NCBI. HMOX1 heme oxygenase 1 [Homo sapiens (human)] Gene ID: 3162. https://www.ncbi.nlm.nih.gov/gene/?term=3162%5Buid%5D. Accessed May 11, 2019.

45. NCBI. HMOX2 heme oxygenase 2 [Homo sapiens (human)] Gene ID: 3163. https://www.ncbi.nlm.nih.gov/gene/?term=3163%5Buid%5D. Accessed May 11, 2019.

46. Kantor E, Bayir H, Ren D, et al. Haptoglobin genotype and functional outcome after aneurysmal subarachnoid hemorrhage. *J Neurosurg*. 2014;120(2):386-390. doi:10.3171/2013.10.JNS13219.

47. Anderson ER, Shah YM. Iron homeostasis in the liver. *Compr Physiol*. 2013;3(1):315-330. doi:10.1002/cphy.c120016.

48. NCBI. HP haptoglobin [Homo sapiens (human)] Gene ID: 3240. https://www.ncbi.nlm.nih.gov/gene/?term=3240%5Buid%5D. Accessed May 11, 2019.

49. Hvidberg V, Maniecki MB, Jacobsen C, Højrup P, Møller HJ, Moestrup SK. Identification of the receptor scavenging hemopexin-heme complexes. *Blood*. 2005;106(7):2572-2579. doi:10.1182/blood-2005-03-1185.

50. NCBI. HPX hemopexin [Homo sapiens (human)] Gene ID: 3263.

https://www.ncbi.nlm.nih.gov/gene/?term=3263%5Buid%5D. Accessed May 11, 2019.

51. Crichton R. *Iron Metabolism: From Molecular Mechanisms to Clinical Consequences: Fourth Edition*. Wiley; 2016. doi:10.1002/9781118925645.

52. NCBI. IREB2 iron responsive element binding protein 2 [Homo sapiens (human)] Gene ID: 3658. https://www.ncbi.nlm.nih.gov/gene/?term=3658%5Buid%5D. Accessed May 11, 2019.

53. Wang G, Guo Z, Tong L, et al. TLR7 (Toll-Like Receptor 7) Facilitates Heme Scavenging Through the BTK (Bruton Tyrosine Kinase)–CRT (Calreticulin)–LRP1 (Low-Density Lipoprotein Receptor–Related Protein-1)–Hx (Hemopexin) Pathway in Murine Intracerebral Hemorrhage. *Stroke*. 2018;49(12):1. doi:10.1161/STROKEAHA.118.022155.

54. NCBI. LRP1 LDL receptor related protein 1 [Homo sapiens (human)] Gene ID: 4035. https://www.ncbi.nlm.nih.gov/gene/?term=4035%5Buid%5D. Accessed May 11, 2019.

55. Shi H, Bencze KZ, Stemmler TL, Philpott CC. A Cytosolic Iron Chaperone That Delivers Iron to Ferritin. *Science (80-)*. 2008;320(5880):1207-1210. doi:10.1126/science.1157643.

56. Piel RB, Shiferaw MT, Vashisht AA, et al. A Novel Role for Progesterone Receptor Membrane Component 1 (PGRMC1): A Partner and Regulator of Ferrochelatase. *Biochemistry*. 2016;55(37):5204-5217. doi:10.1021/acs.biochem.6b00756.

57. NCBI. PGRMC1 progesterone receptor membrane component 1 [Homo sapiens (human)] Gene ID: 10857. https://www.ncbi.nlm.nih.gov/gene/10857. Accessed May 11, 2019.

58. Shayeghi M, Latunde-Dada GO, Oakhill JS, et al. Identification of an Intestinal Heme Transporter. *Cell*. 2005;122(5):789-801. doi:10.1016/j.cell.2005.06.025.

59. NCBI. SLC46A1 solute carrier family 46 member 1 [Homo sapiens (human)] Gene ID: 113235. https://www.ncbi.nlm.nih.gov/gene/?term=113235%5Buid%5D. Accessed May 11, 2019.

60. Soe-Lin S, Apte SS, Mikhael MR, Kayembe LK, Nie G, Ponka P. Both Nramp1 and DMT1 are necessary for efficient macrophage iron recycling. *Exp Hematol*. 2010;38(8):609-617. doi:10.1016/j.exphem.2010.04.003.

61. NCBI. SLC11A1 solute carrier family 11 member 1 [Homo sapiens (human)] Gene ID: 6556. https://www.ncbi.nlm.nih.gov/gene/?term=6556%5Buid%5D. Accessed May 11, 2019.

62. Hubert N, Hentze MW. Previously uncharacterized isoforms of divalent metal transporter (DMT)-1: implications for regulation and cellular function. *Proc Natl Acad Sci U S A*. 2002;99(19):12345-12350. doi:10.1073/pnas.192423399.

63. NCBI. SLC11A2 solute carrier family 11 member 2 [Homo sapiens (human)] Gene ID: 4891. https://www.ncbi.nlm.nih.gov/gene/?term=4891%5Buid%5D. Accessed May 11, 2019.

64. Richardson DR, Lane DJR, Becker EM, et al. Mitochondrial iron trafficking and the integration of iron metabolism between the mitochondrion and cytosol. *Proc Natl Acad Sci U S A*. 2010;107(24):10775-10782. doi:10.1073/pnas.0912925107.

65. NCBI. SLC25A37 solute carrier family 25 member 37 [Homo sapiens (human)] Gene ID: 51312. https://www.ncbi.nlm.nih.gov/gene/51312. Accessed May 11, 2019.

66. Majore S, Bonaccorsi di Patti MC, Valiante M, et al. Characterization of three novel pathogenic

SLC40A1 mutations and genotype/phenotype correlations in 7 Italian families with type 4 hereditary hemochromatosis. *Biochim Biophys Acta - Mol Basis Dis*. 2018;1864(2):464-470. doi:10.1016/j.bbadis.2017.11.006.

67. NCBI. SLC40A1 solute carrier family 40 member 1 [Homo sapiens (human)] Gene ID: 30061. https://www.ncbi.nlm.nih.gov/gene/?term=30061%5Buid%5D. Accessed May 11, 2019.

68. NCBI. SLC48A1 solute carrier family 48 member 1 [Homo sapiens (human)] Gene ID: 55652. https://www.ncbi.nlm.nih.gov/gene/?term=55652%5Buid%5D. Accessed May 11, 2019.

69. Zhang F, Tao Y, Zhang Z, et al. Metalloreductase Steap3 coordinates the regulation of iron homeostasis and inflammatory responses. *Haematol Hematol J*. 2012;97(12):1826. http://connection.ebscohost.com/c/articles/84662557/metalloreductase-steap3-coordinates-regulation-iron-homeostasis-inflammatory-responses.

70. NCBI. STEAP3 STEAP3 metalloreductase [Homo sapiens (human)] Gene ID: 55240. https://www.ncbi.nlm.nih.gov/gene/?term=55240%5Buid%5D. Accessed May 11, 2019.

71. Bartnikas TB, Andrews NC, Fleming MD. Transferrin is a major determinant of hepcidin expression in hypotransferrinemic mice. *Blood*. 2011;117(2):630-637. doi:10.1182/blood-2010-05-287359.

72. Trinder D, Baker E. Transferrin receptor 2: a new molecule in iron metabolism. *Int J Biochem Cell Biol*. 2003;35(3):292-296. doi:10.1016/S1357-2725(02)00258-3.

73. NCBI. TF transferrin [Homo sapiens (human)] Gene ID: 7018. https://www.ncbi.nlm.nih.gov/gene/?term=7018%5Buid%5D. Accessed May 11, 2019.

74. Kleven MD, Jue S, Enns CA. Transferrin Receptors TfR1 and TfR2 Bind Transferrin through Differing Mechanisms. *Biochemistry*. 2018;57(9):1552-1559. doi:10.1021/acs.biochem.8b00006.

75. NCBI. TFRC transferrin receptor [Homo sapiens (human)] Gene ID: 7037. https://www.ncbi.nlm.nih.gov/gene/7037. Accessed May 11, 2019.

76. Nanami M, Ookawara T, Otaki Y, et al. Tumor Necrosis Factor-α–Induced Iron Sequestration and Oxidative Stress in Human Endothelial Cells. *Arterioscler Thromb Vasc Biol*. 2005;25(12).

77. Wessling-Resnick M. Iron Homeostasis and the Inflammatory Response. *Annu Rev Nutr*. 2010;30(1):105-122. doi:10.1146/annurev.nutr.012809.104804.

78. NCBI. TNF tumor necrosis factor [Homo sapiens (human)] Gene ID: 7124. https://www.ncbi.nlm.nih.gov/gene/?term=7124%5Buid%5D. Accessed May 11, 2019.

Appendix B Data-Based Manuscript (Section 3.0) Tables and Figures

Appendix Table 2 Candidate Gene and Data Extraction Information

Gene Symbol	Gene Name	Extraction window[a]	Number of CpG sites	Data status
ACO1	Iron regulatory protein 1	chr9:32382601-32452832	11	Available
ACO2	Iron regulatory protein 2	chr22:41863129-41926993	12	Available
APP	Amyloid precursor protein	chr21:27250861-27545138	21	Available
CALR	Calreticulin	chr19:13047414-13057304	30	Available
CD163	Hemoglobin scavenger receptor	chr12:7621412-7658414	3	Available
CP	Ceruloplasmin	chr3:148888290-148941832	6	Available
CUBN	Cubilin	chr10:16863965-17173816	15	Available
CYBRD1	Duodenal cytochrome b	chr2:172376866-172416643	13	Available
FECH	Ferrochelatase	chr18:55210073-55255969	13	Available
FLVCR1	Feline leukemia virus subgroup C receptor	chr1:213029597-213074705	18	Available
FTH1	Ferritin heavy	chr11:61729757-61737132	20	Available
FTL	Ferritin light	chr19:49466566-49472136	11	Available
FTMT	Mitochondrial ferritin	chr5:121185650-121190523	8	Available
FXN	Frataxin	chr9:71648479-71695993	13	Available
GSTP1	Glutathione S Transferase	chr11:67349066-67356124	12	Available
GLRX5	Glutaredoxin 5		16	Available
HAMP	Hepcidin	chr19:35771410-35778045	NA	Analyzed in pilot work[32]
HEPH	Hephaestin	chrX:65380433-65489230	NA	Data not available
HFE	Human hemochromatosis protein	chr6:26094615-26099056	1	Available
HJV (HFE2)	Hemojuvelin BMP co-Receptor	chr1:145411191-145419545	13	Available
HMOX1	Heme-oxygenase 1	chr22:35775060-35792207	8	Available
HMOX2	Heme-oxygenase 2	chr16:4524341-4562348	22	Available
HP	Haptoglobin	chr16:72086508-72096955	5	Available
HPX	Hemopexin	chr11:6450268-6464254	7	Available
IREB2	Iron responsive element binding	chr15:78728518-78795798	15	Available

LRP1	LDL receptor related protein	chr12:57520282-57609125	91	Available
PCBP1	poly(RC) Binding Protein 1	chr2:70312585-70318334	17	Available
PGRMC1	Progesterone receptor membrane	chrX:118368211-118380429	NA	Data not available
SLC46A1	Heme carrying protein 1	chr17:26719661-26735230	15	Available
SLC11A1	Solute Carrier Family 11 Member 1	chr2:219244752-219263617	26	Available
SLC11A2	Divalent metal transporter 1	chr12:51377775-51424058	15	Available
SLC25A37	Solute Carrier Family 25 Member 37 (Mitoferrin 1)	chr8:23384363-23432063	28	Available
SLC40A1 (FP)	Solute Carrier Family 40 Member 1 (Ferroportin)	chr2:190423316-190447537	14	Available
SLC48A1	Solute carrier family 48 member 1	chr12:48164967-48178536	14	Available
STEAP3	STEAP3 Metalloreductase	chr2:119979384-120025227	29	Available
TF	Transferrin	chr3:133462800-133499850	15	Available
TFRC	Transferrin receptor 1	chr3:195774155-195811032	16	Available
TFR2	Transferrin receptor 2	chr7:100216039-100241201	29	Available
TNF	Tumor necrosis factor	chr6:31542292-31548112	35	Available

NOTE: [a] UCSC Genome Browser Build GRCh37/hg19

Appendix Table 3 Sample Characteristics

Variable	aSAH cohort (n=648)	Targeted discovery sample (n=260)
Age, mean years (SD)	53.2 (11.5)	53.1 (11.0)
Sex, female (n, %)	469 (72.4)	179 (68.8)
Race, White (n, %)	563 (86.9)	225 (86.5)
Treatment, surgery (n, %)	244 (37.7)	101 (38.8)
Fisher grade (n, %)		
2	266 (41.0)	78 (30.0)
3	279 (43.1)	126 (48.5)
4	103 (15.9)	56 (21.5)
WFNS (n, %)		
1	330 (50.9)	108 (41.5)
2	115 (17.7)	57 (21.9)
3	47 (7.3)	23 (8.8)
4	90 (13.9)	44 (16.9)
5	66 (10.2)	28 (10.9)

NOTE: SD, standard deviation; WFNS, World Federation of Neurosurgical Societies System; aSAH, aneurysmal subarachnoid hemorrhage

Appendix Table 4 Cases Control Sample Sizes by Outcome

Outcome	N	n (%), cases	n (%), controls
CV	168	91 (54.2)	77 (45.8)
DCI	258	127 (49.2)	131 (50.8)
GOS3	214	71 (33.2)	143 (66.8)
GOS12	204	53 (26.0)	151 (74.0)
MORT3	232	39 (16.8)	193 (83.2)
MORT12	204	44 (21.6)	160 (78.4)

NOTE: CV, cerebral vasospasm; DCI, delayed cerebral ischemia; GOS3, Glasgow Outcome Scale at 3 months (poor = 1-3); GOS12, Glasgow Outcome Scale at 12 months (poor = 1-3); MORT3, mortality at 3 months; MORT12, mortality at 12 months; Cases indicate the number of participants with poor outcomes (e.g. occurrence of CV or DCI, poor GOS, or death)

Appendix Table 5 DNA Methylation Data Extreme Outlier Summary

Gene Symbol	Extreme Outliers	Gene Symbol	Extreme Outliers
ACO1	31	HMOX1	29
ACO2	60	HMOX2	61
APP	62	*HP*	15
CALR	39	*HPX*	23
CD163	38	*IREB2*	36
CP	20	*LRP1*	189
CUBN	50	*PCBP1*	52
CYBRD1	31	*SLC46A1*	39
FECH	30	*SLC11A1*	45
FLVCR1	28	*SLC11A2*	29
FTH1	35	*SLC25A37*	61
FTL	25	*SLC40A1 (FP)*	43
FTMT	28	*SLC48A1*	43
FXN	31	*STEAP3*	51
GSTP1	28	*TF*	57
GLRX5	52	*TFRC*	43
HFE	0	*TFR2*	60
HJV (HFE2)	28	*TNF*	62

NOTE: Number of extreme DNA methylation outliers (three times the interquartile range) score adjusted and pulled in

Appendix Table 6 Summary of Signficiant and Suggestive Associations

Outcome	Gene	CpG	Position[a]	UCSC Region[b]	Relationship to Island[c]	p
GOS12	*APP*	cg08866780	27543523	TSS1500	Island	0.00013
GOS3	*TNF*	cg08553327	31543647	1stExon	OpenSea	0.00027
GOS3	*TNF*	cg21222743	31543545	1stExon	OpenSea	0.00032
GOS3	*TNF*	cg10717214	31543557	1stExon	OpenSea	0.00039
GOS3	*TNF*	cg08553327	31543647	1stExon	OpenSea	0.00055
MORT12	*APP*	cg08866780	27543523	TSS1500	Island	0.00058
GOS3	*STEAP3*	cg25713625	120022835	3'UTR	S_Shore	0.00061
GOS3	*SLC25A37*	cg18186478	23423960	Body	OpenSea	0.00061
MORT12	*ACO1*	cg13977526	32384244	TSS1500	N_Shore	0.00062
GOS3	*TNF*	cg10650821	31543686	1stExon	OpenSea	0.00083
GOS3	*TNF*	cg12681001	31543540	1stExon	OpenSea	0.00089
MORT3	*APP*	cg08866780	27543523	TSS1500	Island	0.00099
GO3	*LRP1*	cg09150519	57566995	Body	N_Shelf	0.00116
GOS3	*FECH*	cg15864104	55254508	TSS1500	S_Shore	0.00118
GOS3	*TNF*	cg21222743	31543545	1stExon	OpenSea	0.00121
GOS3	*FTH1*	cg24803614	61733475	Body	N_Shore	0.00125
GOS3	*FXN*	cg02268013	71650331	TSS200	N_Shore	0.00133
GOS3	*FXN*	cg21229044	71669714	Body	OpenSea	0.00151
GOS12	*LRP1*	cg12069468	57578649	Body	N_Shore	0.00155
GOS12	*STEAP3*	cg25713625	120022835	3'UTR	S_Shore	0.00182
GOS3	*FLVCR1*	cg23082281	213032677	TSS1500;Body	S_Shore	0.00184
GOS3	*SLC11A2*	cg22826226	51421367	TSS1500	S_Shore	0.00192
GOS3	*LRP1*	cg12069468	57578649	Body	N_Shore	0.00223
MORT12	*TNF*	cg10717214	31543557	1stExon	OpenSea	0.00251
MORT12	*LRP1*	cg12069468	57578649	Body	N_Shore	0.00273
GOS3	*HMOX2*	cg14951292	4525986	TSS1500;5'UTR	N_Shore	0.00289
MORT3	*ACO2*	cg25636833	41864805	TSS200;TSS1500	Island	0.00311
GOS3	*TNF*	cg21467614	31543638	1stExon	OpenSea	0.00316
MORT12	*TNF*	cg21222743	31543545	1stExon	OpenSea	0.00368
MORT12	*TNF*	cg08553327	31543647	1stExon	OpenSea	0.00370
GOS3	*HFE2*	cg26036288	145411967	TSS1500	N_Shelf	0.00388
GOS3	*CYBRD1*	cg15691325	172387940	Body	OpenSea	0.00399
MORT12	*STEAP3*	cg25713625	120022835	3'UTR	S_Shore	0.00410
GOS12	*APP*	cg03779431	27543124	1stExon;5'UTR	Island	0.00411
MORT3	*LRP1*	cg09150519	57566995	Body	N_Shelf	0.00461
GOS12	*ACO1*	cg13977526	32384244	TSS1500	N_Shore	0.00502
GOS3	*LRP1*	cg22339313	57578642	Body	N_Shore	0.00512
GOS12	*HMOX2*	cg14951292	4525986	TSS1500	N_Shore	0.00520
GOS3	*APP*	cg08866780	27543523	TSS1500	Island	0.00596
MORT12	*IREB2*	cg08092050	78734858	Body	S_Shelf	0.00621
GOS3	*IREB2*	cg03700171	78748272	Body	OpenSea	0.00628
GOS12	*TFRC*	cg09470983	195805948	5'UTR	N_Shelf	0.00639
GOS3	*TF*	cg20123637	133467435	Body	S_Shelf	0.00672
GOS3	*IREB2*	cg08092050	78734858	Body	S_Shelf	0.00714
MORT12	*LRP1*	cg06531661	57522120	TSS200	Island	0.00722
MORT3	*FLVCR1*	cg23082281	213032677	TSS1500;Body	S_Shore	0.00729

MORT12	*LRP1*	cg16612926	57601671	Body	OpenSea	0.00779
GOS12	*FXN*	cg21229044	71669714	Body	OpenSea	0.00791
MORT12	*FXN*	cg21229044	71669714	Body	OpenSea	0.00800
GOS3	*GSTP1*	cg06928838	67351490	Body	Island	0.00826
MORT3	*TNF*	cg10717214	31543557	1stExon	OpenSea	0.00828
MORT12	*CALR*	cg27180563	13056485	TSS200	Island	0.00835
DCI	*TF*	cg16262614	133464971	TSS200	Island	0.00840
MORT12	*TNF*	cg12681001	31543540	1stExon	OpenSea	0.00851
GOS3	*TFRC*	cg23502966	195789749	Body	OpenSea	0.00875
MORT12	*SLC25A37*	cg00892891	23423734	Body	OpenSea	0.00877
CV	*FXN*	cg07158339	71650237	TSS1500	N_Shore	0.00880
MORT3	*GLRX5*	cg01304278	96001998	Body;TSS1500	S_Shore	0.00886
MORT3	*TNF*	cg08553327	31543647	1stExon	OpenSea	0.00892
MORT12	*LRP1*	cg09150519	57566995	Body	N_Shelf	0.00931
MORT3	*STEAP3*	cg25713625	120022835	3'UTR	S_Shore	0.00933
MORT3	*TNF*	cg21222743	31543545	1stExon	OpenSea	0.00948
MORT3	*FXN*	cg21229044	71669714	Body	OpenSea	0.00959
MORT3	*FXN*	cg02268013	71650331	TSS200	N_Shore	0.00961
GOS3	*CP*	cg05776336	148939487	1stExon	OpenSea	0.00983
GOS3	*CP*	cg24876069	148916713	Body	OpenSea	0.00992
GOS3	*LRP1*	cg11813198	57572413	Body	S_Shelf	0.01039
GOS3	*TFRC*	cg09470983	195805948	5'UTR	N_Shelf	0.01040
GOS12	*GLRX5*	cg01304278	96001998	Body;TSS1500	S_Shore	0.01049
MORT12	*SLC11A2*	cg22826226	51421367	TSS1500	S_Shore	0.01065
GOS12	*SLC25A37*	cg18186478	23423960	Body	OpenSea	0.01073
GOS3	*CP*	cg19278448	148905466	Body	OpenSea	0.01078
GOS3	*LRP1*	cg21702971	57588243	Body;TSS200	OpenSea	0.01107
GOS12	*GLRX5*	cg02940042	96000874	TSS1500;Body	Island	0.01144
MORT3	*GSTP1*	cg06928838	67351490	Body	Island	0.01164
MORT3	*TF*	cg20123637	133467435	Body	S_Shelf	0.01207
MORT12	*GLRX5*	cg02940042	96000874	TSS1500;Body	Island	0.01252
MORT3	*ACO1*	cg13977526	32384244	TSS1500	N_Shore	0.01293
GOS3	*LRP1*	cg16612926	57601671	Body	OpenSea	0.01344
GOS3	*GLRX5*	cg02940042	96000874	TSS1500;Body	Island	0.01356
MORT3	*IREB2*	cg08092050	78734858	Body	S_Shelf	0.01381
GOS12	*LRP1*	cg22339313	57578642	Body	N_Shore	0.01388
GOS12	*TNF*	cg10717214	31543557	1stExon	OpenSea	0.01388
GOS3	*CYBRD1*	cg15691325	172387940	Body;Body	OpenSea	0.01395
CV	*SLC25A37*	cg24126361	23398346	Body	OpenSea	0.01451
MORT12	*SLC25A37*	cg18186478	23423960	Body	OpenSea	0.01458
MORT3	*FECH*	cg17299799	55214063	3'UTR	OpenSea	0.01491
MORT12	*FXN*	cg02268013	71650331	TSS200	N_Shore	0.01493
GOS12	*HPX*	cg15741583	6461908	Body	OpenSea	0.01513
MORT3	*GLRX5*	cg02940042	96000874	TSS1500;Body	Island	0.01514
GOS12	*LRP1*	cg09150519	57566995	Body	N_Shelf	0.01551
GOS3	*IREB2*	cg13714459	78730363	TSS200	Island	0.01571
GOS3	*CALR*	cg27180563	13056485	TSS200	Island	0.01575
DCI	*CYBRD1*	cg15691325	172387940	Body	OpenSea	0.01581
GOS3	*LRP1*	cg06531661	57522120	TSS200	Island	0.01586

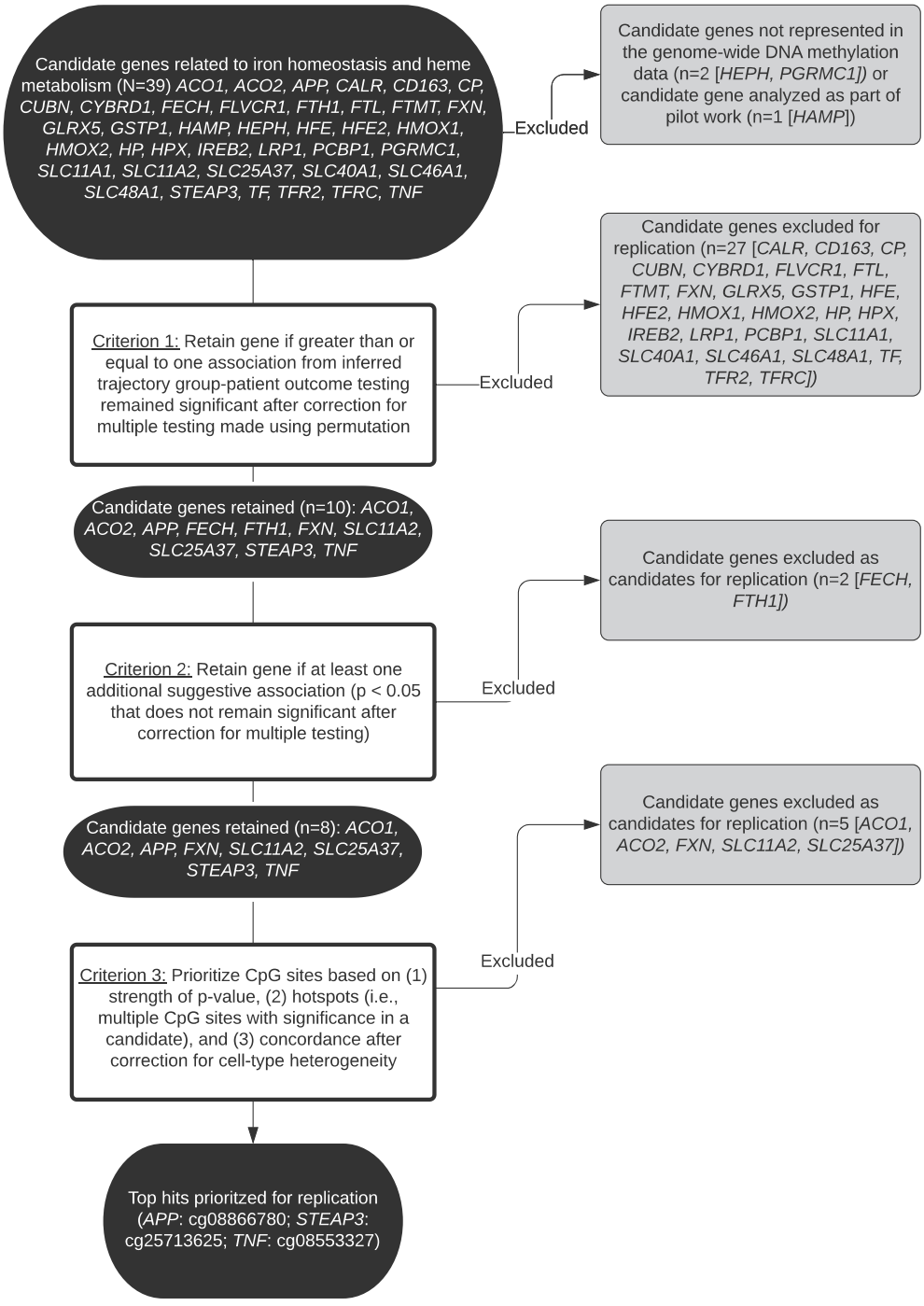
CV	*ACO2*	cg25636833	41864805	TSS200;TSS1500	Island	0.01604
DCI	*FLVCR1*	cg18789758	213030449	Body;TSS1500	N_Shore	0.01622
MORT3	*FP*	cg10752008	190445175	1stExon	N_Shore	0.01625
GOS3	*SLC48A1*	cg01154392	48167362	Body	Island	0.01656
MORT3	*FECH*	cg26043149	55253948	5'UTR;1stExon	Island	0.01669
GOS3	*TNF*	cg12681001	31543540	1stExon	OpenSea	0.01692
MORT3	*SLC11A2*	cg22826226	51421367	TSS1500	S_Shore	0.01696
GOS3	*FLVCR1*	cg18789758	213030449	Body;TSS1500	N_Shore	0.01718
MORT12	*GLRX5*	cg01304278	96001998	Body;TSS1500	S_Shore	0.01737
GOS12	*LRP1*	cg06531661	57522120	TSS200	Island	0.01744
MORT12	*ACO2*	cg25636833	41864805	TSS200;TSS1500	Island	0.01783
MORT3	*PCBP1*	cg26487157	70313295	TSS1500	Island	0.01808
MORT3	*TNF*	cg12681001	31543540	1stExon	OpenSea	0.01827
MORT3	*LRP1*	cg16612926	57601671	Body	OpenSea	0.01835
GOS3	*APP*	cg15407086	27543045	1stExon;5'UTR	Island	0.01839
CV	*CALR*	cg27180563	13056485	TSS200	Island	0.01842
MORT12	*FXN*	cg02268013	71650331	TSS200	N_Shore	0.01876
DCI	*SLC25A37*	cg24917065	23418389	Body	OpenSea	0.01880
GOS3	*FXN*	cg14410042	71654823	Body	S_Shelf	0.01905
GOS12	*TFRC*	cg21494636	195808986	1stExon;5'UTR;TSS200	Island	0.01955
GOS3	*HFE2*	cg00987513	145413337	1stExon;5'UTR	N_Shelf	0.01966
GOS12	*TNF*	cg21222743	31543545	1stExon	OpenSea	0.01986
MORT12	*TNF*	cg21467614	31543638	1stExon	OpenSea	0.02027
GOS3	*FLVCR1*	cg14912380	213068483	Body	OpenSea	0.02041
MORT12	*TNF*	cg10650821	31543686	1stExon	OpenSea	0.02068
GOS12	*TNF*	cg08553327	31543647	1stExon	OpenSea	0.02086
MORT12	*FLVCR1*	cg18789758	213030449	Body;TSS1500	N_Shore	0.02094
MORT3	*TF*	cg20123637	133467435	Body	S_Shelf	0.02094
MORT12	*LRP1*	cg22339313	57578642	Body	N_Shore	0.02131
MORT3	*LRP1*	cg12069468	57578649	Body	N_Shore	0.02135
GOS12	*ACO2*	cg25636833	41864805	TSS200;TSS1500	Island	0.02164
MORT12	*TFR2*	cg19767562	100224437	Body	Island	0.02181
CV	*ACO2*	cg26423539	41864790	TSS200;TSS1500	Island	0.02209
MORT12	*TNF*	cg21222743	31543545	1stExon	OpenSea	0.02218
GOS3	*PCBP1*	cg00396520	70314546	TSS200	Island	0.02226
GOS3	*HPX*	cg15741583	6461908	Body	OpenSea	0.02284
DCI	*HMOX2*	cg06605704	4551147	5'UTR	OpenSea	0.02351
MORT12	*TNF*	cg08553327	31543647	1stExon	OpenSea	0.02433
CV	*CYBRD1*	cg15691325	172387940	Body	OpenSea	0.02464
MORT12	*GSTP1*	cg06928838	67351490	Body	Island	0.02538
MORT12	*CP*	cg24876069	148916713	Body	OpenSea	0.02546
MORT3	*FLVCR1*	cg18789758	213030449	Body;TSS1500	N_Shore	0.02563
MORT3	*TNF*	cg08553327	31543647	1stExon	OpenSea	0.02573
MORT12	*FTL*	cg05617973	49468917	Body	Island	0.02588
MORT12	*FTH1*	cg04230589	61735144	TSS200	Island	0.02601
GOS12	*CP*	cg19278448	148905466	Body	OpenSea	0.02602
DCI	*FTH1*	cg22892043	61731930	3'UTR	N_Shelf	0.02668
MORT3	*IREB2*	cg13714459	78730363	TSS200	Island	0.02694
GOS3	*TNF*	cg02581828	31546769		N_Shore	0.02704

GOS12	*SLC11A2*	cg22826226	51421367	TSS1500	S_Shore	0.02744
MORT12	*FP*	cg10752008	190445175	1stExon	N_Shore	0.02745
GOS12	*LRP1*	cg05943813	57573255	Body	S_Shelf	0.02790
GOS12	*LRP1*	cg21702971	57588243	Body;TSS200	OpenSea	0.02828
CV	*FTH1*	cg04230589	61735144	TSS200	Island	0.02848
GOS3	*LRP1*	cg14621254	57569787	Body	Island	0.02880
GOS12	*TFR2*	cg19767562	100224437	Body	Island	0.02889
DCI	*TNF*	cg12681001	31543540	1stExon	OpenSea	0.02892
GOS3	*HFE2*	cg14036143	145415441	5'UTR;Body	Island	0.02895
GOS12	*LRP1*	cg06531661	57522120	TSS200	Island	0.02900
GOS12	*STEAP3*	cg25713625	120022835	3'UTR	S_Shore	0.02915
MORT12	*TF*	cg17775713	133465469	Body	S_Shore	0.02943
GOS12	*FTL*	cg26532042	49468557	TSS200	Island	0.02987
GOS3	*FTH1*	cg02315732	61732658	Body	N_Shelf	0.02988
GOS12	*SLC48A1*	cg01154392	48167362	Body	Island	0.03037
GOS12	*TFR2*	cg19767562	100224437	Body	Island	0.03048
GOS3	*SLC11A2*	cg25493658	51421225	TSS1500	S_Shore	0.03077
MORT3	*TNF*	cg21222743	31543545	1stExon	OpenSea	0.03087
GOS12	*LRP1*	cg10904028	57590920	Body	OpenSea	0.03137
MORT12	*TF*	cg24512138	133465360	Body	Island	0.03140
MORT12	*STEAP3*	cg25508118	119981769	5'UTR	Island	0.03148
GOS3	*ACO2*	cg25636833	41864805	TSS200;TSS1500	Island	0.03202
GOS12	*FTH1*	cg04230589	61735144	TSS200	Island	0.03212
CV	*CUBN*	cg17244673	17166078	Body	OpenSea	0.03218
GOS3	*FLVCR1*	cg23082281	213032677	TSS1500;Body	S_Shore	0.03254
GOS3	*CYBRD1*	cg24200501	172378036	TSS1500	N_Shore	0.03284
CV	*CP*	cg17439694	148939523	1stExon	OpenSea	0.03322
MORT12	*TF*	cg20123637	133467435	Body	S_Shelf	0.03331
MORT12	*APP*	cg23269692	27372446	Body	OpenSea	0.03349
GOS3	*LRP1*	cg22339313	57578642	Body	N_Shore	0.03372
GOS3	*TNF*	cg10717214	31543557	1stExon	OpenSea	0.03491
MORT12	*TF*	cg24512138	133465360	Body	Island	0.03494
GOS12	*TNF*	cg23204396	31548107		N_Shore	0.03500
GOS3	*FTMT*	cg09441819	121186479	TSS1500	N_Shore	0.03557
MORT3	*CP*	cg19278448	148905466	Body	OpenSea	0.03566
GOS3	*APP*	cg15407086	27543045	1stExon;5'UTR	Island	0.03570
MORT12	*LRP1*	cg01276169	57569898	Body	Island	0.03572
GOS3	*FXN*	cg02268013	71650331	TSS200	N_Shore	0.03594
MORT3	*LRP1*	cg01276169	57569898	Body	Island	0.03616
GOS3	*TF*	cg18286850	133464713	TSS1500	N_Shore	0.03625
DCI	*TNF*	cg21467614	31543638	1stExon	OpenSea	0.03650
GOS3	*TNF*	cg26729380	31543655	1stExon	OpenSea	0.03664
MORT12	*FECH*	cg17299799	55214063	3'UTR	OpenSea	0.03709
CV	*GSTP1*	cg14921275	67350511	TSS1500	N_Shore	0.03710
MORT3	*TNF*	cg10650821	31543686	1stExon	OpenSea	0.03717
GOS12	*IREB2*	cg08092050	78734858	Body	S_Shelf	0.03767
CV	*SLC25A37*	cg24126361	23398346	Body	OpenSea	0.03841
DCI	*CUBN*	cg17436460	17028480	Body	OpenSea	0.03879
DCI	*TNF*	cg09637172	31545252	Body	N_Shelf	0.03914

DCI	*TNF*	cg10717214	31543557	1stExon	OpenSea	0.03964
GOS12	*STEAP3*	cg25508118	119981769	5'UTR	Island	0.03970
GOS3	*SLC25A37*	cg20845639	23430886		S_Shore	0.04010
GOS12	*CALR*	cg27180563	13056485	TSS200	Island	0.04023
CV	*CP*	cg05776336	148939487	1stExon	OpenSea	0.04023
CV	*FTH1*	cg04230589	61735144	TSS200	Island	0.04027
MORT3	*FXN*	cg02268013	71650331	TSS200	N_Shore	0.04030
MORT12	*HMOX2*	cg14951292	4525986	TSS1500	N_Shore	0.04036
GOS3	*SLC46A1*	cg00497630	26733052	1stExon	Island	0.04041
GOS12	*SLC25A37*	cg00892891	23423734	Body	OpenSea	0.04052
MORT12	*TFR2*	cg19767562	100224437	Body	Island	0.04113
MORT12	*TNF*	cg26729380	31543655	1stExon	OpenSea	0.04124
DCI	*TF*	cg08234618	133464842	TSS200	N_Shore	0.04133
MORT12	*FECH*	cg26043149	55253948	5'UTR;1stExon	Island	0.04148
MORT12	*LRP1*	cg12146864	57569768	Body	Island	0.04198
CV	*PCBP1*	cg09175843	70313417	TSS1500	Island	0.04228
MORT3	*TNF*	cg21467614	31543638	1stExon	OpenSea	0.04280
GOS12	*HFE2*	cg00987513	145413337	5'UTR;1stExon	N_Shelf	0.04306
MORT12	*TF*	cg20123637	133467435	Body	S_Shelf	0.04324
GOS12	*TNF*	cg21222743	31543545	1stExon	OpenSea	0.04361
DCI	*GLRX5*	cg12096759	96002581	Body;TSS1500	S_Shore	0.04390
GOS3	*FLVCR1*	cg00156403	213035301	Body	S_Shelf	0.04418
GOS12	*IREB2*	cg13714459	78730363	TSS200	Island	0.04541
MORT12	*FTH1*	cg19407570	61735476	TSS1500	Island	0.04581
GOS12	*LRP1*	cg16612926	57601671	Body	OpenSea	0.04611
DCI	*TF*	cg16262614	133464971	TSS200	Island	0.04632
MORT12	*FLVCR1*	cg14912380	213068483	Body	OpenSea	0.04633
GOS12	*FTL*	cg26532042	49468557	TSS200	Island	0.04639
GOS3	*CP*	cg17439694	148939523	1stExon	OpenSea	0.04651
MORT3	*SLC25A37*	cg18186478	23423960	Body	OpenSea	0.04657
MORT3	*LRP1*	cg06531661	57522120	TSS200	Island	0.04681
MORT3	*CP*	cg19278448	148905466	Body	OpenSea	0.04689
GOS12	*TNF*	cg12681001	31543540	1stExon	OpenSea	0.04730
GOS3	*HFE2*	cg14036143	145415441	5'UTR;Body	Island	0.04739
MORT3	*TNF*	cg26729380	31543655	1stExon	OpenSea	0.04789
GOS12	*TF*	cg20123637	133467435	Body	S_Shelf	0.04794
GOS12	*CALR*	cg27180563	13056485	TSS200	Island	0.04813
GOS3	*TF*	cg16267236	133464463	TSS1500	N_Shore	0.04846
MORT12	*LRP1*	cg10904028	57590920	Body	OpenSea	0.04856
MORT3	*FTH1*	cg19407570	61735476	TSS1500	Island	0.04877
CV	*CP*	cg09457255	148933841	Body	OpenSea	0.04934
MORT12	*CP*	cg19278448	148905466	Body	OpenSea	0.04954

NOTE: Table sorted by smallest p-value; CV, Cerebral vasospasm; DCI, Delayed cerebral ischemia; GOS3, Glasgow Outcome Scale at 3 months; GOS 12, Glasgow Outcome Scale at 12 months; MORT3, Mortality at 3 months; MORT12, Mortality at 12 months; [a] UCSC Genome Browser Build GRCh37/hg19; [b] Trimmed to unique listings, multiple listings indicate information for multiple transcripts; TSS, transcription start site; TSS200 = 0-200 bases upstream of the TSS; TSS1500 = 200-1500 bases upstream of the TSS; 5'UTR = 5' untranslated region, between the TSS and the ATG start site; Body = Between the ATG and stop codon; irrespective of the presence of

introns, exons, TSS, or promoters; 3'UTR = between the stop codon and poly A tail; [c] Shore, 0-2 kb from island; Shelf, 2-4 kb from island; N, upstream (5') of CpG island; S, downstream (3') of CpG island



Appendix Figure 1 Flow Chart for Prioritization of Findings for Replication

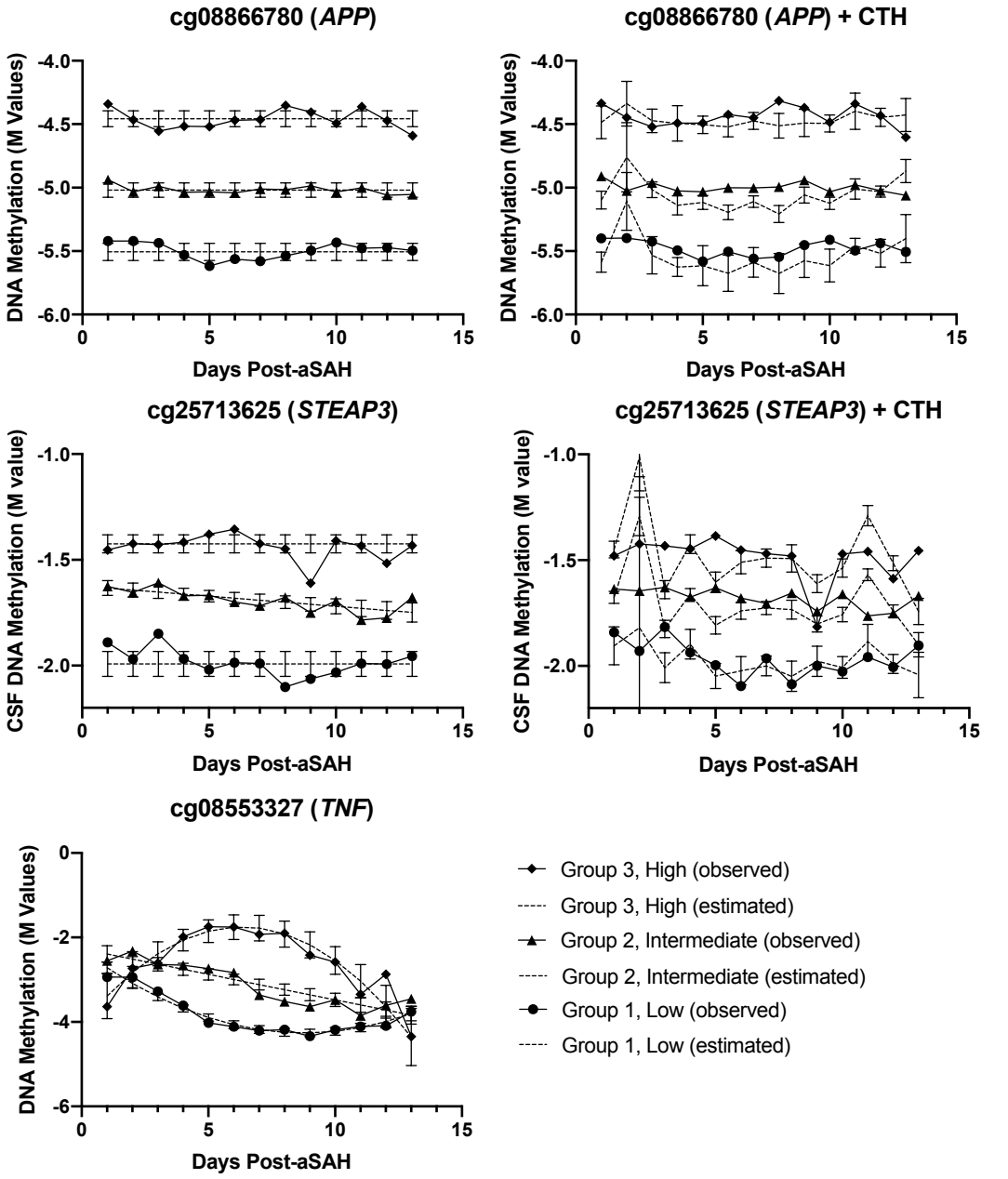
Appendix Table 7 Summary of Prioritization of Top Hits

Candidate gene	CpG Site(s)	Notes	Direction of Effect
APP	cg08866780	Smallest p-value overall; significantly associated with GOS at 12 months and death at 3 and 12 months; suggestively associated with CV and GOS at 3 months; findings persistent after correction for CTH	Three groups (low, intermediate, and high DNA methylation); Increased DNA methylation associated with worse outcomes
TNF	cg12681001; cg21222743; cg10717214; cg21467614; **cg08553327**[a]; cg26729380; cg10650821 [a]Chosen for replication because CpG site had the smallest p-values within the *TNF* gene	Second smallest p-value overall; located in a 'hotspot' of CpGs significant associated with GOS at 3 months and suggestively associated with GOS at 12 months and death at 3 and 12 months; GBTA model at cg08553327 did not pass posterior QC after correction for CTH so unable to compare findings with CTH corrected model, however, two CpG sites within the 'hotspot' did pass posterior QC and findings with GOS at 3 months were persistent after correction for CTH	Three groups (low, intermediate, and high DNA methylation); Intermediate DNA methylation associated with better outcomes compared with both low and high DNA methylation
STEAP3	cg25713625	Third smallest p-value overall; significantly associated with GOS at 3 and 12 months and suggestively associated with death at 3 and 12 months; after correction for CTH, GOS findings were persistent	Three groups (low, intermediate, and high DNA methylation); Increased DNA methylation associated with worse outcomes
SLC25A37	cg18186478	Significantly associated with GOS at 3 months and suggestively associated with GOS at 12 months and death at 3 and 12 months; GBTA model did not pass posterior QC after correction for CTH so unable to compare findings with CTH corrected model	Two groups (low and high DNA methylation); Increased DNA methylation associated with better outcomes
ACO1	cg13977526	Significantly associated with GOS and death at 12 months; suggestively associated with death at 3 months; GBTA model did not pass posterior QC after correction for CTH so unable to compare findings with CTH corrected model	Two groups (low and high DNA methylation); Increased DNA methylation associated with worse outcomes
FXN	cg02268013; cg21229044	Significantly associated with GOS at 3 months and suggestively associated with GOS at 12 months and death at 3 and 12 months; for cg21229044 GBTA controlling for CTH did pass QC, but variability washed out resulting in one group only; GBTA model did not pass posterior QC after correction for CTH for cg02268013 so unable to compare findings with CTH corrected model; odds ratios observed between the two CpG	cg02268013: Two groups (low and high DNA methylation); Increased DNA methylation associated with worse outcomes cg21229044: Three groups (low, intermediate, and high DNA methylation); Increased DNA methylation associated with better outcomes

		sites are in opposite directions with respect to the effect of DNA methylation	
SLC11A2	cg22826226	Significantly associated with GOS at 3 months; suggestively associated with GOS at 12 months and death at 3 and 12 months; GBTA model did not pass posterior QC after correction for CTH so unable to compare findings with CTH corrected model	Two groups (low and high DNA methylation); Increased DNA methylation associated with better outcomes
ACO2	cg25636833	Significantly associated with death at 3 months; suggestively associated with GOS at 3 and 12 months and death at 12 months; findings persistent after correction for CTH	Three groups (low, intermediate, and high DNA methylation); Increased DNA methylation associated with worse outcomes

NOTE: Table sorted by smallest p-value

Appendix Figure 2 DNA Methylation Trajectory Plots for Top Hits



NOTE: Plots of inferred trajectory groups computed using group-based trajectory analysis; plots on the left are unadjusted for cell-type heterogeneity (CTH); plots on the right are adjusted for CTH; the CTH-adjusted trajectory model for cg08553327 did not pass posterior model QC and therefore is not included here; bars plot the 95% confidence interval for the estimate

Appendix Table 8 Participant Characteristics by Trajectory Group for Top Hits

CpG Site	Variable	Unadjusted for CTH			Adjusted for CTH		
		Group 1 (Low)	Group 2 (Intermediate)	Group 3 (High)	Group 1 (Low)	Group 2 (Intermediate)	Group 3 (High)
cg08866780 (*APP*)	Group membership, n (%)	63 (24.2)	139 (53.5)	58 (22.3)	67 (25.8)	140 (53.8)	53 (20.4)
	Age, Mean (SD)	52.7 (10.5)	52.5 (11.0)	54.8 (11.6)	52.0 (10.7)	52.4 (11.2)	56.3 (10.7)
	Sex, Female, n (%)	45 (71.4)	91 (65.5)	43 (74.1)	46 (68.7)	95 (67.9)	38 (71.7)
	Race, White, n (%)	55 (87.3)	120 (86.3)	50 (86.2)	59 (88.1)	122 (87.1)	44 (83.0)
	Fisher Grade						
	2	20 (31.7)	43 (30.9)	15 (25.9)	21 (31.3)	44 (31.4)	13 (24.5)
	3	28 (44.4)	67 (48.2)	31 (53.4)	31 (46.3)	68 (48.6)	27 (50.9)
	4	15 (23.8)	29 (20.9)	12 (20.7)	15 (22.4)	28 (20.0)	13 (24.5)
	Intervention, Coil, n (%)	46 (73.0)	76 (54.7)	33 (56.9)	45 (67.2)	77 (55.0)	33 (62.3)
cg25713625 (*STEAP3*)	Group membership, n (%)	29 (11.1)	166 (63.8)	65 (25.0)	32 (12.3)	160 (61.5)	68 (26.2)
	Age, mean (SD)	52.3 (12.2)	53.0 (10.8)	53.6 (11.2)	52.3 (13.6)	53.0 (10.6)	53.6 (10.7)
	Sex, female, n (%)	13 (44.8)	118 (71.1)	48 (73.8)	15 (46.9)	116 (72.5)	48 (70.6)
	Race, White, n (%)	20 (69.0)	147 (88.6)	58 (89.2)	23 (71.9)	140 (87.5)	62 (91.2)
	Fisher Grade						
	2	6 (20.7)	54 (32.5)	18 (27.7)	10 (31.3)	45 (28.1)	23 (33.8)
	3	14 (48.3)	80 (48.2)	32 (49.2)	13 (40.6)	83 (51.9)	30 (44.1)
	4	9 (31.0)	32 (19.3)	15 (23.1)	9 (28.1)	32 (20.0)	15 (22.1)
	Intervention, Coil, n (%)	16 (55.2)	99 (59.6)	40 (61.5)	18 (56.3)	97 (60.6)	40 (58.8)
cg08553327 (*TNF*)	Group membership, n (%)	175 (67.3)	70 (26.9)	15 (5.8)	NA		
	Age, mean (SD)	53.9 (11.1)	51.1 (10.5)	52.9 (11.9)			
	Sex, female, n (%)	119 (68.0)	48 (68.6)	12 (80.0)			
	Race, White, n (%)	152 (86.9)	59 (84.3)	14 (93.3)			
	Fisher Grade						
	2	51 (29.1)	21 (30.0)	6 (40.0)			
	3	87 (49.7)	33 (47.1)	6 (40.0)			
	4	37 (21.1)	16 (22.9)	3 (20.0)			
	Intervention, Coil, n (%)	114 (65.1)	34 (48.6)	7 (46.7)			

NOTE: Participant characteristics by trajectory group for our top three hits; SD, standard deviation;

Appendix Table 9 Results of Binary Logistic Regression and Global Analysis Exploring Associations of cg08866780 (*APP*) Unadjusted and CTH-Adjusted Trajectory Groups with Patient Outcomes While Controlling for Age, Sex, Race, and Fisher Grade

cg08866780 (*APP*)							
Base Model, Polynomial Order 000							
	Group 2 (Intermediate) vs. Group 1 (Low)			**Group 3 (High) vs. Group 1 (Low)**			
Outcome	**OR**	**95% CI**	***p*[c]**	**OR**	**95% CI**	***p*[c]**	**Global *p*[d]**
CV	0.443	0.19 to 0.989	0.051	0.366	0.14 to 1.00	0.056	0.069
DCI	0.803	0.44 to 1.48	0.479	1.156	0.56 to 2.41	0.698	0.488
GOS3	0.349	0.16 to 0.74	0.0060[b]	0.76	0.32 to 1.82	0.541	0.0118[b]
GOS12	0.194	0.08 to 0.44	**0.0001**[a]	0.826	0.39 to 1.98	0.672	**0.0001**[a]
MORT3	0.209	0.08 to 0.52	**0.0010**[a]	1.080	0.44 to 2.67	0.868	**0.0002**[a]
MORT12	0.206	0.08 to 0.50	**0.0006**[a]	0.977	0.39 to 2.43	0.960	**0.0002**[a]
CTH-Adjusted Model, Polynomial Order 000							
	Group 2 (Intermediate) vs. Group 1 (Low)			**Group 3 (High) vs. Group 1 (Low)**			
Outcome	**OR**	**95% CI**	***p*[c]**	**OR**	**95% CI**	***p*[c]**	**Global *p*[d]**
CV	0.813	0.37 to 1.76	0.602	0.654	0.25 to 1.67	0.374	0.672
DCI	1.230	0.68 to 2.24	0.494	1.400	0.67 to 2.97	0.377	0.656
GOS3	0.332	0.16 to 0.70	0.0040[b]	0.683	0.27 to 1.67	0.4072	0.0097[b]
GOS12	0.347	0.15 to 0.77	0.0097[b]	1.070	0.43 to 2.68	0.8853	0.0066[b]
MORT3	0.359	0.15 to 0.87	0.0237[b]	1.284	0.50 to 3.31	0.6028	0.0091[b]
MORT12	0.337	0.14 to 0.79	0.0128[b]	1.143	0.44 to 2.95	0.7817	0.0072[b]

NOTE: [a] Significant associations (meeting empirical significance level) have been bolded; [b] Suggestive associations (unadjusted p<0.05) have been underlined; [c] Empirical significance threshold for logistic regression analysis=0.001 (calculated based on the minimum of 12 p-values, including both Group 2 vs. Group 1 and Group 3 vs. Group 1 comparisons, in permutation testing); [d] Empirical significance threshold for global analysis=0.002 (calculated based on the minimum of 6 p-values in permutation testing); OR, Odds Ratio; CI, Confidence Interval; CTH, cell type heterogeneity; CV, cerebral vasospasm; DCI, delayed cerebral ischemia; GOS3, Glasgow Outcome Scale at 3-months (poor = 1-3); GOS12, Glasgow Outcome Scale at 12-months (poor = 1-3); MORT3, mortality at 3-months; MORT12, mortality at 12-months; OR, odds ratio; CI, confidence interval; p, p-value based on an alpha of 0.05

Appendix Table 10 Results of Binary Logistic Regression and Global Analysis Exploring Associations of cg25713625 (*STEAP3*) Unadjusted and CTH-Adjusted Trajectory Groups with Patient Outcomes While Controlling for Age, Sex, Race, and Fisher Grade

cg25713625 (*STEAP3*)							
Base Model, Polynomial Order 010							
	Group 2 (Intermediate) vs. Group 1 (Low)			**Group 3 (High) vs. Group 1 (Low)**			
Outcome	**OR**	**95% CI**	***p*[c]**	**OR**	**95% CI**	***p*[c]**	**Global *p*[d]**
CV	0.662	0.21 to 2.03	0.471	1.126	0.33 to 3.80	0.847	0.351
DCI	1.032	0.45 to 2.41	0.941	1.256	0.49 to 3.20	0.629	0.789
GOS3	2.982	0.86 to 12.56	0.100	11.743	3.14 to 21.65	**0.0006**[a]	**0.00005**[a]
GOS12	6.346	1.47 to 31.32	0.029[b]	15.595	3.35 to 22.54	**0.0018**[a]	**0.0005**[a]
MORT3	7.147	1.00 to 32.53	0.076	19.081	3.01 to 25.12	0.0093[b]	**0.0013**[a]
MORT12	4.859	1.00 to 35.80	0.064	12.762	2.71 to 25.13	0.0041[b]	**0.0015**[a]
CTH-Adjusted Model, Polynomial Order 000							
	Group 2 (Intermediate) vs. Group 1 (Low)			**Group 3 (High) vs. Group 1 (Low)**			
Outcome	**OR**	**95% CI**	***p*[c]**	**OR**	**95% CI**	***p*[c]**	**Global *p*[d]**
CV	0.433	0.14 to 1.26	0.131	0.832	0.25 to 2.70	0.760	0.129
DCI	0.612	0.27 to 1.36	0.231	0.913	0.37 to 2.20	0.840	0.260
GOS3	2.589	0.80 to 9.35	0.115	4.143	1.25 to 5.01	0.0271[b]	0.062
GOS12	3.960	1.00 to 10.58	0.055	5.972	1.52 to 6.25	0.0182[b]	0.0326[b]
MORT3	3.768	0.90 to 11.54	0.108	4.847	1.05 to 7.50	0.069	0.122
MORT12	3.120	0.86 to 15.63	0.115	4.249	1.06 to 5.25	0.059	0.123

NOTE: [a] Significant associations (meeting empirical significance level) have been bolded; [b] Suggestive associations (unadjusted p<0.05) have been underlined; [c] Empirical significance threshold for logistic regression analysis=0.003 (calculated based on the minimum of 12 p-values, including both Group 2 vs. Group 1 and Group 3 vs. Group 1 comparisons, in permutation testing); [d] Empirical significance threshold for global analysis=0.003 (calculated based on the minimum of 6 p-values in permutation testing); OR, Odds Ratio; CI, Confidence Interval; CTH, cell type heterogeneity; CV, cerebral vasospasm; DCI, delayed cerebral ischemia; GOS-3, Glasgow Outcome Scale at 3-months (poor = 1-3); GOS-12, Glasgow Outcome Scale at 12-months (poor = 1-3); MORT3, mortality at 3-months; MORT12, mortality at 12-months; OR, odds ratio; CI, confidence interval; p, p-value based on an alpha of 0.05

Appendix Table 11 Results of Binary Logistic Regression and Global Analysis Exploring Associations of cg08553327 (*TNF*) Unadjusted Trajectory Groups with Patient Outcomes While Controlling for Age, Sex, Race, and Fisher Grade

cg08553327 (*TNF*)							
Base Model, Polynomial Order 212							
	Group 2 (Intermediate) vs. Group 1 (Low)			**Group 3 (High) vs. Group 1 (Low)**			
Outcome	**OR**	**95% CI**	***p*[c]**	**OR**	**95% CI**	***p*[c]**	**Global *p*[d]**
CV	0.850	0.42 to 1.72	0.651	1.223	0.31 to 5.36	0.777	0.844
DCI	0.827	0.47 to 1.46	0.516	3.066	0.98 to 11.64	0.069	0.101
GOS3	3.686	1.78 to 7.87	**0.0005**[a]	13.134	3.51 to 39.02	**0.0003**[a]	**0.00001**[a]
GOS12	2.080	0.98 to 4.43	0.0568	4.488	1.24 to 16.50	0.0209[b]	0.0244[b]
MORT3	2.488	1.11 to 5.57	0.0257[b]	5.603	1.47 to 20.44	0.0089[b]	0.0094[b]
MORT12	2.496	1.12 to 5.56	0.0243[b]	7.123	1.89 to 27.94	0.0037[b]	0.0038[b]

NOTE: [a] Significant associations (meeting empirical significance level) have been bolded; [b] Suggestive associations (unadjusted p<0.05) have been underlined; [c] Empirical significance threshold for logistic regression analysis=0.001 (calculated based on the minimum of 12 p-values, including both Group 2 vs. Group 1 and Group 3 vs. Group 1 comparisons, in permutation testing); [d] Empirical significance threshold for global analysis=0.002 (calculated based on the minimum of 6 p-values in permutation testing); OR, Odds Ratio; CI, Confidence Interval; CTH, cell type heterogeneity; CV, cerebral vasospasm; DCI, delayed cerebral ischemia; GOS-3, Glasgow Outcome Scale at 3-months (poor = 1-3); GOS-12, Glasgow Outcome Scale at 12-months (poor = 1-3); MORT3, mortality at 3-months; MORT12, mortality at 12-months; OR, odds ratio; CI, confidence interval; p, p-value based on an alpha of 0.05

Appendix Table 12 Patient Outcome Distributions by Unadjusted Trajectory Groups for cg08866780 (*APP*), cg25713625 (*STEAP3*), and cg08553327 (*TNF*)

	cg08866780 (*APP*)			cg25713625 (*STEAP3*)			cg08553327 (*TNF*)		
	Group 1 n=63 (24.2)	Group 2 n=139 (53.5)	Group 3 n=58 (22.3)	Group 1 n=29 (11.2)	Group 2 n=166 (63.8)	Group 3 n=65 (25.0)	Group 1 n=175 (67.3)	Group 2 n=70 (27.6)	Group 3 n=15 (5.8)
CV, n (%)									
Yes	29 (69.0)	43 (30.9)	19 (32.8)	10 (34.5)	54 (32.5)	27 (41.5)	59 (33.7)	26 (37.1)	6 (40.0)
No	13 (20.6)	41 (29.5)	23 (39.7)	8 (27.6)	52 (31.3)	17 (26.2)	48 (27.4)	25 (35.7)	4 (26.7)
Unknown	21 (33.3)	55 (39.6)	16 (27.6)	11 (37.9)	60 (36.1)	21 (32.3)	68 (38.9)	19 (27.1)	5 (33.3)
DCI, n (%)									
Yes	32 (50.8)	64 (46.0)	31 (53.4)	14 (48.3)	79 (47.6)	34 (52.3)	85 (48.6)	31 (44.3)	11 (73.3)
No	30 (47.6)	75 (54.0)	26 (44.8)	15 (51.7)	86 (51.8)	30 (46.2)	90 (51.4)	37 (52.9)	4 (26.7)
Unknown	1 (1.6)	0	1 (1.7)	0	1 (0.6)	1 (1.5)	0	2 (2.9)	0
GOS3, n (%)									
Poor	25 (39.7)	27 (19.4)	19 (32.8)	4 (13.8)	38 (22.9)	29 (44.6)	34 (19.4)	28 (40.0)	9 (60.0)
Good	30 (47.6)	86 (61.9)	27 (46.6)	19 (65.5)	101 (60.8)	23 (35.4)	106 (60.6)	33 (47.1)	4 (26.7)
Unknown	8 (12.7)	26 (18.7)	12 (20.7)	6 (20.7)	27 (16.3)	13 (20.0)	35 (20.0)	9 (12.9)	2 (13.3)
GOS12, n (%)									
Poor	21 (33.3)	14 (10.1)	18 (31.0)	2 (6.9)	29 (17.5)	22 (33.8)	29 (16.6)	18 (25.7)	6 (40.0)
Good	29 (46.0)	96 (69.1)	26 (44.8)	22 (75.9)	97 (58.4)	32 (49.2)	107 (61.1)	37 (52.9)	7 (46.7)
Unknown	13 (20.6)	29 (20.9)	14 (24.1)	5 (17.2)	40 (24.1)	11 (16.9)	39 (22.3)	15 (21.4)	2 (13.3)
MORT3, n (%)									
Yes	15 (23.8)	9 (6.5)	15 (25.9)	1 (3.4)	21 (12.7)	17 (26.2)	19 (10.9)	15 (21.4)	5 (33.3)
No	42 (66.7)	116 (83.5)	35 (60.3)	23 (79.3)	130 (78.3)	40 (61.5)	135 (77.1)	49 (70.0)	9 (60.0)
Unknown	6 (9.5)	14 (10.1)	8 (13.8)	5 (17.2)	15 (9.0)	8 (12.3)	21 (12.0)	6 (8.6)	1 (6.7)
MORT12, n (%)									
Yes	17 (27.0)	11 (7.9)	16 (27.6)	2 (6.9)	23 (13.9)	19 (29.2)	22 (12.6)	16 (22.9)	6 (40.0)
No	33 (52.4)	99 (71.2)	28 (48.3)	22 (75.9)	103 (62.0)	35 (53.8)	114 (65.1)	39 (55.7)	7 (46.7)
Unknown	6 (9.5)	14 (10.1)	8 (13.8)	5 (17.2)	15 (9.0)	8 (12.3)	21 (12.0)	6 (8.6)	1 (6.7)

NOTE: CV, Cerebral Vasospasm; DCI, Delayed Cerebral Ischemia; GOS-3, Glasgow Outcome Scale at 3-months (poor = 1-3); GOS-12, Glasgow Outcome Scale at 12-months (poor = 1-3); MORT3, mortality at 3-months; MORT12, mortality at 12-months

Appendix C Manuscript 1: Genetic Variability and Trajectories of DNA Methylation May Support a Role for *HAMP* in Patient Outcomes After Aneurysmal Subarachnoid Hemorrhage

Neurocrit Care (2020) 32:550–563
https://doi.org/10.1007/s12028-019-00787-4

NEUROCRITICAL CARE SOCIETY

ORIGINAL WORK

Genetic Variability and Trajectories of DNA Methylation May Support a Role for *HAMP* in Patient Outcomes After Aneurysmal Subarachnoid Hemorrhage

Lacey W. Heinsberg[1*], Annie I. Arockiaraj[2], Elizabeth A. Crago[3], Dianxu Ren[4], John R. Shaffer[2,5], Paula R. Sherwood[3], Susan M. Sereika[4], Daniel E. Weeks[2,6] and Yvette P. Conley[1,2]

Abstract

Background/Objective: Preclinical evidence suggests that iron homeostasis is an important biological mechanism following aneurysmal subarachnoid hemorrhage (aSAH); however, this concept is underexplored in humans. This study examined the relationship between patient outcomes following aSAH and genetic variants and DNA methylation in the hepcidin gene (*HAMP*), a key regulator of iron homeostasis.

Methods: In this exploratory, longitudinal observational study, participants with verified aSAH were monitored for acute outcomes including cerebral vasospasm (CV) and delayed cerebral ischemia (DCI) and evaluated post-discharge at 3 and 12 months for long-term outcomes of death and functional status using the Modified Rankin Scale (mRS; poor = 3–6) and Glasgow Outcome Scale (GOS; poor = 1–3). Participants were genotyped for two genetic variants, and DNA methylation data were collected from serial cerebrospinal fluid over 14 days post-aSAH at eight methylation sites within *HAMP*. Participants were grouped based on their site-specific DNA methylation trajectory, with and without correcting for cell-type heterogeneity (CTH), and the associations between genetic variants and inferred DNA methylation trajectory groups and patient outcomes were tested. To correct for multiple testing, an empirical significance threshold was computed using permutation testing.

Results: Genotype data for rs10421768 and rs7251432 were available for 241 and 371 participants, respectively, and serial DNA methylation data were available for 260 participants. Acute outcome prevalence included CV in 45% and DCI in 37.1% of the overall sample. Long-term outcome prevalence at 3 and 12 months included poor GOS in 23% and 21%, poor mRS in 31.6% and 27.3%, and mortality in 15.1% and 18.2%, respectively, in the overall sample. Being homozygous for the rs7251432 variant allele was significantly associated with death at 3 months ($p = 0.003$) and was the only association identified that passed adjustment for multiple testing mentioned above. Suggestive associations (defined as trending toward significance, p value < 0.05, but not meeting empirical significance thresholds) were identified between the homozygous variant allele for rs7251432 and poor GOS and mRS at 3 months (both $p = 0.04$) and death at 12 months ($p = 0.02$). For methylation trajectory groups, no associations remained significant after correction for multiple testing. However, for methylation trajectory groups not adjusted for CTH, suggestive associations

*Correspondence: law145@pitt.edu
[1] Department of Health Promotion and Development, School of Nursing, University of Pittsburgh, 440 Victoria Building, 3500 Victoria Street, Pittsburgh, PA 15261, USA
Full list of author information is available at the end of the article

Springer

were identified between cg18149657 and poor GOS and mRS at 3 months ($p = 0.003$ and $p = 0.04$, respectively) and death at 3 months ($p = 0.04$), and between cg26283059 and DCI ($p = 0.01$). For methylation trajectory groups adjusted for CTH, suggestive associations were identified between cg02131995 and good mRS at 12 months ($p = 0.02$), and between cg26283059 and DCI ($p = 0.01$).

Conclusions: This exploratory pilot study offers preliminary evidence that *HAMP* may play a role in patient outcomes after aSAH. Replication of this study and mechanistic investigation of the role of *HAMP* in patient outcomes after aSAH are needed.

Keywords: Hepcidins, Polymorphism, DNA methylation, Subarachnoid hemorrhage, Patient outcome assessment

Aneurysmal subarachnoid hemorrhage (aSAH) is a substantial public health problem affecting approximately 30,000 people in the USA annually, who are on average 55 years of age [1], and is a leading contributor to loss of productive life-years [2]. There are functional, emotional, and psychological strains for survivors and their families, as well as a profound impact on personal and public healthcare expenditures [1–3]. Among survivors, recovery is variable. Unfortunately, healthcare providers are often unable to identify patients at risk of poor outcomes after aSAH [4].

Evidence suggests that iron homeostasis may be an important predictor of aSAH outcomes [5]. In aSAH, blood accumulates in the subarachnoid space and is catabolized by heme-oxygenase (inducted in response to cellular stress) into carbon monoxide, biliverdin, and labile free iron [6]. Under normal physiologic conditions, iron is tightly bound to carrier proteins as ferric iron and is recycled within the body [7]. However, because of physiologic changes associated with aSAH, including acidic brain tissue pH, hypoxia, and an influx of catecholamines in the extracellular fluid, iron is liberated to the less stable, non-protein-bound ferrous form [8]. This non-protein-bound iron has the potential to damage nearby tissues. The hepcidin antimicrobial peptide gene (*HAMP*) provides instructions for the production of a peptide called hepcidin [7], which is often referred to as the "master" of iron homeostasis. Although primarily synthesized and secreted by hepatocytes, hepcidin mRNA has been detected in normal brain tissue [9] and systematically regulates plasma levels of iron posttranslationally via interaction with its receptor, ferroportin [10].

Hepcidin levels have recently emerged as an important regulator in brain iron homeostasis [11], and there is some evidence that increased hepcidin expression is associated with poor health outcomes after neurologic injury [12, 13]. In preclinical models, increased hepcidin expression in the brain secondary to aSAH has been associated with lower neurological outcome scores [12]. Likewise, in humans, increased serum hepcidin levels have been associated with poor outcomes following intracranial hemorrhage [13]. We hypothesize that patients with aSAH may not uniformly respond to iron overload or potential therapeutics to mitigate the effects of iron overload potentially due to genetic variability in *HAMP*.

Moreover, because DNA methylation is extensively involved in oxidative stress responses [14], a common pathological complication associated with aSAH, variability of *HAMP* DNA methylation could impact a patient's ability to handle a large influx of iron following aSAH. Notably, DNA methylation of *HAMP* within the cerebrospinal fluid (CSF) after aSAH has not been studied previously. Examining DNA methylation within the CSF, which is a more proximal tissue to aneurysm rupture (compared with traditionally studied blood), may offer a unique perspective into the pathophysiology of aSAH.

If a relationship between genetic or epigenetic variability of *HAMP* and outcomes in the aSAH population can be proven, it may be possible to predict who is at risk of poorer outcomes and to develop interventions that target *HAMP* to improve outcomes. It would also be helpful to know whether variability in *HAMP* impacts long-term outcomes directly or indirectly via acute aSAH outcomes such as cerebral vasospasm (CV) and delayed cerebral ischemia (DCI). To our knowledge, no studies to date have examined genetic and epigenetic variability of *HAMP* and their relationships to the development of acute outcomes of CV and DCI, and long-term outcomes of death and functional status following aSAH in humans. The purpose of this exploratory pilot study is to addresses this knowledge gap by investigating these relationships.

Methods

Study Design

This study was a longitudinal, observational, candidate gene polymorphism and DNA methylation association study that capitalized on existing genomic, clinical, and outcome data collected from a large cohort of aSAH participants. This study examined outcomes in the acute phase (inpatient stay from day of insult up to 14 days post-injury) and long-term outcome phase (interview at 3 and 12 months).

Setting and Sample

Participants were prospectively recruited between 2000 and 2013 from UPMC Presbyterian neurovascular intensive care unit in Pittsburgh, PA. Participants were eligible if they were diagnosed via cerebral angiogram with aSAH, between 18 and 75 years of age, and able to read and speak English (to facilitate functional status survey). Participants were excluded if the cause of their SAH was trauma, arteriovenous malformation, mycotic aneurysm, or unknown. Informed consent was obtained from all individual participants included in this study.

All participants were managed according to aSAH practice guidelines which included prophylactic nimodipine; blood pressure management to maintain adequate systolic pressure (< 140 mmHg before aneurysm securing and > 140 mmHg after aneurysm repair); invasive intracranial pressure monitoring as clinically indicated; and normovolemia guided by fluid balance. Sample sizes for each portion of the study vary as described below.

Genotype Data Collection

Single nucleotide polymorphisms (SNPs) tagging the *HAMP* gene were selected using the National Institute of Environmental Health Sciences Linkage Disequilibrium Tag SNP Selection database [15]. A linkage disequilibrium threshold of $r^2 = 0.8$ was designated in addition to a *5-prime* and *3-prime* flanking region of 2000 base pairs. Two tagging SNPs (rs10421768 and rs7251432) were identified as representing the genetic variability in *HAMP* [16].

DNA was extracted from peripheral leukocytes using a simple salting out procedure [17]. Genome-wide data were collected for a subset of study participants with Affymetrix Gene Chip Assay SNP 6.0 (Affymetrix, Santa Clara, CA, USA) as previously described [18]. Genotype data for SNP rs10421768 were extracted from these data. Because data for SNP rs7251432 were not available in the genome-wide data, this SNP was genotyped using an ABI Prism® Sequence Detection System (Applied Bioscience, Carlsbad, CA, USA) to conduct allelic discrimination using a TaqMan® assay. Duplicate controls on each plate were included to assess consistency and integrity of genotyping data, and data were double-called by two individuals blinded to phenotype data. Genotyping success rates for rs10421768 and rs7251432 were 93.1% and 96.3%, respectively. Given the low amount of ancestral diversity in our sample and that allelic frequencies in *HAMP* differ based on ancestry, non-Caucasian participants were removed from SNP analyses in an attempt to control for population substructure [19].

DNA Methylation Data Collection

Genome-wide DNA methylation data were collected on a subset of study participants using bisulfite converted DNA extracted from biosamples of bagged CSF collected from a ventricular drain changed every 12 h on days 0 to 14 following aSAH. Unlike protein biomarkers, DNA methylation is a chemical modification that is highly stable after biospecimen collection and methods of biosample collection, time to processing, and storage of DNA confer insignificant effects on DNA methylation [20]. The CSF samples were centrifuged, and the cellular component separated from the supernatant fluid and stored in a −80° freezer until DNA extraction from the cellular component. DNA extraction was performed with the Qiagen Midi kit (Qiagen, Valencia, CA, USA).

DNA methylation data were collected with the Infinium Human Methylation450 Beadchip and scanned using an Illumina iSCAN (Illumina, Incorporated, San Diego, CA, USA) by the Center for Inherited Disease Research. All samples from the same individual were on the same chip and technical replicates were included to assess data reliability. The raw data were analyzed using Genome Studio Software (Illumina, Incorporated, San Diego, CA, USA), and data cleaning and quality control were performed. Our data-cleaning pipeline was implemented in R using functions from several packages, including Minfi [21] and ENmix [22], and included identification and exclusion of low-quality probes and poorly performing samples and functional normalization. For each sample, cell-type make-up, as a percentage of five cell types, was estimated using Houseman's reference-free method; [23] these estimates were later used in the analyses adjusted for cell-type heterogeneity.

DNA methylation data within the *HAMP* transcript region ± 2000 kb upstream and downstream (GRCh37/hg19, chr19:35771410:35778045) were extracted from the genome-wide data. A total of eight CpG sites within *HAMP* (cg04668516, cg02131995, cg18149657, cg23677000, cg04085447, cg17907567, cg26283059, and cg27273033) passed quality control procedures and were included in the current analyses. Participants with genome-wide methylation data were included in the current study if they had between two and four CSF DNA methylation measurements (to allow for group-based trajectory analysis) over 14 days following aSAH.

Outcome Phenotyping

Demographic data, including age, sex, and self-reported race, and clinical data, including severity of hemorrhage as measured by Fisher grade, were obtained from the medical record. Clinical data throughout the inpatient stay (through 14 days post-aSAH) were collected by a

trained study nurse. Following discharge, a study staff member trained in neuropsychological testing collected long-term outcome measures during face-to-face interviews or a telephone call with the patient or proxy at 3 and 12 months following aSAH. Interviewers were initially trained and routinely observed for ongoing quality assurance.

A total of eight (two acute and six long-term) outcomes were examined in the current study. The primary acute outcome measures in this study were CV and DCI. CV was defined as the presence of cerebral vessel narrowing of $\geq 25\%$ on cerebral angiogram as evaluated by a neurosurgeon. DCI was defined as the presence of neurological deterioration, after exclusion of non-ischemic causes, which was accompanied by evidence of abnormal cerebral blood flow. In evaluating DCI, neurological deterioration measures were abstracted from the medical record and included any of the following: a decline in level of consciousness as measured by an increase ≥ 2 points on the National Institutes of Health Stroke Scale, a drop in the Glasgow Coma Scale, and/or new and persistent (>1 h) focal neurological deficit. Cerebral blood flow measures used to diagnose DCI included either cerebral angiography performed by the treating physician, or transcranial Doppler performed daily for 14 days by a trained study staff member.

Long-term outcome measures of death and functional status were measured at 3 and 12 months after aSAH. Mortality data were obtained from medical records, caregiver report, or the Social Security Death Index. Global functional outcomes were measured using the Glasgow Outcome Scale (GOS) [24] and the Modified Rankin Scale (mRS) [25]. The GOS rates level of functioning on a scale of 1 (death) to 5 (good recovery) and has a well-established validity in patients with neurological insult [24]. The mRS measures mental and functional deficits using a scale from 0 (no symptoms) to 6 (death) and has a established validity in stroke patients [26, 27]. All study staff involved in data collection were blinded to participant genotype and DNA methylation status.

Statistical Analysis

Demographics and Clinical Characteristics

Statistical analysis was conducted using R statistical software (version 3.4.1, Vienna, Austria) and SAS (version 9.4, SAS Institute Incorporated, Cary, NC, USA). CV, DCI, and death were treated as binary outcomes (occurrence vs. no occurrence), and the mRS and GOS scores were dichotomized as good (mRS 0 to 2; GOS 4 to 5) or poor (mRS 3 to 6; GOS 1 to 3). Standard descriptive statistics were computed. A more detailed statistical methodological narrative is available in the Supplemental Material.

HAMP Single Nucleotide Polymorphisms

All analyses exploring the associations between SNPs and outcomes were conducted using R. Allele frequencies and Hardy–Weinberg Equilibrium (HWE) were evaluated for each SNP. All statistical models used binary logistic regression while adjusting for age, sex, and Fisher grade. The relationship between genetic variability in *HAMP* and patient outcomes was evaluated based on presence (vs. absence) of the minor allele for both SNPs. Given the heavily skewed distribution of genotypes for rs10421768, no further analyses were conducted for this SNP. For rs7251432, genotypic analysis (two-degree-of-freedom test of genetic association) was conducted. Finally, based on results from the genotypic analysis, a model testing a recessive association for rs7251432 was computed.

Given the exploratory nature of this pilot study and correlation between outcomes of interest, correction for multiple testing was made by computing empirical significance thresholds using permutation analysis (10,000 permutations). Associations with p values < 0.05 that did not meet the empirical significance thresholds were considered to be "suggestive" (i.e., trending) associations.

For the genotypic model, a likelihood ratio test (LRT) was used to produce a global p value of the overall model fit by comparing the full model (including the SNP) with a restricted model (omitting the SNP). As described above, a global empirical significance threshold of 0.008 was computed for this global p value to correct for testing eight correlated outcomes.

HAMP DNA Methylation

Group-based Trajectory Analysis Using group-based trajectory analysis (GBTA) as implemented in the PROC TRAJ macro in SAS, the change pattern over time was used to infer distinct trajectory groups for DNA methylation at each CpG site using a censored normal model [28, 29]. DNA methylation data were available on days 0 to 14 following aSAH. Due to limited available information on days 0 and 14 (only 2 and 17 observations were available, respectively), we removed these days from our trajectory analysis. M-values were used for all methylation analyses. GBTA was largely automated for 39 possible models (Supplemental Table 1) with a maximum of three groups and comprehensive combinations of polynomial orders of 0 (intercept-only), 1 (linear), and 2 (quadratic). A secondary evaluation of model adequacy was performed to ensure models moving forward for binary logistic regression analysis met quality control standards. An expanded methods section includes a more detailed statistical methodological narrative (Supplemental Material) in addition to explanatory flowcharts of DNA methylation analysis workflow (Supplemental Figs. 1a, 1b, and 1c).

Cell-Type Heterogeneity Cell-type heterogeneity (CTH) is an important consideration in methylation studies as differential methylation between cell types may impact overall methylation levels and confound results. In cases where genome-wide data are collected, there are methods to control for CTH [30]. However, if only CpG site-specific DNA methylation data are collected, as would be the case with a biomarker used clinically, controlling for CTH is not possible. In an effort to explore the potential clinical utility of methylation values unadjusted for CTH (as they would likely be in the hospital setting) as well as adjusted for CTH, the aforementioned procedures were conducted without adjustment for CTH ("unadjusted models") and with adjustment for CTH ("CTH-adjusted models") for all eight CpG sites. As described above, using previously collected genome-wide data, CTH data were computed using the R statistical software and Houseman's reference-free method [23].

Trajectory Group—Patient Outcome Associations The relationship between site-specific DNA methylation trajectory groups and patient outcomes were examined in R using binary logistic regression while controlling for age, sex, race, and Fisher grade. As stated above, p values < 0.05 were considered suggestive; however, permutation analysis was used to compute an empirical significance threshold of 0.002, which corrects for testing eight correlated outcomes for trajectory group–patient outcome associations. Again, models were evaluated using the LRT to produce a global p value to determine the overall effect of the predicted trajectory group membership within the model. A global empirical significance threshold of 0.001 was computed for this global p value to correct for testing eight correlated outcomes.

Results

Demographics and Clinical Characteristics

Because this study capitalized on existing data as previously described, sample sizes for each portion of the study varied. In our sample of 591 participants, genotype data for rs10421768 and rs7251432 were available for 241 and 371 Caucasian participants, respectively, and DNA methylation data were available for 260 participants (Table 1). The final sample sizes differed between analyses based on outcome data availability and analysis-specific sample sizes are presented (Tables 2, 3, 4 and 5). The overall sample was 93.4% Caucasian and 71.6% female with a mean(± SD) age of 53.4 ± 11.4 years. Fisher grades of 2, 3, and 4 comprised 40.8%, 43.0% and 16.2% of the overall sample, respectively. World Federation Neurosurgical Societies (WFNS) grading system scores of 1, 2, 3, 4, and 5 comprised 51.6%, 17.3%, 7.1%, 13.5%, and 10.5% of the overall sample, respectively. Of the overall sample, 37.7% had surgical intervention (vs. coil embolization). Older age was associated with poor GOS at 3 and 12 months ($p = 0.04$ and $p = 0.03$, respectively) and death at 3 months ($p = 0.02$) in the overall sample. No significant associations between sex or intervention (surgical vs. coil embolization) and outcome measures were found; however, sex was included in all logistic regression models as a covariate given the presence of an estrogen response element governing hepcidin expression

Table 1 Demographics and clinical characteristics

Variable	Overall sample $N = 591$	rs10421768 $n = 241$	rs7251432 $n = 371$	DNA methylation $n = 260$
Age, mean years (SD)	53.4 (11.4)	54.3 (11.4)	53.5 (11.3)	53.06 (11.0)
Sex, female (*n*, %)	423 (71.6)	163 (67.6)	255 (68.7)	179 (68.8)
Race, Caucasian (*n*, %)	552 (93.4)	241 (100)	371 (100)	225 (86.5)
Treatment, surgery (*n*, %)	223 (37.7)	86 (35.7)	117 (31.5)	101 (38.8)
Fisher grade (n, %)				
2	241 (40.8)	69 (28.6)	137 (36.9)	78 (30.0)
3	254 (43.0)	128 (53.1)	172 (46.4)	126 (48.5)
4	96 (16.2)	44 (18.3)	62 (16.7)	56 (21.5)
WFNS (n, %)				
1	305 (51.6)	131 (54.4)	195 (52.6)	108 (41.5)
2	102 (17.3)	42 (17.4)	67 (18.1)	57 (21.9)
3	42 (7.1)	13 (5.4)	26 (7.0)	23 (8.8)
4	80 (13.5)	29 (12.0)	48 (12.9)	44 (16.9)
5	62 (10.5)	26 (10.8)	35 (9.4)	28 (10.9)

WFNS, World Federation of Neurosurgical Societies System

[31]. Non-Caucasian race was associated with poor mRS at 3 months ($p=0.04$) and poor GOS at 12 months ($p=0.02$). Higher Fisher grade was associated with poor acute outcomes of CV ($p=0.006$) and DCI ($p=0.04$) as well as with poor long-term outcomes of GOS, mRS, and death measured at 3 and 12 months ($p<0.001$ for all outcome measures and time points).

HAMP Single Nucleotide Polymorphisms

Allelic Associations

The association between genetic variability in *HAMP* and patient outcomes was first evaluated based on presence (vs. absence) of the minor allele (Supplemental Table 3). For tagging SNPs rs10421768 (minor allele, G) and rs7251432 (minor allele, A), the minor allele frequencies were 23.0% and 43.1%, respectively. No significant associations were found between the presence of the minor allele and patient outcomes for either SNP after controlling for age, sex, and Fisher grade. A figure depicting the locations of these SNPs is presented (Supplemental Fig. 2).

Genotypic Associations

rs10421768 The distribution of rs10421768 genotypes showed AA as the most common (56.8%), followed by AG (40.2%), and GG (2.9%). In our sample, HWE was violated for rs10421768 ($p=0.036$). Given the sparse sample size for the GG genotype, we were not able to perform statistically meaningful subgroup analyses based on genotype for rs10421768.

rs7251432 The distribution of rs7251432 genotypes showed AG as the most common (41.5%), followed by GG (36.1%), and AA (22.4%). In our sample, HWE was violated ($p=0.003$). Next, we tested associations between rs7251432 genotype and patient outcomes in a genotypic model (Table 2). Compared with the GG genotype, suggestive associations (defined as trending toward significance with a p value <0.05 but not meeting empirical significance thresholds) were identified between the AA genotype and mRS at 3 months ($p=0.05$) and death at 3 months ($p=0.02$) after adjusting for age, sex, and Fisher grade (with odds ratios consistent with a recessive model). However, these p values did not meet the empirical significance threshold of 0.005. No differences in patient outcomes were identified between the AG genotype and the GG genotype.

The global test of SNP association within the models based on the LRT identified a suggestive overall effect for only mortality at 3 months (global $p=0.01$). However, this global p value did not meet the global empirical significance threshold of $p=0.008$ calculated in permutation testing.

Recessive Associations

rs7251432 Next, we tested for association between rs7251432 and patient outcomes under a recessive model (Table 3). Compared with combined GG and AG genotype groups, the AA genotype group was significantly associated with death at 3 months ($p=0.003$) and sugges-

Table 2 Results from binary logistic regression and global analysis exploring associations of rs7251432 genotype with patient outcomes while controlling for age, sex, and Fisher grade

Outcome	*N*	*n*, cases	*n*, controls	rs7251432						Global p^c
				AG vs. GG			*AA vs. GG*			
				OR	*95% CI*	p^b	OR	*95% CI*	p^b	
CV	185	83	102	1.43	0.72–2.86	0.30	1.91	0.87–4.26	0.11	0.25
DCI	325	109	216	0.87	0.51–1.48	0.62	1.08	0.57–2.01	0.81	0.79
GOS-3	277	72	205	0.74	0.37–1.45	0.38	1.71	0.82–3.59	0.15	0.09
GOS-12	264	60	204	0.69	0.33–1.43	0.32	1.60	0.73–3.54	0.24	0.12
mRS-3	277	87	190	1.12	0.60–2.09	0.73	2.01	1.00–4.09	0.05	0.13
mRS-12	264	76	188	0.84	0.43–1.63	0.61	1.55	0.74–3.24	0.24	0.27
Death-3	277	48	229	0.78	0.34–1.77	0.56	2.65	1.17–6.08	0.02[a]	0.01[a]
Death-12	264	55	209	0.74	0.34–1.59	0.44	2.00	0.89–1.59	0.10	0.06

[a] Suggestive association (unadjusted $p<0.05$)

[b] Empirical significance threshold for logistic regression analysis = 0.005 (calculated based on the minimum of 16 p values, including both the AG versus GG and AA versus GG comparisons, in permutation testing)

[c] Empirical significance threshold for global analysis = 0.008 (calculated based on the minimum of 8 p values in permutation testing)

Cases indicate the number of participants with poor outcomes (e.g., occurrence of CV or DCI, poor GOS or mRS, or death)

CI, confidence interval; CV, cerebral vasospasm; Death-3, mortality at 3 months; Death-12, mortality at 12 months; DCI, delayed cerebral ischemia; GOS-3, Glasgow Outcome Scale at 3 months (poor = 1–3); GOS-12, Glasgow Outcome Scale at 12 months (poor = 1–3); mRS-3, Modified Rankin Scale at 3 months (poor = 3–6); mRS-12, Modified Rankin Scale at 12 months (poor = 3–6); OR, odds ratio; p, p value based on alpha of 0.05

Table 3 Results from binary logistic regression analysis exploring associations of rs7251432 collapsed genotype (recessive model) with patient outcomes while controlling for age, sex, and Fisher grade

Outcome	*N*	*n*, cases	*n*, controls	rs7251432		
				AA genotype vs. reference group combined GG + AG genotypes		
				OR	95% CI	*p*[b]
CV	185	83	102	1.60	0.78–3.31	0.20
DCI	325	109	216	1.15	0.65–2.02	0.62
GOS-3	277	72	205	1.99	1.01–3.86	0.04[a]
GOS-12	264	60	204	1.93	0.94–3.91	0.07
mRS-3	277	87	190	1.90	1.01–3.56	0.04[a]
mRS-12	264	76	188	1.69	0.87–3.26	0.12
Death-3	277	48	229	2.98	1.43–6.17	0.003[a,c]
Death-12	264	55	209	2.32	1.11–4.79	0.02[a]

[a] Suggestive association (unadjusted $p < 0.05$)

[b] Empirical significance threshold for logistic regression analysis = 0.01 (calculated based on the minimum of 8 *p* values in permutation testing)

[c] Results meeting empirical significance threshold of 0.01

Cases indicate the number of participants with poor outcomes (e.g., occurrence of CV or DCI, poor GOS or mRS, or death)

CI, confidence interval; CV, cerebral vasospasm; Death-3, mortality at 3 months; Death-12, mortality at 12 months; DCI, delayed cerebral ischemia; GOS-3, Glasgow Outcome Scale at 3 months (poor = 1–3); GOS-12, Glasgow Outcome Scale at 12 months (poor = 1–3); mRS-3, Modified Rankin Scale at 3 months (poor = 3–6); mRS-12, Modified Rankin Scale at 12 months (poor = 3–6); OR odds ratio; *p*, *p* value based on alpha of 0.05

tively associated with poor GOS at 3 months ($p = 0.04$), poor mRS at 3 months ($p = 0.04$), and death at 12 months ($p = 0.02$) after adjusting for age, sex, and Fisher grade. Only the *p* value for death at 3 months met the empirical significance threshold of 0.01.

HAMP DNA Methylation

Group-based Trajectory Analysis

Based on the GBTA procedures outlined (Supplemental Fig. 1), distinct methylation trajectory groups passing model adequacy assessment were inferred at six CpG sites for unadjusted models (cg02131995, cg17907567, cg18149657, cg27273033, cg04668516, and cg26283059) and at all eight CpG sites for CTH-adjusted models (Supplemental Table 2). Of the 14 candidate models presented, three could not be tested for association with outcomes because only one trajectory group was inferred at those CpG sites.

Trajectory Group—Patient Outcome Associations

Out of the 11 candidate models eligible for binary logistic regression (five unadjusted models and six CTH-adjusted models), we found suggestive associations between trajectory groups and patient outcomes in four models (two unadjusted models and two CTH-adjusted models). Given the breadth of this study, we have only discussed the CpG sites with suggestive patient outcome associations below. A figure depicting the locations of these CpG sites is presented (Supplemental Fig. 2).

Unadjusted models Binary logistic regression analysis exploring associations of unadjusted methylation trajectory groups with patient outcomes was performed (Table 4). Trajectory groups at two CpG sites (cg18149657 and cg26283059) had suggestive associations (unadjusted $p < 0.05$) with patient outcomes after aSAH. For both cg18149657 and cg26283059, three distinct trajectory groups were identified, and trajectory plots are presented (Fig. 1). At cg18149657, predicted membership assignment to Group 3 (vs. reference Group 1) was suggestively associated with poor GOS and mRS at 3 months ($p = 0.003$ and $p = 0.04$ respectively) and death at 3 months ($p = 0.04$) after controlling for age, sex, race, and Fisher grade, while at cg26283059, predicted membership assignment to Group 3 (vs. reference Group 1) was suggestively associated with DCI ($p = 0.01$). However, these results did not meet the empirical significance threshold for unadjusted models of 0.002 calculated in permutation testing. No other associations were identified between unadjusted methylation trajectory groups and patient outcomes after aSAH (data not shown).

Based on the global test of trajectory group significance within the models using the LRT, a suggestive overall effect was identified for only GOS at 3 months at cg18149657 (global $p = 0.004$). At cg26283059, a suggestive overall effect was identified for not only DCI (global $p = 0.003$), but also CV (global $p = 0.02$). However, these global *p* values did not meet the global empirical significance threshold for unadjusted models of $p = 0.001$ calculated in permutation testing.

Table 4 Results from binary logistic regression and global analysis exploring associations of unadjusted methylation trajectory groups with patient outcomes while controlling for age, sex, race, and Fisher grade

Outcome	*N*	*n*, cases	*n*, controls	cg18149657 (Unadjusted Model with Polynomial Order 222)						
				Group 2 (10.4%) versus reference group 1 (83.9%)			Group 3 (5.8%) versus reference group 1 (83.9%)			Global p^c
				OR	95% CI	p^b	OR	95% CI	p^b	
CV	179	36	143	0.56	0.21–1.41	0.23	1.86	0.45–9.40	0.41	0.29
DCI	249	116	133	0.80	0.34–1.83	0.60	1.17	0.38–3.61	0.78	0.82
GOS-3	209	72	137	1.97	0.71–5.4	0.19	8.40	2.21–41.60	0.003[a]	0.004[a]
GOS-12	199	53	146	1.58	0.57–4.13	0.36	2.23	0.60–7.89	0.21	0.35
mRS-3	209	83	126	1.57	0.59–4.18	0.36	3.89	1.10–16.03	0.04[a]	0.08
mRS-12	199	67	132	2.32	0.94–5.73	0.07	2.36	0.68–8.26	0.17	0.09
Death-3	209	39	170	1.56	0.45–4.68	0.45	3.86	1.006–14.30	0.04[a]	0.12
Death-12	199	46	153	1.65	0.57–4.44	0.33	2.86	0.76–10.37	0.11	0.22
Outcome	***N***	***n*, cases**	***n*, controls**	**cg26283059 (Unadjusted Model with Polynomial Order 000)**						
				Group 2 (65.8%) versus reference group 1 (6.5%)			**Group 3 (27.7%) versus reference group 1 (6.5%)**			**Global p^c**
				OR	**95% CI**	**p^b**	**OR**	**95% CI**	**p^b**	
CV	179	36	143	0.87	0.22–3.44	0.84	2.39	0.56–10.34	0.23	0.02[a]
DCI	249	116	133	1.84	0.64–6.15	0.28	4.65	1.48–16.66	0.01[a]	0.003[a]
GOS-3	209	72	137	2.55	0.68–12.70	0.20	2.21	0.54–11.65	0.30	0.39
GOS-12	199	53	146	1.53	0.41–7.54	0.55	1.20	0.29–6.33	0.81	0.72
mRS-3	209	83	126	1.23	0.39–4.17	0.73	0.98	0.28–3.63	0.98	0.80
mRS-12	199	67	132	1.08	0.33–3.91	0.90	0.74	0.20–2.93	0.66	0.60
Death-3	209	39	170	4.04	0.70–77.97	0.20	3.83	0.60–76.22	0.23	0.32
Death-12	199	46	153	2.07	0.48–14.45	0.38	2.20	0.47–16.21	0.36	0.62

[a] Suggestive association (unadjusted $p < 0.05$)

[b] Empirical significance threshold for logistic regression analysis = 0.002 (calculated based on the minimum of 16 *p* values, including both Group 2 vs. Group 1 and Group 3 vs. Group 1 comparisons, in permutation testing)

[c] Empirical significance threshold for global analysis = 0.001 (calculated based on the minimum of 8 p values in permutation testing)

Cases indicate the number of participants with poor outcomes (e.g., occurrence of CV or DCI, poor GOS or mRS, or death)

CI, confidence interval; CV, cerebral vasospasm; Death-3, mortality at 3 months; Death-12, mortality at 12 months; DCI, delayed cerebral ischemia; GOS-3, Glasgow Outcome Scale at 3 months (poor = 1–3); GOS-12, Glasgow Outcome Scale at 12 months (poor = 1–3); mRS-3, Modified Rankin Scale at 3 months (poor = 3–6); mRS-12, Modified Rankin Scale at 12 months (poor = 3–6); OR odds ratio; *p*, *p* value based on alpha of 0.05

CTH-adjusted models Binary logistic regression analysis exploring associations of CTH-adjusted methylation trajectory groups with patient outcomes was performed (Table 5). Trajectory groups at two CpG sites (cg02131995 and cg26283059) had suggestive associations with patient outcomes after aSAH. For both cg02131995 and cg26283059, three distinct trajectory groups were identified, and trajectory plots are presented (Fig. 2). At cg02131995, predicted membership assignment to Group 2 (vs. reference Group 1) was suggestively associated with good mRS at 12 months ($p = 0.02$), while at cg26283059, predicted membership assignment to Group 3 (vs. reference Group 1) was suggestively associated with DCI ($p = 0.01$). However, these results did not meet the empirical significance threshold for CTH-adjusted models of 0.002 calculated in permutation testing. No other associations were identified between CTH-adjusted methylation trajectory groups and patient outcomes after aSAH (data not shown).

Based on the global test of trajectory group significance within the models using the LRT, a suggestive overall effect was identified for mRS at 12 months at cg02131995 (global $p = 0.05$), whereas a suggestive overall effect was identified for DCI at cg26283059 (global $p = 0.01$). However, these global *p* values did not meet the global empirical significance threshold for unadjusted models of $p = 0.001$ calculated in permutation testing.

Table 5 Results from binary logistic regression and global analysis exploring associations of cth-adjusted methylation trajectory groups with patient outcomes while controlling for age, sex, race, and Fisher grade

Outcome	*N*	*n*, cases	*n*, controls	cg02131995 (CTH-adjusted Model with Polynomial Order 000)						Global p^c
				Group 2 (30%) versus reference group 1 (59.2%)			Group 3 (10.8%) versus reference group 1 (59.2%)			
				OR	95% CI	p^b	OR	95% CI	P^b	
CV	179	36	143	0.75	0.38–1.47	0.40	1.30	0.49–3.54	0.60	0.52
DCI	249	116	133	0.73	0.41–1.29	0.28	2.07	0.89–5.06	0.10	0.08
GOS-3	209	72	137	0.77	0.37–1.56	0.47	0.94	0.34–2.50	0.91	0.77
GOS-12	199	53	146	0.51	0.22–1.11	0.10	0.77	0.26–2.08	0.61	0.24
mRS-3	209	83	126	1.10	0.56–2.13	0.78	1.43	0.56–3.65	0.45	0.75
mRS-12	199	67	132	0.42	0.19–0.85	0.02[a]	0.84	0.32–2.13	0.72	0.05
Death-3	209	39	170	0.53	0.20–1.26	0.17	0.75	0.22–2.23	0.63	0.34
Death-12	199	46	153	0.47	0.19–1.08	0.08	0.70	0.22–1.95	0.51	0.20
Outcome	***N***	***n*, cases**	***n*, controls**	**cg26283059 (CTH-adjusted Model with Polynomial Order 000)**						**Global p^c**
				Group 2 (64.6%) versus reference group 1 (10.4%)			**Group 3 (25%) versus reference group 1 (10.4%)**			
				OR	**95% CI**	**p^b**	**OR**	**95% CI**	**P^b**	
CV	179	36	143	0.82	0.28–2.40	0.72	2.02	0.60–6.83	0.25	0.07
DCI	249	116	133	1.55	0.66–3.87	0.33	3.49	1.32–9.77	0.01[a]	0.01[a]
GOS-3	209	72	137	1.29	0.45–4.10	0.65	1.70	0.52–6.02	0.39	0.65
GOS-12	199	53	146	1.18	0.40–4.04	0.77	1.14	0.33–4.32	0.84	0.96
mRS-3	209	83	126	1.05	0.40–2.86	0.92	1.07	0.36–3.29	0.91	0.99
mRS-12	199	67	132	0.86	0.32–2.42	0.76	0.74	0.24–2.35	0.61	0.87
Death-3	209	39	170	2.42	0.61–16.35	0.27	3.47	0.77–25.00	0.14	0.28
Death-12	199	46	153	1.25	0.40–4.83	0.72	1.73	0.48–7.29	0.42	0.65

[a] Suggestive association (unadjusted $p < 0.05$)

[b] Empirical significance threshold for logistic regression analysis = 0.002 (calculated based on the minimum of 16 *p* values, including both Group 2 vs. Group 1 and Group 3 vs. Group 1 comparisons, in permutation testing)

[c] Empirical significance threshold for global analysis = 0.001 (calculated based on the minimum of 8 *p* values in permutation testing)

Cases indicate the number of participants with poor outcomes (e.g., occurrence of CV or DCI, poor GOS or mRS, or death)

CI, confidence interval; CTH, cell-type heterogeneity; CV, cerebral vasospasm; Death-3, mortality at 3 months; Death-12, mortality at 12 months; DCI, delayed cerebral ischemia; GOS-3 Glasgow Outcome Scale at 3 months (poor = 1–3); GOS-12 Glasgow Outcome Scale at 12 months (poor = 1–3); mRS-3 Modified Rankin Scale at 3 months (poor = 3–6); mRS-12, Modified Rankin Scale at 12 months (poor = 3–6); OR, odds ratio; *p*, *p* value based on an alpha of 0.05

Discussion

HAMP Single Nucleotide Polymorphisms

To our knowledge, this is the first demonstration that variability in *HAMP* genotype may be associated with patient outcomes following aSAH in humans. In this exploratory pilot study, individuals with the AA genotype had between two and three times higher odds of death at 3 and 12 months following aSAH, even after accounting for age and injury severity, as well as two times higher odds of poor GOS and mRS at 3 months. Although the associations observed for this SNP were relatively consistent across time points and outcomes, it is important to note that only the association between the AA genotype and death at 3 months remained significant after correction for multiple testing described above. While little is known about rs7251432, this SNP (located in an intronic region of *HAMP*) has been associated with Kawasaki disease and nominal increased risk of abdominal aortic aneurysm (AAA), which are characterized by inflammation of blood vessels [32, 33]. Interestingly, Kawasaki and aSAH have been linked in three case reports of young children [34–36]. Similarly, cerebral aneurysm and AAA have been shown to share biological pathways in their hemodynamic pathogenesis [37].

Additionally, one study showed that the AA genotype of rs7251432 was associated with higher hemoglobin levels in a simulated hypoxic environment compared with GG or AG genotypes [38–40]. While the literature surrounding the effects of hemoglobin concentrations after aSAH is mixed, one theory suggests that lower

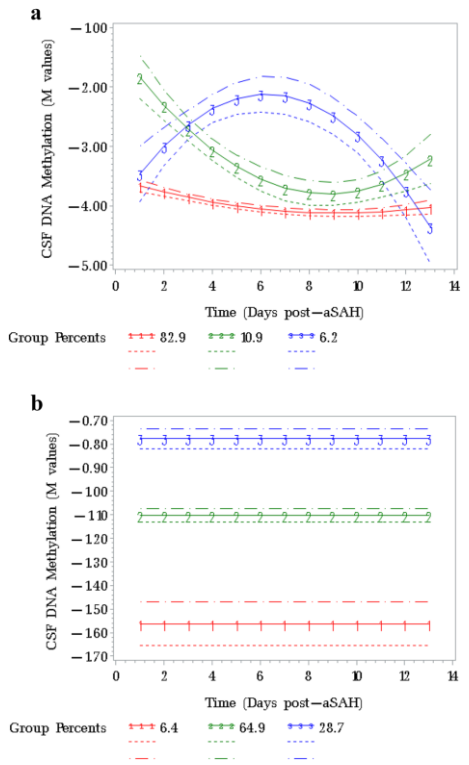


Fig. 1 Methylation Trajectory Group Plots for Unadjusted Models. **a** Methylation Trajectory Group Plot at cg18149657 (Unadjusted Model with Polynomial Order 222[a]), **b** Methylation Trajectory Group Plot at cg26283059 (Unadjusted Model with Polynomial Order 000[a]), [a] Postulated model indicates the number of groups and polynomial orders (trajectory group shapes) of 0 (intercept-only), 1 (linear), or 2 (quadratic). Note: Solid lines represent group-specific average trajectories; dashed lines represent 95% Confidence Interval for group-specific trajectories; group number displayed on group-specific solid fitted lines

hemoglobin levels may offer protective effects, including up-regulation of nitric oxide and decreased blood viscosity which improves brain perfusion [41]. In aSAH, brain tissue oxygen pressure and pH can be altered leading to ischemic changes and secondary injury [42]. This is important because if patients with the rs7251432 AA genotype have higher hemoglobin levels in the hypoxic environment created by aSAH, this could be a potential explanation for the increased odds of death in this subset because they do not have the potential protective response described above. Although this study did not evaluate hemoglobin levels, this is an important future direction of this work.

For rs10421768, we identified modest positive odds ratios of around 1.2 or higher (Supplemental Table 3). Although these modest associations were not significant in our small sample size, there could be a real effect that could be detected in larger sample sizes. Notably, this SNP has been previously identified to play a role in iron homeostasis in cases of iron overload [43] warranting future investigation in a larger sample.

In this sample, both rs10421768 and rs7251432 violated HWE. After a third party reviewed the double-called raw genotype data, we determined that genotyping error was unlikely. A potential explanation for this violation, however, could be that our sample of aSAH cases is enriched for associated variants and may not be representative of the general population, further supporting the potential role of genetic variability of *HAMP* with aSAH.

HAMP DNA Methylation

This study suggests variability in *HAMP* DNA methylation trajectories may be associated with patient outcomes following aSAH, although none of the associations identified survived correction for multiple testing. For the unadjusted model at cg18149657, Group 3 had between two and eight times increased odds of unfavorable 3-month outcomes. This group had the highest methylation levels across groups peaking between days 4 and 10 (Fig. 1a), a key window in recovery after aSAH in which DCI often occurs. However, the global test of trajectory group significance identified only a suggestive association for GOS at 3 months. An interesting finding in this study is that at cg26283059, suggestive associations were found in both the unadjusted model and CTH-adjusted model between Group 3 and DCI. At this site and similar to above, participants in Group 3 had the highest overall methylation levels and had increased odds of DCI for both the unadjusted model and CTH-adjusted model (Figs. 1b and 2b). Based on the global p values for unadjusted methylation trajectory groups at cg26283059, post hoc analyses revealed that by changing the reference group from Group 1 to Group 3, we achieved suggestively significant results for not only DCI, but also CV (data not shown). Again, we observe consistent results that hypermethylation is associated with both acute outcomes supporting our findings further. These results are inconsistent with our expectations that hypomethylation (leading to increased gene expression) versus the observed hypermethylation would be associated with unfavorable outcomes, suggesting other regulatory mechanisms might be at play.

Beyond this study, few studies to date have specifically examined DNA methylation of *HAMP* and of those

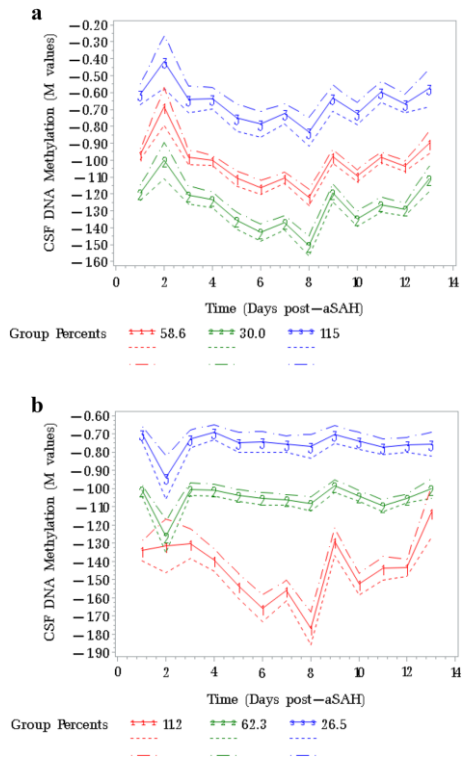


Fig. 2 Methylation Trajectory Group Plots for CTH-Adjusted Models. **a** Methylation Trajectory Group Plot at cg02131995 (CTH-adjusted Model with Polynomial Order 000[a]), **b** Methylation Trajectory Group Plot at cg26283059 (CTH-adjusted Model with Polynomial Order 000[a]). [a]Postulated model indicates the number of groups and polynomial orders (trajectory group shapes) of 0 (intercept-only), 1 (linear), or 2 (quadratic). *CTH* cell-type heterogeneity; solid lines represent group-specific average trajectories; dashed lines represent 95% Confidence Interval for group-specific trajectories; group number displayed on group-specific solid fitted lines; intercept-only polynomials appear jagged due to the inclusion of CTH data as time-varying covariates

that have, none have been in aSAH. However, similar to rs7251432 discussed above, hypomethylation of CpG sites within the *HAMP* promoter region (cg23677000 and cg04085447) has been associated with cases of Kawasaki disease [44]. Recently, it has become more clear that complex signaling pathways, plasma and tissue iron levels, inflammatory cytokines, erythropoiesis, and estrogen all play an important role in hepcidin expression by inducing or inhibiting hepcidin synthesis [31, 45, 46]. For example, endothelins are potent vasoconstrictors released during inflammatory responses that have been shown to be associated with poor outcomes following aSAH [45, 47]. Specific to this study, endothelin-1 has been shown to lead to increases in mRNA levels of iron homeostasis genes (including *HAMP*) in the brain [45]. Also, estrogen has been implicated as a regulator of hepcidin expression via transcriptional inhibition through a functional estrogen response element in the promoter region of the *HAMP* gene [31]. These potential confounders will be important to consider in future work of a larger sample size.

It is important to interpret these results with caution as none of our methylation findings remained significant after correction for multiple testing described above. It is also important to note that methylation was measured in the CSF as opposed to the major site of hepcidin production (the liver) or more commonly examined tissue of blood which makes it difficult to compare the results of this study with existing studies of *HAMP*. Finally, it should be acknowledged that Housman's reference-free method for correcting for CTH [23] has not been validated in CSF and cell count data for associated DNA methylation data were unavailable. Nevertheless, given the proximal location of CSF within the central nervous system, CSF DNA methylation of *HAMP* could have important clinical implications and warrants more robust validation of the potential prognostic performance of these loci.

Limitations

This exploratory pilot study is significant because it suggests that *HAMP* may influence patient outcomes after aSAH; however, several important limitations should be acknowledged. First, by selecting specific DNA methylation targets located within the relatively narrow window of 2000 base pairs upstream and downstream of the *HAMP* gene region, we did not exhaust all potential regulatory regions of *HAMP*. Expanding this region to capture more CpG sites should be an area of future research in a larger study. Similarly, this study focused on genetic and epigenetic variability of *HAMP* and did not examine other loci relevant to the iron homeostasis pathway. While *HAMP* was carefully chosen based on the existing literature and its importance in iron homeostasis, interpretability of the results is limited without considering the intricate genetic network potentially moderating patient outcomes. Similarly, this study did not measure hepcidin or iron levels. Future investigation of hepcidin and iron levels, as well as additional genetic loci and their interactions within the iron homeostasis pathway, is warranted.

Furthermore, by capitalizing on data and samples from an existing cohort, we were limited to existing outcome

and covariate data previously collected. Specifically, there is a large amount of missing data due to loss to follow-up. Because aSAH outcomes vary between and within patients over time, we did not fill in missing data with imputation or by carrying the last observation forward to avoid biasing the results. Similarly, the available CV data used in this study captured only the cases verified with angiogram (vs. ultrasound). Therefore, it is likely that we did not capture a portion of the overall sample who developed CV identified with transcranial Doppler. Similarly, severity of injury is an important predictor of outcomes after aSAH. In the available data, in addition to the Fisher grade, we had access to the WFNS grading system which is a score based on the patient's presenting clinical condition and combines measures from the Glasgow Coma Scale and focal neurological deficits. Both WFNS and Fisher grade have been shown to be important predictors of outcomes after aSAH and are highly correlated in our data. In post hoc analyses, we repeated the analyses while controlling for WFNS (instead Fisher grade) and found that the results were consistent across outcomes and timepoints. In an effort to avoid multicollinearity with our models, we ultimately chose to control for Fisher grade (as opposed to WFNS) because it offers greater insight into the amount of blood within the subarachnoid space and is more scientifically meaningful to the conceptual foundation of this study.

An additional limitation is the relatively small sample size of this study. Specifically, there was an insufficient number of participants with the rs10421768 GG genotype for reliable analysis. Given this underrepresentation, we were unable to perform genotypic association analysis in this sample size. Next, race was self-reported by study participants and use of this variable limits the ability to stringently control for population substructure. Similarly, given the underrepresentation of non-Caucasian populations, we were unable to perform subgroup analyses for the SNP portion of this study in more diverse ancestries, which limits the generalizability of the findings. Future efforts are needed to replicate these findings in larger, more diverse samples. Finally, studies have demonstrated that smoking and body mass index (BMI) impact DNA methylation [48, 49] and that treatment strategies (e.g., induced hypertension, administration of a calcium channel blocker) impact the incidence of acute complications after aSAH. Unfortunately, our small sample size and amount of missing data prevented us from controlling for these factors.

Lastly and of particular notability, the interpretability of the results of this study are limited because of the large number of tests conducted. A simple Bonferroni correction for multiple testing would be too conservative due to the high correlation between outcomes in this cohort. Specifically, CV and DCI have a correlation of >0.75. Likewise, mRS and GOS have a correlation of >0.70 and were assessed at both 3 months and 12 months. Moreover, death is imbedded in both mRS and GOS scores. Despite these correlations, it was ultimately decided to retain both mRS and GOS in this study to allow for between study comparisons in the future. In an effort to carefully adjust for the multiple testing while properly taking the correlation between tests into account, we performed permutation testing to compute an empirical significance threshold within each model examined. While many of our results are consistent across correlated outcomes and time points suggesting possible associations of genetic and epigenetic variability of *HAMP* with patient outcomes after aSAH, it is imperative that this study be replicated in a larger sample to confirm these findings.

Conclusion

Outcomes following aSAH are variable creating an unyielding need for reliable biomarkers of poor outcomes. Hepcidin has surfaced as an important preclinical predictor of outcomes following experimental brain injury [12]. The results of this pilot study offer support to this theory in humans; however, due to the exploratory nature of this study, more research is needed before definitive conclusions can be made about the role of *HAMP* in outcomes after aSAH. In our sample, we identified an association between the AA genotype of rs7251432 and death at 3 months that remained significant after correction for multiple testing. If these findings can be replicated in an independent sample, this approach has the potential to provide a clinical tool for prediction of those at high-risk of poor outcomes following aSAH. We also believe that our exploratory study found interesting trends that may warrant further investigation into the role of DNA methylation of *HAMP* in patient outcomes after aSAH.

Electronic supplementary material

The online version of this article (https://doi.org/10.1007/s12028-019-00787-4) contains supplementary material, which is available to authorized users.

Author details

[1] Department of Health Promotion and Development, School of Nursing, University of Pittsburgh, 440 Victoria Building, 3500 Victoria Street, Pittsburgh, PA 15261, USA. [2] Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA. [3] Department of Acute and Tertiary Care, School of Nursing, University of Pittsburgh, Pittsburgh, PA, USA. [4] Department of Health and Community Systems, School of Nursing, University of Pittsburgh, Pittsburgh, PA, USA. [5] Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA, USA. [6] Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA.

Acknowledgements

We would also like to acknowledge the anonymous reviewers who took the time to critically evaluate this paper. Their detailed feedback greatly improved the clarity and quality of this work for the scientific community.

Author Contributions

LWH contributed to the study conception and design, analysis and interpretation of data, and drafted, critically revised, and gave final approval for the manuscript. AIA contributed to the analysis of data and critically revised and gave final approval for the manuscript. EAC contributed to the acquisition and interpretation of data and critically revised and gave final approval for the manuscript. DR contributed to the analysis and interpretation of data and critically revised and gave final approval for the manuscript. JRS contributed to analysis and interpretation of the data and critically revised and gave final approval for the manuscript. PRS contributed to the acquisition of data and critically revised and gave final approval for the manuscript. SMS contributed to the analysis and interpretation of data and critically revised and gave final approval for the manuscript. DEW contributed to the study design, analysis and interpretation of data, and critically revised and gave final approval for the manuscript. YPC contributed to the study conception and design, interpretation of data, and critically revised and gave final approval for the manuscript. All authors agree to be accountable for all aspects of the work in ensuring that questions relating to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All authors meet authorship criteria, have read and approved the published work, and certify that they have participated sufficiently in the work to take responsibility for the content including the concept, design, analysis, writing, or revision (see author contributions above).

Source of Support

Research reported in this publication was supported by the National Instiute of Nursing Research of the National Institutes of Health under Award Numbers F31NR017311, R01NR004339, R01NR013610, and T32NR009759 and with additional support from the International Society of Nurses in Genetics, University of Pittsburgh Leslie A. Hoffman Endowed Research Award, Nightingale Awards of Pennsylvania, and Eta Chapter, Sigma Theta Tau, Inc. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or supporting foundations.

Conflict of interest

LW Heinsberg reports grants from the National Institute of Nursing Research, University of Pittsburgh, Eta Chapter, Sigma Theta Tau, Inc., and International Society of Nurses in Genetics during the conduct of this study. YP Conley and PR Sherwood reports grants from the National Institute of Nursing Research. EA Crago reports grants from the University of Pittsburgh. The remaining authors report no conflicts of interest to disclose.

Ethical Approval

Institutional review board approval at the University of Pittsburgh has been obtained (IRB Approval Number 021039), and we have adhered to ethical considerations in the protection of all human subjects involved.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Published online: 25 July 2019

References

1. Zacharia BE, Hickman ZL, Grobelny BT, et al. Epidemiology of aneurysmal subarachnoid hemorrhage. Neurosurg Clin N Am. 2010;21(2):221–33. https://doi.org/10.1016/j.nec.2009.10.002.
2. Rosenbaum BP, Kelly ML, Kshettry VR, Weil RJ. Neurologic disorders, in-hospital deaths, and years of potential life lost in the USA, 1988–2011. J Clin Neurosci. 2014;21(11):1874–80. https://doi.org/10.1016/j.jocn.2014.05.006.
3. le Roux AA, Wallace MC. Outcome and cost of aneurysmal subarachnoid hemorrhage. Neurosurg Clin N Am. 2010;21(2):235–46. https://doi.org/10.1016/j.nec.2009.10.014.
4. Lantigua H, Ortega-Gutierrez S, Schmidt JM, et al. Subarachnoid hemorrhage: who dies, and why? Crit Care. 2015;19(1):309. https://doi.org/10.1186/s13054-015-1036-0.
5. Gomes JA, Selim M, Cotleur A, et al. Brain iron metabolism and brain injury following subarachnoid hemorrhage: iCeFISH-pilot (CSF iron in SAH). Neurocrit Care. 2014;21(2):285–93. https://doi.org/10.1007/s12028-014-9977-8.
6. Wagner KR, Sharp FR, Ardizzone TD, Lu A, Clark JF. Heme and iron metabolism: role in cerebral hemorrhage. J Cereb Blood Flow Metab. 2003;23(6):629–52. https://doi.org/10.1097/01.WCB.0000073905.87928.6D.
7. Ganz T. Hepcidin and iron regulation, 10 years later. Blood. 2011;117(17):4425–33. https://doi.org/10.1182/blood-2011-01-258467.
8. Bishop GM, Robinson SR. Quantitative analysis of cell death and ferritin expression in response to cortical iron: implications for hypoxia-ischemia and stroke. Brain Res. 2001;907(1–2):175–87. https://doi.org/10.1016/S0006-8993(01)02303-4.
9. Hänninen MM, Haapasalo J, Haapasalo H, et al. Expression of iron-related genes in human brain and brain tumors. BMC Neurosci. 2009;10(1):36. https://doi.org/10.1186/1471-2202-10-36.
10. Nemeth E, Tuttle MS, Powelson J, et al. Hepcidin regulates cellular iron efflux by binding to ferroportin and inducing its internalization. Science. 2004;306(5704):2090–3. https://doi.org/10.1126/science.1104742.
11. Vela D. Hepcidin, an emerging and important player in brain iron homeostasis. J Transl Med. 2018;16(1):25. https://doi.org/10.1186/s12967-018-1399-5.
12. Tan G, Liu L, He Z, Sun J, Xing W, Sun X. Role of hepcidin and its downstream proteins in early brain injury after experimental subarachnoid hemorrhage in rats. Mol Cell Biochem. 2016;418(1–2):31–8. https://doi.org/10.1007/s11010-016-2730-1.
13. Xiong X-YY, Chen J, Zhu W-YY, et al. Serum hepcidin concentrations correlate with serum iron level and outcome in patients with intracerebral hemorrhage. Neurol Sci. 2015;36(10):1843–9. https://doi.org/10.1007/s10072-015-2266-2.
14. Zhao H, Han Z, Ji X, Luo Y. Epigenetic regulation of oxidative stress in ischemic stroke. Aging Dis. 2016;7(3):295–306. https://doi.org/10.14336/AD.2015.1009.
15. Sciences NI of EH. LD TAG SNP Selection (TagSNP). https://snpinfo.niehs.nih.gov/snpinfo/snptag.html.
16. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. https://analysistools.nci.nih.gov/LDlink/?var1=rs7251432&var2=rs10421768&pop=CEU&tab=ldpair. Accessed 2 Apr 2018.
17. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16(3):1215. https://doi.org/10.1093/nar/16.3.1215.
18. Kim H, Crago E, Kim M, et al. Cerebral vasospasm after sub-arachnoid hemorrhage as a clinical predictor and phenotype for genetic association study. Int J Stroke. 2013;8(8):620–5. https://doi.org/10.1111/j.1747-4949.2012.00823.x.
19. Little J, Higgins JPT, Ioannidis JPA, et al. STrengthening the REporting of Genetic Association studies (STREGA)–an extension of the STROBE statement. *Eur J Clin Invest*. 2009;39(4):247-266. http://www.ncbi.nlm.nih.gov/pubmed/19297801. Accessed 12 Sep 2017.
20. Groen K, Lea RA, Maltby VE, Scott RJ, Lechner-Scott J. Letter to the editor: blood processing and sample storage have negligible effects on methylation. Clin Epigenet. 2018;10(1):22. https://doi.org/10.1186/s13148-018-0455-6.
21. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: a flexible and comprehensive bioconductor package for the analysis of infinium DNA methylation microarrays. Bioinformatics. 2014;30(10):1363–9. https://doi.org/10.1093/bioinformatics/btu049.
22. Xu Z, Niu L, Li L, Taylor JA. ENmix: a novel background correction method for Illumina HumanMethylation450 BeadChip. Nucleic Acids Res. 2016;44(3):e20. https://doi.org/10.1093/nar/gkv907.
23. Houseman EA, Molitor J, Marsit CJ. Reference-free cell mixture adjustments in analysis of DNA methylation data. Bioinformatics. 2014;30(10):1431–9. https://doi.org/10.1093/bioinformatics/btu029.
24. Jennett B, Bond M. Assessment of outcome after severe brain damage. *Lancet (London, England)*. 1975;1(7905):480-484. http://www.ncbi.nlm.nih.gov/pubmed/46957. Accessed 3 Jan 2017.
25. van Swieten JC, Koudstaal PJ, Visser MC, Schouten HJ, van Gijn J. Interobserver agreement for the assessment of handicap in stroke patients.

Stroke. 1988;19(5):604-607. http://www.ncbi.nlm.nih.gov/pubmed/3363593. Accessed 15 May 2017.
26. Banks JL, Marotta CA. Outcomes validity and reliability of the modified rankin scale: implications for stroke clinical trials - A literature review and synthesis. Stroke. 2007;38(3):1091–6. https://doi.org/10.1161/01.STR.0000258355.23810.c6.
27. Van Swieten JC, Koudstaal PJ, Visser MC, Schouten H, Van Gijn J. Inter-observer agreement for the assessment of handicap in stroke patients. Stroke. 1988;19(5):604–7. https://doi.org/10.1161/01.STR.19.5.604.
28. Jones BL, Nagin DS, Roeder K. A SAS procedure based on mixture models for estimating developmental trajectories. Sociol Methods Res. 2001;29(3):374–93. https://doi.org/10.1177/0049124101029003005.
29. Jones BL, Nagin DS. Advances in group-based trajectory modeling and an SAS procedure for estimating them. Sociol Methods Res. 2007;35(4):542–71. https://doi.org/10.1177/0049124106292364.
30. Titus AJ, Gallimore RM, Salas LA, Christensen BC. Cell-type deconvolution from DNA methylation: a review of recent applications. Hum Mol Genet. 2017;26(R2):R216–24. https://doi.org/10.1093/hmg/ddx275.
31. Hou Y, Zhang S, Wang L, et al. Estrogen regulates iron homeostasis through governing hepatic hepcidin expression via an estrogen response element. Gene. 2012;511(2):398–403. https://doi.org/10.1016/j.gene.2012.09.060.
32. Lillvis JH, Kyo Y, Tromp G, et al. Analysis of positional candidate genes in the AAA1 susceptibility locus for abdominal aortic aneurysms on chromosome 19. BMC Med Genet. 2011;12(1):14. https://doi.org/10.1186/1471-2350-12-14.
33. Huang Y-H, Yang KD, Hsu Y-W, et al. Correlation of HAMP gene polymorphisms and expression with the susceptibility and length of hospital stays in Taiwanese children with Kawasaki disease. *Oncotarget*. July 2017. https://doi.org/10.18632/oncotarget.17700.
34. Ishida A, Matsuo S, Kawamura S, Nishikawa T. Subarachnoid hemorrhage due to nonbranching aneurysm of the middle cerebral artery in a young adult with a history of Kawasaki disease. Surg Neurol Int. 2014;5:5. https://doi.org/10.4103/2152-7806.125285.
35. Ahn JH, Phi JH, Kang H-S, et al. A ruptured middle cerebral artery aneurysm in a 13-month-old boy with Kawasaki disease. J Neurosurg Pediatr. 2010;6(2):150–3. https://doi.org/10.3171/2010.5.PEDS1012.
36. Tanaka S, Sagiuchi T, Kobayashi I. Ruptured pediatric posterior cerebral artery aneurysm 9 years after the onset of Kawasaki disease: a case report. Child's Nerv Syst. 2007;23(6):701–6. https://doi.org/10.1007/s00381-006-0263-8.
37. Tanweer O, Wilson TA, Metaxa E, Riina HA, Meng H. A comparative review of the hemodynamics and pathogenesis of cerebral and abdominal aortic aneurysms: lessons to learn from each other. J Cerebrovasc Endovasc Neurosurg. 2014;16(4):335–49. https://doi.org/10.7461/jcen.2014.16.4.335.
38. Gerile W, Yang H, Longyan Y, Jing N, Jingling W, Hongshu J. Association between rs7251432 Polymorphism in hepcidin gene and change in hemogram after the HiHiLo training in men of Han Nationality. *Chin J Sport Med*. 2010;29(4):386-390. http://caod.oriprobe.com/articles/25996307/Association_between_rs7251432_Polymorphism_in_Hepcidin_Gene_and_Change.htm. Accessed June 6, 2017.
39. Chen Y, Jiang C, Luo Y, Liu F, Gao Y. Interaction of CARD14, SENP1 and VEGFA polymorphisms on susceptibility to high altitude polycythemia in the Han Chinese population at the Qinghai-Tibetan Plateau. Blood Cells Mol Dis. 2016;57:13–22. https://doi.org/10.1016/j.bcmd.2015.11.005.
40. Festic E, Rabinstein AA, Freeman WD, et al. Blood transfusion is an important predictor of hospital mortality among patients with aneurysmal subarachnoid hemorrhage. Neurocrit Care. 2013;18(2):209–15. https://doi.org/10.1007/s12028-012-9777-y.
41. Rebel A, Ulatowski JA, Kwansa H, Bucci E, Koehler RC. Cerebrovascular response to decreased hematocrit: effect of cell-free hemoglobin, plasma viscosity, and CO2. Am J Physiol Heart Circ Physiol. 2003;285(4):H1600–8. https://doi.org/10.1152/ajpheart.00077.2003.
42. Hoffman WE, Wheeler P, Edelman G, Charbel FT, Torres NJ, Ausman JI. Hypoxic brain tissue following subarachnoid hemorrhage. *Anesthesiology*. 2000;92(2):442–446. http://www.ncbi.nlm.nih.gov/pubmed/10691231. Accessed 3 April 2018.
43. Andreani M, Radio FC, Testi M, et al. Association of hepcidin promoter c.-582 A > G variant and iron overload in thalassemia major. *Haematologica*. 2009;94(9):1293–1296. https://doi.org/10.3324/haematol.2009.006270.
44. Huang YH, Kuo HC, Li SC, Cai XY, Liu SF, Kuo HC. HAMP promoter hypomethylation and increased hepcidin levels as biomarkers for Kawasaki disease. J Mol Cell Cardiol. 2018;117:82–7. https://doi.org/10.1016/j.yjmcc.2018.02.017.
45. Bickford JS, Ali NF, Nick JA, et al. Endothelin-1-mediated vasoconstriction alters cerebral gene expression in iron homeostasis and eicosanoid metabolism. Brain Res. 2014;1588:25–36. https://doi.org/10.1016/j.brainres.2014.09.022.
46. Liu J, Sun B, Yin H, Liu S. Hepcidin: a promising therapeutic target for iron disorders: a systematic review. Medicine (Baltimore). 2016;95(14):e3150. https://doi.org/10.1097/MD.0000000000003150.
47. Thampatty BP, Sherwood PR, Gallek MJ, et al. Role of endothelin-1 in human aneurysmal subarachnoid hemorrhage: associations with vasospasm and delayed cerebral ischemia. Neurocrit Care. 2011;15(1):19–27. https://doi.org/10.1007/s12028-011-9508-9.
48. Li S, Wong EM, Bui M, et al. Inference about causation between body mass index and DNA methylation in blood from a twin family study. *International Journal of Obesity*. https://www.biorxiv.org/content/early/2017/11/21/223040. Published November 21, 2018. Accessed 26 Aug 2018.
49. Li S, Wong EM, Bui M, et al. Causal effect of smoking on DNA methylation in peripheral blood: a twin and family study. Clin Epigenetics. 2018;10(1):18. https://doi.org/10.1186/s13148-018-0452-9.

Appendix C.1 License Agreement for Manuscript 1

Copyright Transfer Statement (CTS) for your article in Neurocritical Care (787)

Springer <springerauthorquery@springeronline.com>
Wed 7/10/2019 9:14 AM
To: Heinsberg, Lacey Wright <LAW145@pitt.edu>

Copyright Transfer

10.07.2019

visit us at springer.com

Confirmation of your Copyright Transfer

Dear Author,

Please note: This e-mail is a confirmation of your copyright transfer and was sent to you only for your own records.

1. Publication
The copyright to this article, (including any supplementary information and graphic elements therein (e.g. illustrations, charts, moving images) (the 'Article'), is hereby assigned for good and valuable consideration to Springer Science+Business Media, LLC, part of Springer Nature and Neurocritical Care Society (the 'Assignee'). Headings are for convenience only.

2. Grant of Rights
In consideration of the Assignee evaluating the Article for publication, the Author(s) grant the Assignee without limitation the exclusive (except as set out in clauses 3, 4 and 5 a) iv), assignable and sub-licensable right, unlimited in time and territory, to copy-edit, reproduce, publish, distribute, transmit, make available and store the Article, including abstracts thereof, in all forms of media of expression now known or developed in the future, including pre- and reprints, translations, photographic reproductions and extensions. Furthermore, to enable additional publishing services, such as promotion of the Article, the Author(s) grant the Assignee the right to use the Article (including the use of any graphic elements on a stand-alone basis) in whole or in part in electronic form, such as for display in databases or data networks (e.g. the Internet), or for print or download to stationary or portable devices. This includes interactive and multimedia use as well as posting the Article in full or in part or its abstract on social media, and the right to alter the Article to the extent necessary for such use. The Assignee may also let third parties share the Article in full or in part or its abstract on social media and may in this context sub-license the Article and its abstract to social media users. Author(s) grant to Assignee the right to re-license Article metadata without restriction (including but not limited to author name, title, abstract, citation, references, keywords and any additional information as determined by Assignee).

3. Self-Archiving
Author(s) are permitted to self-archive a pre-print and an author's accepted manuscript version of

their Article.

a. A pre-print is the author's version of the Article before peer-review has taken place ("Pre-Print"). Prior to acceptance for publication, Author(s) retain the right to make a Pre-Print of their Article available on any of the following: their own personal, self-maintained website; a legally compliant, non-commercial pre-print server such as but not limited to arXiv and bioRxiv. Once the Article has been published, the Author(s) should update the acknowledgement and provide a link to the definitive version on the publisher's website: "This is a pre-print of an article published in [insert journal title]. The final authenticated version is available online at: https://doi.org/[insert DOI]".
b. An Author's Accepted Manuscript (AAM) is the version accepted for publication in a journal following peer review but prior to copyediting and typesetting that can be made available under the following conditions:
    a. Author(s) retain the right to make an AAM of their Article available on their own personal, self-maintained website immediately on acceptance,
    b. Author(s) retain the right to make an AAM of their Article available for public release on any of the following 12 months after first publication ("Embargo Period"): their employer's internal website; their institutional and/or funder repositories. AAMs may also be deposited in such repositories immediately on acceptance, provided that they are not made publicly available until after the Embargo Period.

An acknowledgement in the following form should be included, together with a link to the published version on the publisher's website: "This is a post-peer-review, pre-copyedit version of an article published in [insert journal title]. The final authenticated version is available online at: http://dx.doi.org/[insert DOI]".

4. Authors' Retained Rights

Author(s) retain the following non-exclusive rights for the published version provided that, when reproducing the Article or extracts from it, the Author(s) acknowledge and reference first publication in the Journal:

a. to reuse graphic elements created by the Author(s) and contained in the Article, in presentations and other works created by them;
b. they and any academic institution where they work at the time may reproduce the Article for the purpose of course teaching (but not for inclusion in course pack material for onward sale by libraries and institutions); and
c. to reproduce, or to allow a third party Assignee to reproduce the Article in whole or in part in any printed volume (book or thesis) written by the Author(s).

5. Warranties

The Author(s) warrant and represent that:

a. (i) the Author(s) are the sole copyright owners or have been authorised by any additional copyright owner(s) to assign the rights defined in clause 2, (ii) the Article does not infringe any intellectual property rights (including without limitation copyright, database rights or trade mark rights) or other third party rights and no licence from or payments to a third party are required to publish the Article, (iii) the Article has not been previously published or licensed, (iv) if the Article contains material from other sources (e.g. illustrations, tables, text quotations), Author(s) have obtained written permissions to the extent necessary from the copyright holder(s), to license to the Assignee the same rights as set out in Clause 2 but on a non-exclusive basis and without the right to use any graphic elements on a stand-alone basis and have cited any such material correctly;
b. all of the facts contained in the Article are according to the current body of science true and accurate;
c. nothing in the Article is obscene, defamatory, violates any right of privacy or publicity, infringes any other human, personal or other rights of any person or entity or is otherwise unlawful and that informed consent to publish has been obtained for all research participants;
d. nothing in the Article infringes any duty of confidentiality which any of the Author(s) might owe to anyone else or violates any contract, express or implied, of any of the Author(s). All of the institutions in which work recorded in the Article was created or carried out have authorised and approved such research and publication; and

e. the signatory (the Author or the employer) who has signed this agreement has full right, power and authority to enter into this agreement on behalf of all of the Author(s).

6. Cooperation
The Author(s) shall cooperate fully with the Assignee in relation to any legal action that might arise from the publication of the Article, and the Author(s) shall give the Assignee access at reasonable times to any relevant accounts, documents and records within the power or control of the Author(s). The Author(s) agree that the distributing entity is intended to have the benefit of and shall have the right to enforce the terms of this agreement.

7. Author List
After signing, changes of authorship or the order of the authors listed will not be accepted unless formally approved in writing by the Assignee.

8. Edits & Corrections
The Author(s) agree(s) that the Assignee may retract the Article or publish a correction or other notice in relation to the Article if the Assignee considers in its reasonable opinion that such actions are appropriate from a legal, editorial or research integrity perspective.

This is an automated e-mail; please do not reply to this account. If you have any questions, please go to our help pages.

Thank you very much.

Kind regards,

Springer Author Services

Article Details

Journal title	Article title
Neurocritical Care	Genetic variability and trajectories of DNA methylation may support a role for HAMP in patient outcomes after aneurysmal subarachnoid hemorrhage
DOI	**Corresponding Author**
10.1007/s12028-019-00787-4	Lacey Heinsberg
Copyright transferred to	**Transferred on**
Springer Science+Business Media, LLC, part of Springer Nature and Neurocritical Care Society	Wed Jul 10 15:13:58 CEST 2019

Springer Nature Customer Service Center

Tiergartenstr. 15-17
69121 Heidelberg
Germany
phone: +49 6221 345 0
fax: +49 6221 345 4229
customerservice@springernature.com

Springer New York, LCC

233 Spring Street
New York, NY 10013
USA
phone: +1 212 460 1500 or 800-SPRINGER
(Weekdays 8:30am - 5:30pm ET)
fax: +1 212-460-1700
customerservice@springernature.com

Appendix D Manuscript 2: Genetic Variability of the Iron Homeostasis Pathway in Patient Outcomes After Aneurysmal Subarachnoid Hemorrhage

Neurocrit Care
https://doi.org/10.1007/s12028-020-00961-z

NEUROCRITICAL CARE SOCIETY

ORIGINAL WORK

Genetic Variability in the Iron Homeostasis Pathway and Patient Outcomes After Aneurysmal Subarachnoid Hemorrhage

Lacey W. Heinsberg[1*], Sheila A. Alexander[2], Elizabeth A. Crago[2], Ryan L. Minster[3], Samuel M. Poloyac[4], Daniel E. Weeks[3,5] and Yvette P. Conley[1,3]

Abstract

Background/Objective: Iron can be detrimental to most tissues both in excess and in deficiency. The brain in particular is highly susceptible to the consequences of excessive iron, especially during blood brain barrier disruption after injury. Preliminary evidence suggests that iron homeostasis is important during recovery after neurologic injury; therefore, the exploration of genetic variability in genes involved in iron homeostasis is an important area of patient outcomes research. The purpose of this study was to examine the relationship between tagging single nucleotide polymorphisms (SNPs) in candidate genes related to iron homeostasis and acute and long-term patient outcomes after aneurysmal subarachnoid hemorrhage (aSAH).

Methods: This study was a longitudinal, observational, candidate gene association study of participants with aSAH that used a two-tier design including tier 1 (discovery, $n = 197$) and tier 2 (replication, $n = 277$). Participants were followed during the acute outcome phase for development of cerebral vasospasm and delayed cerebral ischemia (DCI) and during the long-term outcome phase for death and gross functional outcome using the Glasgow Outcome Scale (GOS; poor = 1–3). Genetic association analyses were performed using a logistic regression model adjusted for age, sex, and Fisher grade. Approximate Bayes factors (ABF) and Bayesian false discovery probabilities (BFDP) were used to prioritize and interpret results.

Results: In tier 1, 235 tagging SNPs in 28 candidate genes were available for analysis and 26 associations (20 unique SNPs in 12 genes) were nominated for replication in tier 2. In tier 2, we observed an increase in evidence of association for three associations in the ceruloplasmin (*CP*) and cubilin (*CUBN*) genes. We observed an association of rs17838831 (*CP*) with GOS at 3 months (tier 2 results, odds ratio [OR] = 2.10, 95% confidence interval [CI] = 1.14–3.86, $p = 0.018$, ABF = 0.52, and BFDP = 70.8%) and GOS at 12 months (tier 2 results, OR = 1.86, 95% CI 0.98–3.52, $p = 0.058$, ABF = 0.72, and BFDP = 77.3%) as well as rs10904850 (*CUBN*) with DCI (tier 2 results, OR = 0.70, 95% CI 0.48–1.02, $p = 0.064$, ABF = 0.59, and BFDP = 71.8%).

Conclusions: Among the genes examined, our findings support a role for *CP* and *CUBN* in patient outcomes after aSAH. In an effort to translate these findings into clinical utility and improve outcomes after aSAH, additional research is needed to examine the functional roles of these genes after aSAH.

Keywords: Polymorphism, Iron, Subarachnoid hemorrhage, Patient outcome assessment, Bayes theorem

*Correspondence: law145@pitt.edu
[1] Department of Health Promotion and Development, School of Nursing, University of Pittsburgh, 440 Victoria Building, 3500 Victoria Street, Pittsburgh, PA 15261, USA
Full list of author information is available at the end of the article

Aneurysmal subarachnoid hemorrhage (aSAH) is a type of hemorrhagic stroke most commonly resulting from a ruptured aneurysm [1]. Although it does not account for a large percentage of strokes, it does account for a

substantial percentage of stroke-related death and disability. Specifically, aSAH is responsible for between 5 and 10% of strokes, as many as 25% of all stroke-related deaths, and has a 30-day mortality of at least 30% [1–3]. Moreover, greater than 50% of aSAH survivors have long-term functional deficits [1–3]. Important gaps in our knowledge as healthcare providers includes a lack of information surrounding the biology of poor outcomes and the inability to comprehensively predict which patients will do poorly after aSAH. This gap in knowledge limits the potential for targeted, early interventions from the physicians, nurses, physical, occupational, and speech therapists, and other members of the stroke recovery and rehabilitation team [4].

In an effort to identify potential biomarkers of poor outcomes after aSAH, we have focused on the iron homeostasis pathway which has been shown to be important in recovery following SAH in both a preclinical study and a small pilot study in humans [5, 6]. The scientific premise of this study was described in detail as part of our pilot work [7]. In brief, the brain is highly susceptible to damage from excessive iron, especially after neurologic injury and disruption of the blood brain barrier. In healthy humans, iron is typically bound to carrier proteins as ferric iron. However, after aSAH, catabolized blood from the subarachnoid space breaks down into several products including non-protein-bound ferrous iron (i.e., free iron) that can be toxic to nearby tissues [8–10]. Because of the toxicity of free iron, we posited that genetic variability within the iron homeostasis pathway may impact iron management after a large influx of free iron into the subarachnoid space following aSAH, subsequently accounting for a potentially important portion of variability in patient outcomes in this population. Exploration of the potential relationship between genetic variability in the iron homeostasis pathway and acute and long-term patient outcomes after aSAH may support interventions to improve outcomes and reduce the burden of this substantial public health problem. The purpose of this study was to examine genetic variability of a large number of candidate genes within the iron homeostasis pathway and patient outcomes after aSAH in humans.

Methods

Study Design

This study was a longitudinal, observational, candidate gene association study that assessed the relationship between genetic variability and patient outcomes acutely (between 0 and 14 days post-aSAH) and in the long term (at 3 and 12 months following aSAH) using a two-tier design (discovery and replication). In tier 1 (discovery), we used existing genome-wide genotype and patient outcome data collected from a larger study of aSAH participants [11] and analyzed the relationship between tagging single nucleotide polymorphisms (SNPs) in candidate genes involved in iron homeostasis and acute and long-term patient outcomes. In tier 2 (replication), we used existing patient outcome data and stored biosamples from an independent (i.e., non-overlapping) test sample of participants from the same cohort to replicate findings for the top hits identified in tier 1. A depiction of the study workflow is presented in Fig. 1.

Setting and Sample

This study was approved by the Institutional Review Board of the University of Pittsburgh, and informed consent was obtained from all research participants. Participants included in this study were prospectively recruited from University of Pittsburg Medical Center Presbyterian neurovascular intensive care unit in Pittsburgh, Pennsylvania between 2000 and 2013. Our inclusion and exclusion criteria, as well as standard treatment for patients, were previously described [7]. In brief, participants were eligible for this study if they were over the age of 18 years and diagnosed with aSAH from aneurysm rupture using cerebral angiogram. Participants were ineligible for this study if their SAH was caused by a source other than an aneurysm or if they had a history of a significant neurological disorder. Because minor allele frequencies can differ based on ancestry, and this population substructure can confound statistical analyses, the current analyses were limited to participants who self-reported their race as White.

Genotype Data Collection

Tier 1 (Discovery)

For tier 1, this study used existing genome-wide genotype data collected previously on a subset of available participants ($n = 244$) using the Affymetrix Gene Chip Assay SNP 6.0 (Affymetrix, Santa Clara, CA, USA) as described elsewhere [11]. Our initial tier 1 analyses made the assumption that the genome-wide data had undergone thorough quality control (QC). We then realized that the QC pipeline was out of date. To ensure our findings were rigorous, we completed the tier 1 genome-wide data QC and then repeated the genetic association analyses. Genome-wide data QC was completed using PLINK Version 1.9 [12, 13] R statistical software [14], the *plinkQC* package [15], and established QC thresholds [16]. The Affymetrix array yielded data for 904,087 variants for 244 participants. As part of our QC pipeline, 47 participants and 303,429 SNPs not meeting QC standards were removed from the genome-wide data. Details of our data QC pipeline are presented (Supplemental Material, Section I). The final genome-wide dataset consisted of 197

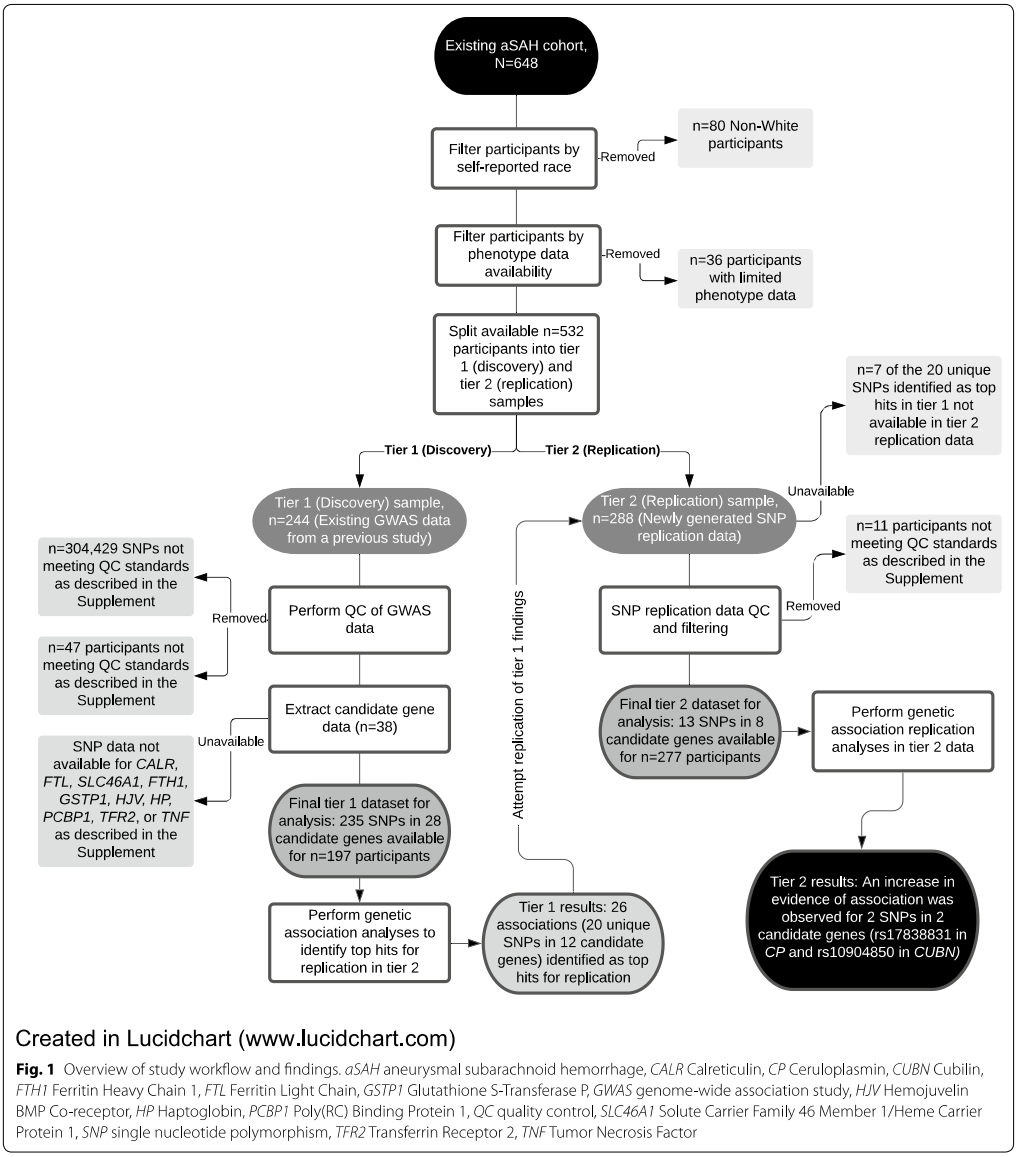


Fig. 1 Overview of study workflow and findings. *aSAH* aneurysmal subarachnoid hemorrhage, *CALR* Calreticulin, *CP* Ceruloplasmin, *CUBN* Cubilin, *FTH1* Ferritin Heavy Chain 1, *FTL* Ferritin Light Chain, *GSTP1* Glutathione S-Transferase P, *GWAS* genome-wide association study, *HJV* Hemojuvelin BMP Co-receptor, *HP* Haptoglobin, *PCBP1* Poly(RC) Binding Protein 1, *QC* quality control, *SLC46A1* Solute Carrier Family 46 Member 1/Heme Carrier Protein 1, *SNP* single nucleotide polymorphism, *TFR2* Transferrin Receptor 2, *TNF* Tumor Necrosis Factor

participants and 600,658 SNPs for extraction of tier 1 data.

Building on the pilot work for this project [7], candidate genes ($n=38$) for this study were selected based on their known biological roles in iron homeostasis (Supplemental Material, Section II). For each candidate gene, tagging SNPs within the gene region ± 1 kb were identified using the UCSC Genome Browser [17] and the National Cancer Institute Division of Cancer Genetics and Epidemiology SNPchip and SNPclip tools [18] using a linkage disequilibrium threshold of $r^2=0.8$. Following genome-wide data QC, tagging SNP data were extracted

from the genome-wide data for tier 1 analysis. None of the selected tagging SNPs located in the calreticulin, ferritin light, or heme carrying protein 1 transcript regions were available in the raw genome-wide data. Following data QC, none of the selected tagging SNPs located in the ferritin heavy, glutathione S transferase, hemojuvelin BMP co-receptor, haptoglobin, poly(RC) binding protein 1, transferrin receptor 2, or tumor necrosis factor transcript regions were available (Supplemental Material, Section II). Final tier 1 data for analysis were available for 235 tagging SNPs in 28 candidate genes (Table 1).

Tier 2 (Replication)

For tier 2, DNA was extracted for an independent test sample of aSAH participants ($n=288$) from buffy coat using a simple salting out procedure [19]. Participants were genotyped for top hits identified from the earlier pre-QC tier 1 analysis using iPLEX on the MassARRAY Typer 4.0 (Agena Bioscience) platform and software [20] according to the manufacturer's standard protocols at the University of Pittsburgh Genomics Research Core. Duplicates were included on the plate with no inconsistencies identified, and all genotypes were called by two blinded individuals. The assay design with primer and probe sequences are presented (Supplemental Material, Section III). Our raw genotype data included 13 SNPs for 288 participants. As part of our QC pipeline, 11 participants and 0 SNPs not meeting QC standards were removed. Details of our genotype data collection methods and QC pipeline are presented (Supplemental Material, Section III). The final tier 2 dataset consisted of 277 participants and 13 SNPs. Note that tier 2 replication SNPs were chosen, and the genotype data generated, before implementation of the updated QC pipeline in tier 1. Therefore, in the analyses presented here (i.e., post-QC), some of our top associations in the post-QC tier 1 are missing tier 2 replication data because they were not present in the pre-QC tier 1 results used to select SNPs for tier 2. This limitation is discussed in more detail in the discussion section. A depiction of the study workflow is presented in Fig. 1.

Patient Outcomes

This study used demographic (e.g., age, sex, race), social (e.g., marital status), clinical (e.g., severity of injury measured using the clinical grading scales Fisher grade and World Federation of Neurosurgical Societies [WFNS] grade), and treatment (e.g., intervention [clip versus coil], medications administered) data extracted from the medical record as well as patient outcome data collected during the acute and long-term phases.

Acute outcome measures used in this study were cerebral vasospasm (CV) and delayed cerebral ischemia (DCI) within 14 days of aSAH. Both outcomes are important acute complications that can occur during the recovery phase following aSAH and are indicators of potentially poor recovery in the long term [4]. CV was defined as cerebral vessel narrowing of $\geq 25\%$ during cerebral angiogram performed and measured by a neurosurgeon [7]. DCI was defined as the co-occurrence of (1) non-ischemic neurological deterioration such as an increase of ≥ 2 points on the National Institutes of Health Stroke Scale or a new and persistent (present for >1 h) neurological deficit and (2) abnormal cerebral blood flow measured using cerebral angiogram or transcranial Doppler [7].

Table 1 Iron homeostasis candidate genes examined

Gene	Name	Gene	Name
ACO1	Aconitase 1	*HMOX1*	Heme-oxygenase 1
ACO2	Aconitase 2	*HMOX2*	Heme-oxygenase 2
APP	Amyloid precursor protein	*HPX*	Hemopexin
CD163	Hemoglobin scavenger receptor	*IREB2*	Iron responsive element binding protein 2
CP	Ceruloplasmin	*LRP1*	LDL receptor related protein
CUBN	Cubilin	*PGRMC1*	Progesterone receptor membrane
CYBRD1	Duodenal cytochrome b	*SLC11A1*	Solute carrier family 11 member 1
FECH	Ferrochelatase	*SLC11A2*	Divalent metal transporter 1
FLVCR1	Feline leukemia virus subgroup C receptor	*SLC25A37*	Solute carrier family 25 member 37 (Mitoferrin 1)
FTMT	Mitochondrial ferritin	*SLC40A1*	Solute carrier family 40 member 1 (Ferroportin)
FXN	Frataxin	*SLC48A1*	Solute carrier family 48 member 1
GLRX5	Glutaredoxin 5	*STEAP3*	STEAP3 metalloreductase
HEPH	Hephaestin	*TF*	Transferrin
HFE	Human hemochromatosis protein	*TFRC*	Transferrin receptor 1

In the long-term outcome phase, trained study staff performed patient interviews in person or via telephone at 3 and 12 months following aSAH. To measure global functional status, the Glasgow Outcomes Scale (GOS), which has established validity in people with neurological injury, was used as a quantitative measure of participants' ability to function on a scale of 1 (death) to 5 (good recovery) [21]. If participants were unable to participate in the interview, their caregiver or proxy was interviewed. Death data were obtained from the medical record, caregiver/family report, or the Social Security Death Index. All study staff involved in patient recruitment and patient outcome data collection were blinded to participant genotype.

Statistical Analysis

Descriptive and Preliminary Analyses

Statistical analyses were conducted using R statistical software [14] and PLINK [12, 13]. The patient outcomes of CV, DCI, and death were treated as binary (occurrence versus no occurrence), and GOS scores were dichotomized as good (4–5) or poor (1–3). Standard descriptive statistics were computed in R for all variables, and data were screened for assumptions of logistic regression and examined for outliers. Preliminary analyses were conducted to identify potential confounders/covariates. Hardy-Weinberg Equilibrium was evaluated for all SNPs as part of our QC pipelines.

Genetic Association Analyses

Genetic association analyses were performed in PLINK [12, 13] using a logistic regression model adjusted for age, sex, and Fisher grade. Only additive models (treating SNPs as ordinal based on variant allele dosage) were considered. The approximate Bayes factor (ABF) was used to compute Bayesian false discovery probabilities (BFDP) for each SNP-phenotype association which was subsequently used to prioritize and 'flag' SNPs for replication in tier 2 [22, 23]. The ABF is an approximation to the Bayes factor where an ABF < 1 indicates an increase in evidence for association and an ABF > 1 indicates a decrease in evidence for association [22, 23]. The ABF is used to compute the BFDP which can be interpreted as a probability of false discovery regardless of power, sample size, or how many other SNPs were tested ultimately preventing the need for correction for multiple testing [22, 23].

For the purpose of transparency, replication, and future application of the study methodology in other populations, we present an expanded and detailed explanation of the genetic association analyses performed here including calculation of the ABF and BFDP as well as the follow-up flagging approach, including formulae (Supplemental Material, Section IV). Because the direction of effect of associations observed in tier 1 is not integrated in tier 2 replication calculations, we also performed a mega-analysis combining tier 1 and tier 2 data to aid in interpretation. Lastly, in genetic association studies, ancestry can be an important confounder of results given that minor allele frequencies often differ by ancestry and race. In order to explore the influence of ancestry (versus self-reported race) and aid in interpretation of results, we performed an ancestry sensitivity analysis as described in detail (Supplemental Material, Section IV).

Results

Sample Characteristics

Sample sizes for each association test varied between tier 1 and tier 2 as well as between associations depending on genotyping success rate and outcome data availability; results detailing specific samples sizes for all associations are presented (Supplemental Material, Section V). An overview of the available sample sizes and associated demographic and clinical characteristics is presented (Table 2). For tier 1, we had an overall sample size of 197 participants. Our tier 1 sample had a mean age of 54.4 (±11.3) years, was 69% female, and Fisher grades of 2, 3, or 4 accounted for 29.4%, 53.3%, and 17.3%, respectively. For tier 2, we had an overall samples size of 277 participants. Our tier 2 sample had a mean age of 54.1 (±11.1) years, was 74.7% female, and Fisher grades of 2, 3, or 4 accounted for 52%, 33.9%, and 14.1%, respectively. Age,

Table 2 Demographic and clinical characteristics for tier 1 and tier 2

Variable	Tier 1 (discovery), *n* = 197	Tier 2 (replication), *n* = 277
Age, mean years (SD)	54.4 (11.3)	54.1 (11.1)
Sex, female (*n*, %)	136 (69.0)	207 (74.7)
Treatment, clip (*n*, %)	79 (40.1)	84 (30.3)
Fisher grade (*n*, %)		
2	58 (29.4)	144 (52.0)
3	105 (53.3)	94 (33.9)
4	34 (17.3)	39 (14.1)
WFNS grade (*n*, %)		
1	104 (52.8)	143 (51.6)
2	37 (18.8)	44 (15.9)
3	10 (5.1)	25 (9.0)
4	26 (13.2)	37 (13.4)
5	20 (10.2)	28 (10.1)
Married, yes (*n*, %)	131 (66.8)[a]	174 (62.8)

SD standard deviation, *WFNS* World Federation of Neurological Societies

[a] Marital status of one participant was unknown in the tier 1 sample; percentage calculated from *n* = 196 of known marital status

WFNS scores, and marital status were similar between the groups as shown but we observed differences in treatment between the groups with 40.1% in tier 1 receiving treatment via clipping compared with only 30.3% in tier 2. In a pooled analysis of tier 1 and tier 2 participants, older age and higher Fisher grade were associated with poor outcomes at all time points. Sex was not associated with outcomes in our sample but was included as a covariate because of the importance of estrogen response elements in iron homeostasis [24].

Genetic Association Analyses

In our tier 1 data, 235 tagging SNPs passing QC procedures were available from 28 candidate genes listed in Table 1. Sample sizes ranged from 106 to 189 and 239 to 273 in tiers 1 and 2, respectively, depending on the SNP and patient outcome of interest. Expanded details for all SNPs examined are presented (Supplemental Material, Section V).

Based on our calibration approach described above, 26 associations from tier 1 were flagged as noteworthy of investigation and replication in tier 2 (Table 3). Models presented in Table 3 include age, sex, and Fisher grade as covariates. The 26 flagged associations were from 20 unique SNPs positioned in 12 of our candidate genes including aconitase 1 (*ACO1*), amyloid precursor protein (*APP*), ceruloplasmin (*CP*), cubilin (*CUBN*), cytochrome B reductase 1 (*CYBRD1*), ferrochelatase (*FECH*), ferritin mitochondrial (*FTMT*), frataxin (*FXN*), glutaredoxin 5 (*GLRX5*), hemopexin (*HPX*), LDL receptor-related protein 1 (*LRP1*), and transferrin (*TF*). An expanded results section has been included detailing all tier 1 results (Supplemental Material, Section V).

The tier 2 replication results are presented and are ranked by the tier 2 BFDP (Table 3). Six associations were flagged as increasing evidence of association from tier 1 to tier 2 based on an ABF < 1 and a decrease in the BFDP. However, it should be noted that three of our top six associations had odds ratios (OR) in the opposite direction in tier 2 compared with tier 1. Specifically, rs11087985 (*APP*) with death at 3 and 12 months and rs13302577 (*ACO1*) with death at 12 months had OR < 1 in tier 2, suggesting the minor allele confers protection compared with an OR > 1 in tier 1, suggesting the minor allele confers risk. As a result, the signals for these associations were canceled out in the mega-analysis (Supplemental Material, Section V). The opposite effect directions observed in these associations reduces our confidence that these are actually notable signals of interest. Three of the six associations, however, were important in tier 1, tier 2, and a mega-analysis combining tier 1 and tier 2 data. Specifically, rs17838831 (*CP*) was associated with GOS at 3 months in tier 1 (OR = 2.83, 95% confidence interval [CI] = 1.33–5.99, $p = 0.007$, ABF = 0.52, BFDP = 82.2%), tier 2 (OR = 2.10, 95% CI 1.14–3.86, $p = 0.018$, ABF = 0.52, BFDP = 70.8%), and a mega-analysis (OR = 2.32, 95% CI 1.45–3.70, $p = 0.0004$, ABF = 0.10, BFDP = 47.8%). Similarly, we observed an association between the same SNP and poor GOS at 12 months in tier 1 (OR = 3.09, 95% CI 1.39–6.87, $p = 0.006$, ABF = 0.53, BFDP = 82.7%), tier 2 (OR = 1.86, 95% CI 0.98–3.52, $p = 0.058$, ABF = 0.72, BFDP = 77.3%), and a mega-analysis (OR = 2.22, 95% CI 1.36–3.63, $p = 0.001$, ABF = 0.18, BFDP = 61.6%). Finally, we also observed an association between rs10904850 in *CUBN* and DCI in tier 1 (OR = 0.57, 95% CI 0.35–0.93, $p = 0.024$, ABF = 0.48, BFDP = 81.2%), tier 2 (OR = 0.70, 95% CI 0.48–1.02, $p = 0.064$, ABF = 0.59, BFDP = 71.8%), and a mega-analysis (OR = 0.65, 95% CI 0.48–0.87, $p = 0.004$, ABF = 0.11, BFDP = 50.1%). For the remaining 12 associations, we observed a tier 2 ABF > 1 and an increase in the BFDP compared with tier 1 which suggests that the tier 2 data decreased the evidence for association. As with tier 1, details of tier 2 and the mega-analysis results are presented (Supplemental Material, Section V). A depiction of the study workflow, including an overview of the main results, is presented in Fig. 1.

In an ancestry sensitivity analysis of tier 1 data, our top hits (rs17838831 [*CP*] with GOS at 3 and 12 months; and rs10904850 [*CUBN*] with DCI) remained top hits in the sensitivity analysis. Overall, we observed strong correlation between the results with an 82.5% concordance between the top 40 hits from the original analysis compared with the sensitivity analysis (i.e., 33 of the top 40 associations were common between the two analyses). The details of this sensitivity analysis and results are presented (Supplemental Material, Section VI).

Discussion

We selected the candidate genes for this study based on attributes that are relevant to our phenotypes of interest during recovery from aSAH. Of our candidate genes, the associations from the *CP* and *CUBN* genes stand out in our results suggesting potential importance to aSAH recovery. Specifically, one SNP in *CP*, rs17838831 (located on chromosome 3 at position 148939861 [GRCh37/hg19]), was important in the tier 1 discovery sample, tier 2 replication sample, and mega-analysis for GOS at 3 and 12 months. Ceruloplasmin, the protein encoded for by *CP*, is a multicopper oxidase that accounts for a majority of serum copper and is heavily involved in ferroxidase activity, mitochondrial function, and antioxidant and anti-inflammatory mechanisms [25, 26]. Existing literature demonstrates the importance of *CP* in iron homeostasis. Specifically, even when total body iron stores are normal, low plasma levels of *CP*

Table 3 Results of binary logistic regression exploring associations of candidate tagging SNPs with patient outcomes while controlling for age, sex, and Fisher grade

SNP	Gene	Outcome	Tier 1 (discovery)								Tier 2 (replication)							
			n, case	*n*, control	MAF	OR	95% CI	*p*	ABF	BFDP	*n*, case	*n*, control	MAF	OR	95% CI	*p*	ABF	BFDP
rs17838831	*CP*	GOS3	42	120	0.15	2.83	1.33–5.99	0.007	0.52	82.2	69	184	0.14	2.10	1.14–3.86	0.018	0.52	70.8[a]
rs10904850	*CUBN*	DCI	71	118	0.31	0.57	0.35–0.93	0.024	0.48	81.2	108	165	0.34	0.70	0.48–1.02	0.064	0.59	71.8[a]
rs11087985	*APP*	MORT3	27	150	0.37	2.79	1.39–5.63	0.004	0.43	79.4	42	223	0.34	0.63	0.36–1.11	0.113	0.81	75.8[b]
rs13302577	*ACO1*	MORT12	28	130	0.35	2.17	1.14–4.13	0.019	0.55	83.3	45	194	0.33	0.53	0.28–0.99	0.046	0.67	77.0[b]
rs17838831	*CP*	GOS12	32	128	0.15	3.09	1.39–6.87	0.006	0.53	82.7	59	180	0.14	1.86	0.98–3.52	0.058	0.72	77.3[a]
rs1411675	*FXN*	GOS3	42	119	0.40	0.40	0.22–0.75	0.004	0.36	76.2	69	184	0.46	1.01	0.64–1.60	0.975	1.31	80.8
rs8177248	*TF*	DCI	71	118	0.36	1.56	1.04–2.41	0.031	0.47	80.8	107	165	0.37	0.81	0.56–1.17	0.261	1.04	81.5
rs11087985	*APP*	MORT12	29	131	0.37	2.21	1.15–4.26	0.018	0.56	83.3	45	194	0.34	0.69	0.40–1.19	0.182	0.91	82.0[b]
rs7870295	*FXN*	GOS3	41	119	0.38	0.41	0.22–0.76	0.005	0.39	77.7	69	184	0.42	0.98	0.62–1.55	0.915	1.31	82.0
rs3991	*APP*	GOS3	42	119	0.24	2.14	1.19–3.85	0.011	0.44	80	69	184	0.22	1.09	0.63–1.88	0.763	1.21	82.9
rs3847364	*CUBN*	MORT3	27	147	0.47	0.36	0.17–0.76	0.007	0.52	82.3	42	222	0.45	0.83	0.50–1.40	0.485	1.15	84.2
rs3847364	*CUBN*	MORT12	29	128	0.47	0.37	0.17–0.78	0.009	0.54	83	45	194	0.45	0.82	0.49–1.38	0.453	1.13	84.7
rs8177224	*TF*	GOS3	42	120	0.34	1.89	1.09–3.28	0.024	0.52	82.5	69	183	0.32	1.03	0.63–1.70	0.898	1.27	85.7
rs12476341	*CYBRD1*	DCI	70	118	0.23	1.97	1.13–3.43	0.018	0.48	81.2	108	165	0.25	1.01	0.68–1.52	0.952	1.39	85.7
rs8177224	*TF*	MORT3	27	150	0.34	2.12	1.13–3.98	0.020	0.55	83.3	42	222	0.32	0.95	0.54–1.67	0.849	1.21	85.8
rs7870295	*FXN*	GOS12	31	126	0.38	0.41	0.21–0.82	0.012	0.54	82.8	59	180	0.42	0.94	0.58–1.52	0.794	1.27	86.0
rs10435797	*ACO1*	DCI	70	118	0.34	0.62	0.39–0.98	0.039	0.53	82.8	105	164	0.30	1.06	0.72–1.55	0.783	1.340	87.0
rs1560550	*FTMT*	DCI	71	117	0.50	0.62	0.39–0.97	0.038	0.53	82.5	107	165	0.45	1.03	0.73–1.44	0.883	1.53	87.8
rs2035675	*HPX*	MORT12	28	126	0.23	2.08	1.16–3.75	0.015	0.48	81.3	NA							
rs533952	*FECH*	DCI	70	118	0.30	0.55	0.33–0.92	0.022	0.48	81.3	NA							
rs2035675	*HPX*	MORT3	26	145	0.23	2.08	1.15–3.77	0.015	0.49	81.6	NA							
rs10876966	*LRP1*	DCI	71	118	0.26	1.77	1.06–2.97	0.030	0.53	82.7	NA							
rs2306692	*LRP1*	DCI	71	115	0.17	0.44	0.23–0.85	0.014	0.53	82.7	NA							
rs1736439	*FECH*	DCI	67	118	0.39	0.60	0.38–0.97	0.038	0.54	83	NA							
rs716315	*GLRX5*	GOS12	32	128	0.45	2.27	1.17–4.40	0.016	0.55	83.1	NA							
rs1028932	*ACO1*	MORT3	27	150	0.49	2.11	1.12–3.97	0.020	0.56	83.3	NA							

ABF approximate Bayes factor, *BFDP* Bayesian false discovery probability, *CI* confidence interval, *DCI* delayed cerebral ischemia, *GOS3* Glasgow Outcome Scale at 3 months, *GOS12* Glasgow Outcome Scale at 12 months, *MORT3* death at 3 months, *MORT12* death at 12 months, *n case* count for affected individuals (cases), *n control* count for unaffected individuals (controls), *NA* not applicable as no tier 2 replication data existed, *OR* odds ratio, *p p* value based on alpha of 0.05, *SNP* single nucleotide polymorphism

[a] Increase in evidence of association in Tier 2 based on ABF < 1 and drop in BFDP as well as directionally consistent OR and significant results in mega-analysis

[b] Increase in evidence of association in Tier 2 based on ABF < 1 and drop in *BFDP* but discordant OR directions and no evidence of association in mega-analysis. Note: All tier 1 associations presented had a C value of < 5

cause hypoferremia [27]. Moreover, aceruloplasminemia, a disorder of impaired iron homeostasis and classical example of the clinical features of *CP* dysfunction, is characterized by iron accumulation in microglia and neurons and increased reactive oxygen species which are important disrupted mechanisms after aSAH [25]. Further, *CP* knockout mice show evidence of increased lipid peroxidation and iron accumulation, functions important during dysregulated iron and lipid metabolism and ferroptosis after aSAH [25, 28]. Outside of its role in iron homeostasis, *CP* is thought to help deliver copper to damaged areas of infection, inflammation, or trauma [26]. Although there has been less mechanistic work investigating the plausible role of *CP* after aSAH specifically, interestingly, lower *CP* levels in the cerebrospinal fluid of aSAH patients have been associated with development of deep cerebral infarcts in a small pilot study [6]. Moreover, rs17838831 has been shown to be strongly associated with plasma levels of *CP* though the allele and direction of association were not reported [29]. In tier 1, tier 2, and the mega-analysis of this current study, with each dose of the rs17838831 variant allele, participants had between a 2.10 and 2.83 times higher odds of poor GOS at 3 months. Similarly, we observed an association between the same SNP and poor GOS at 12 months in tier 1 with OR between 1.86 and 3.09. Given our findings, this SNP warrants further investigation in patient outcomes after aSAH.

In addition, a SNP in *CUBN*, rs10904850 (located on chromosome 10 at position 16997707 [GRCh37/hg19]), was important for DCI in tier 1, tier 2, and the mega-analysis. *CUBN*, the protein encoded for by *CUBN*, plays an important role in iron homeostasis by facilitating uptake of transferrin iron and clearance of hemoglobin at the kidneys [30, 31]. Interestingly, in a multi-ethnic study of iron disorders, rs10904850 was associated with serum iron in African Americans but not in other ethnic groups; [32] similar to above, the allele and direction of this association are not clear [32]. The influence of *CUBN* in chronic kidney disease has been shown [31] but has not been investigated after aSAH specifically. After aSAH, acute kidney injury occurs in upwards of 25% of patients and even subtle decreases in creatine clearance have been associated with poor outcomes after aSAH [33]. Based on this research, it is possible that this SNP may impact iron homeostasis by influencing efficiency of reabsorption of iron in the kidney. In tier 1, tier 2, and a mega-analysis, with each dose of the rs17838831 variant allele, participants had between 0.57 and 0.70 times lower odds of DCI. Given our findings, this SNP warrants further investigation in patient outcomes after aSAH.

Of note, there are some important differences between the tier 1 and tier 2 samples. Although we controlled for age, sex, and severity of injury as measured by Fisher grade, between-sample differences may be important to the interpretation of study findings. Specifically, our samples were 69% and 74.7% female for tier 1 and tier 2, respectively. Though research suggests that the overall differences in outcomes between men and women are null and we controlled for sex in our analyses [34], estrogen is known to play a role in iron homeostasis and may be an important consideration here [24]. Moreover, an unexpected observation between the cohorts included discordant Fisher grade distributions. Specifically, Fisher grades of 2 and 3 were 29.4% and 53.3%, respectively, for tier 1 and 52% and 33.9%, respectively, for tier 2. Severity of injury is an important predictor of patient outcomes and we had several measures available in our cohort that were not only significantly associated with patient outcomes after aSAH, but also more similar between tier 1 and tier 2 (e.g., WFNS). Ultimately, we chose to control for Fisher grade as opposed to WFNS because it is a more direct measure of the amount of blood within the subarachnoid space, more closely associated with iron homeostasis within the body, and more relevant to the scientific premise of this study. In an attempt to explore the role of our choice of measure for severity of injury in our analyses, we repeated the tier 1 genetic association analysis controlling for WFNS rather than Fisher grade. Importantly, our associations between rs17838831 (*CP*) and GOS at 3 and 12 months and rs10904850 (*CUBN*) and DCI remained in the top hits in our sensitivity analysis. Overall, we identified an 80% concordance between the top 40 associations in tier 1 (Supplemental Material, Section VII).

Although there are many strengths to this study including embedded replication and the use of Bayesian statistical methods to aid in interpretation of results, there are some important limitations that should be acknowledged. First, while our data QC pipeline resulted in rigorous analyses of more accurate data, it did reduce our tier 1 sample size and power significantly (though our overall sample size remains quite large compared with similar patient outcomes work in the aSAH population). Specifically, the small sample size in tier 1 (i.e., post-QC) may have prevented us from detecting signals of association for SNPs with small effect sizes which were therefore not carried forward for tier 2 replication. Similarly, because we were limited to examining SNPs available in the tier 1 genome-wide genotype data, we were not able to comprehensively examine all SNPs within our candidate genes. It is possible that SNPs located in our candidate genes, but not examined as part of this study, may be associated with patient outcomes after aSAH. For example, we had no data available for the haptoglobin gene which has received a great deal of recent attention

and been shown both experimentally and clinically to be important in outcomes after aSAH [35, 36]. Additionally, we lacked tier 2 replication data for some associations flagged as noteworthy in tier 1 (i.e., SNPs in *HPX*, *FECH*, *LRP1*, and *GLRX5*) so we were unable to determine whether all tier 1 SNP-phenotype associations could be replicated or not. Therefore, while *CP* and *CUBN* rose to the top in our analyses, we cannot necessarily eliminate the remaining list of candidate genes as plausible future targets of investigation in aSAH recovery research. Future areas of investigation should include attempting replication of associations observed in tier 1 that we were lacking tier 2 data for and exploring genetic variability of genes with inadequate tier 1 data for inclusion (Supplemental Material, Section II).

Next, this study was limited only to genetic variation of the iron homeostasis pathway. A future area of research should be to examine other omic mechanisms including levels of gene products in serum or cerebrospinal fluid and subsequent associations with patient outcomes. Finally, given that minor allele frequencies often differ based on race and ancestry, this study was limited to only participants who self-reported their race as White which restricts the generalizability of findings. A strength of this study, however, was the ability to perform a tier 1 sensitivity analysis controlling for ancestry using principal components computed from the genome-wide data. Identifying a concordance of 82.5% between the main analysis and ancestry sensitivity analysis offers some evidence of the utility of self-reported race for the Pittsburgh, Pennsylvania population when lacking genome-wide data. However, these results also underscore the importance that ancestry can play in genetic association studies as well as the critical need to replicate of findings.

Conclusion

Patient outcomes after aSAH vary widely and reliable and stable biomarkers to identify patients who may do poorly are needed to improve supportive care. In this study, SNPs in the *CP* and *CUBN* genes were flagged as important for future investigation. Specifically, we observed associations between rs17838831 (*CP*) and GOS at 3 months and 12 months and rs10904850 (*CUBN*) and DCI after aSAH in a discovery and replication sample, and in a mega-analysis. In order to translate this work to clinical practice in the future, functional investigation of *CP* and *CUBN* after aSAH is warranted.

Electronic supplementary material
The online version of this article (https://doi.org/10.1007/s12028-020-00961-z) contains supplementary material, which is available to authorized users.

Author details
[1] Department of Health Promotion and Development, School of Nursing, University of Pittsburgh, 440 Victoria Building, 3500 Victoria Street, Pittsburgh, PA 15261, USA. [2] Department of Acute and Tertiary Care, School of Nursing, University of Pittsburgh, Pittsburgh, PA, USA. [3] Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA. [4] Department of Pharmaceutical Sciences, School of Pharmacy, University of Pittsburgh, Pittsburgh, PA, USA. [5] Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA.

Acknowledgements
We would like to acknowledge Sandra Deslouches for her expertise and work in the laboratory, Tiffany Wang for her help formatting the Supplemental Tables associated with this publication, and the anonymous reviewers who took the time to critically evaluate this paper as their feedback improved the clarity and quality of this work.

Author Contributions
All authors meet authorship criteria, have read and approved the submitted manuscript, and certify that they have participated sufficiently in the work to take responsibility for the content including the concept, design, analysis, writing, or revision. Lacey W. Heinsberg contributed to the study conception and design, acquisition, analysis, and interpretation of data, and drafted, critically revised, and gave final approval for the manuscript. Sheila A. Alexander contributed to the study design, interpretation of data, and critically revised and gave final approval for the manuscript. Elizabeth A. Crago contributed to the acquisition and interpretation of data and critically revised and gave final approval for the manuscript. Ryan L. Minster contributed to the analysis and interpretation of data and critically revised and gave final approval for the manuscript. Samuel M. Poloyac contributed to acquisition and interpretation of data and critically revised and gave final approval for the manuscript. Daniel E. Weeks contributed to the study design, analysis and interpretation of data, and critically revised and gave final approval for the manuscript. Yvette P. Conley contributed to the study conception and design, acquisition and interpretation of data, and critically revised and gave final approval for the manuscript. All authors agree to be accountable for all aspects of the work in ensuring that questions relating to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Source of Support
Research reported in this publication was supported by the National Institute of Nursing Research of the National Institutes of Health under Award Nos. F31NR017311, R01NR004339, R01NR013610, and T32NR009759 with additional support from the Nightingale Awards of Pennsylvania, Center for Jonas Nursing and Veterans Healthcare, Jayne F. Wiggins Memorial Award, and Sigma Theta Tau—Eta Chapter. The content is solely the responsibility of the authors and does not represent the official views of the National Institutes of Health or supporting foundations.

Conflict of interest
LW Heinsberg reports Grants from the National Institutes of Health, University of Pittsburgh Jayne F. Wiggins Memorial Scholarship, Eta Chapter, Sigma Theta Tau, Inc., Jonas Foundation, and the Nightingale Awards of Pennsylvania during the conduct of this study. YP Conley, DE Weeks, and EA Crago reports Grants from the National Institutes of Health. SA Alexander, RL Minster, and SM Poloyac report nothing to disclose.

Ethical Conduct of Research
Informed consent was obtained from all study participants. Institutional Review Board approval at the University of Pittsburgh is in place (IRB approval number STUDY19100368) and we have adhered to ethical considerations in the protection of all human subjects involved.

Publisher's Note

Published online: 03 April 2020

References

1. Boling B, Groves TR. Management of subarachnoid hemorrhage. Crit Care Nurse. 2019;39(5):58–67.
2. Zacharia BE, Hickman ZL, Grobelny BT, et al. Epidemiology of aneurysmal subarachnoid hemorrhage. Neurosurg Clin N Am. 2010;21(2):221–33.
3. van Gijn J, Kerr RS, Rinkel GJE. Subarachnoid haemorrhage. Lancet. 2007;369(9558):306–18.
4. Teo M, Teo M, Turner C, et al. What factors determine treatment outcome in aneurysmal subarachnoid hemorrhage in the modern era? A Post Hoc STASH analysis. World Neurosurg. 2017;105:270–81.
5. Tan G, Liu L, He Z, Sun J, Xing W, Sun X. Role of hepcidin and its downstream proteins in early brain injury after experimental subarachnoid hemorrhage in rats. Mol Cell Biochem. 2016;418(1):31–8.
6. Gomes JA, Selim M, Cotleur A, et al. Brain iron metabolism and brain injury following subarachnoid hemorrhage: iCeFISH-pilot (CSF iron in SAH). Neurocrit Care. 2014;21(2):285–93.
7. Heinsberg LW, Arockiaraj AI, Crago EA, et al. Genetic variability and trajectories of DNA methylation may support a role for HAMP in patient outcomes after aneurysmal subarachnoid hemorrhage. Neurocrit Care 2019;32(2):550–63.
8. Wagner KR, Sharp FR, Ardizzone TD, Lu A, Clark JF. Heme and iron metabolism: role in cerebral hemorrhage. J Cereb Blood Flow Metab. 2003;23(6):629–52.
9. Bishop GM, Robinson SR. Quantitative analysis of cell death and ferritin expression in response to cortical iron: implications for hypoxia-ischemia and stroke. Brain Res. 2001;907(1–2):175–87.
10. Garton T, Keep RF, Hua Y, Xi G. Brain iron overload following intracranial haemorrhage. Stroke Vasc Neurol. 2016;1(4):172–84.
11. Kim H, Crago E, Kim M, et al. Cerebral vasospasm after sub-arachnoid hemorrhage as a clinical predictor and phenotype for genetic association study. Int J Stroke. 2013;8(8):620–5.
12. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81(3):559–75.
13. Purcell S. PLINK. Available from: https://www.cog-genomics.org/plink2. Accessed 20 Mar 2020.
14. Team RC. R: a language and environment for statistical computing. 2018; Available from: https://www.r-project.org/. Accessed 22 Jan 2018.
15. Meyer H. plinkQC: genotype quality control with "PLINK." 2019.
16. Marees AT, de Kluiver H, Stringer S, et al. A tutorial on conducting genome-wide association studies: quality control and statistical analysis. Int J Methods Psychiatr Res. 2018;27(2):e1608.
17. Karolchik D, Hinrichs A, Furey T, et al. The UCSC table browser data retrieval tool. Nucleic Acids Res. 2004;32:493–6.
18. Machiela M, Chanock S. LDlink a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinformatics. 2015;31:3555–7.
19. Miller SA, Dykes DD, Polesky HF. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 1988;16(3):1215.
20. Gabriel S, Ziaugra L, Tabbaa D. SNP genotyping using the sequenom massARRAY iPLEX platform. Curr Protoc Hum Genet. 2009;60(suppl 60):2121–21218. https://doi.org/10.1002/0471142905.hg0212s60.
21. Jennett B, Bond M. Assessment of outcome after severe brain damage. Lancet (London, England). 1975;1(7905):480–4.
22. Wakefield J. A Bayesian measure of the probability of false discovery in molecular genetic epidemiology studies. Am J Hum Genet. 2007;81:208–27.
23. Wakefield J. Bayes factors for genome-wide association studies: comparison with *P* values. Genet Epidemiol. 2008;33(1):79–86.
24. Qian Y, Yin C, Chen Y, et al. Estrogen contributes to regulating iron metabolism through governing ferroportin signaling via an estrogen response element. Cell Signal. 2015;27(5):934–42.
25. Patel BN, Dunn RJ, Jeong SY, Zhu Q, Julien JP, David S. Ceruloplasmin regulates iron levels in the CNS and prevents free radical injury. J Neurosci. 2002;22(15):6578–86.
26. Besold AN, Culbertson EM, Culotta VC. The Yin and Yang of copper during infection. J Biol Inorg Chem. 2016;21(2):137–44.
27. Cherukuri S, Potla R, Sarkar J, Nurko S, Harris ZL, Fox PL. Unexpected role of ceruloplasmin in intestinal iron absorption. Cell Metab. 2005;2(5):309–19.
28. Bickford JS, Ali NF, Nick JA, et al. Endothelin-1-mediated vasoconstriction alters cerebral gene expression in iron homeostasis and eicosanoid metabolism. Brain Res. 2014;1588:25–36.
29. Adamsson Eryd S, Sjögren M, Smith JG, et al. Ceruloplasmin and atrial fibrillation: evidence of causality from a population-based Mendelian randomization study. J Intern Med. 2014;275(2):164–71. https://doi.org/10.1111/joim.12144.
30. Verroust PJ, Christensen EI. Megalin and cubilin–the story of two multipurpose receptors unfolds. Nephrol Dial Transpl. 2002;17(11):1867–71.
31. Böger CA, Chen MH, Tin A, et al. CUBN is a gene locus for albuminuria. J Am Soc Nephrol. 2011;22(3):555–70.
32. McLaren CE, McLachlan S, Garner CP, et al. Associations between single nucleotide polymorphisms in iron-related genes and iron status in multiethnic populations. PLoS ONE. 2012;7(6):e338339.
33. Zacharia BE, Ducruet AF, Hickman ZL, et al. Renal dysfunction as an independent predictor of outcome after aneurysmal subarachnoid hemorrhage: a single-center cohort study. Stroke. 2009;40(7):2375–81.
34. Hamdan A, Barnes J, Mitchell P. Subarachnoid hemorrhage and the female sex: Analysis of risk factors, aneurysm characteristics, and outcomes. J Neurosurg. 2014;121(6):1367–73.
35. Morton MJ, Hostettler IC, Kazmi N, et al. Haptoglobin genotype and outcome after aneurysmal subarachnoid haemorrhage. J Neurol Neurosurg Psychiatry. 2019;91(3):305–13.
36. Hugelshofer M, Buzzi RM, Schaer CA, et al. Haptoglobin administration into the subarachnoid space prevents hemoglobin-induced cerebral vasospasm. J Clin Invest. 2019;129(12):5219–35.

Appendix D.1 License Agreement for Manuscript 2

Copyright Transfer Statement (CTS) for your article in Neurocritical Care (961)

Springer <springerauthorquery@springeronline.com>
Tue 3/17/2020 2:53 PM
To: Heinsberg, Lacey Wright <LAW145@pitt.edu>

Copyright Transfer

17.03.2020

visit us at springer.com

Confirmation of your Copyright Transfer

Dear Author,

Please note: This e-mail is a confirmation of your copyright transfer and was sent to you only for your own records.

1. Publication
The copyright to this article, (including any supplementary information and graphic elements therein (e.g. illustrations, charts, moving images) (the 'Article'), is hereby assigned for good and valuable consideration to Springer Science+Business Media, LLC, part of Springer Nature and Neurocritical Care Society (the 'Assignee'). Headings are for convenience only.

2. Grant of Rights
In consideration of the Assignee evaluating the Article for publication, the Author(s) grant the Assignee without limitation the exclusive (except as set out in clauses 3, 4 and 5 a) iv), assignable and sub-licensable right, unlimited in time and territory, to copy-edit, reproduce, publish, distribute, transmit, make available and store the Article, including abstracts thereof, in all forms of media of expression now known or developed in the future, including pre- and reprints, translations, photographic reproductions and extensions. Furthermore, to enable additional publishing services, such as promotion of the Article, the Author(s) grant the Assignee the right to use the Article (including the use of any graphic elements on a stand-alone basis) in whole or in part in electronic form, such as for display in databases or data networks (e.g. the Internet), or for print or download to stationary or portable devices. This includes interactive and multimedia use as well as posting the Article in full or in part or its abstract on social media, and the right to alter the Article to the extent necessary for such use. The Assignee may also let third parties share the Article in full or in part or its abstract on social media and may in this context sub-license the Article and its abstract to social media users. Author(s) grant to Assignee the right to re-license Article metadata without restriction (including but not limited to author name, title, abstract, citation, references, keywords and any additional information as determined by Assignee).

3. Self-Archiving
Author(s) are permitted to self-archive a pre-print and an author's accepted manuscript version of

their Article.

a. A pre-print is the author's version of the Article before peer-review has taken place ("Pre-Print"). Prior to acceptance for publication, Author(s) retain the right to make a Pre-Print of their Article available on any of the following: their own personal, self-maintained website; a legally compliant, non-commercial pre-print server such as but not limited to arXiv and bioRxiv. Once the Article has been published, the Author(s) should update the acknowledgement and provide a link to the definitive version on the publisher's website: "This is a pre-print of an article published in [insert journal title]. The final authenticated version is available online at: https://doi.org/[insert DOI]".
b. An Author's Accepted Manuscript (AAM) is the version accepted for publication in a journal following peer review but prior to copyediting and typesetting that can be made available under the following conditions:
    a. Author(s) retain the right to make an AAM of their Article available on their own personal, self-maintained website immediately on acceptance,
    b. Author(s) retain the right to make an AAM of their Article available for public release on any of the following 12 months after first publication ("Embargo Period"): their employer's internal website; their institutional and/or funder repositories. AAMs may also be deposited in such repositories immediately on acceptance, provided that they are not made publicly available until after the Embargo Period.

An acknowledgement in the following form should be included, together with a link to the published version on the publisher's website: "This is a post-peer-review, pre-copyedit version of an article published in [insert journal title]. The final authenticated version is available online at: http://dx.doi.org/[insert DOI]".

4. Authors' Retained Rights

Author(s) retain the following non-exclusive rights for the published version provided that, when reproducing the Article or extracts from it, the Author(s) acknowledge and reference first publication in the Journal:

a. to reuse graphic elements created by the Author(s) and contained in the Article, in presentations and other works created by them;
b. they and any academic institution where they work at the time may reproduce the Article for the purpose of course teaching (but not for inclusion in course pack material for onward sale by libraries and institutions); and
c. to reproduce, or to allow a third party Assignee to reproduce the Article in whole or in part in any printed volume (book or thesis) written by the Author(s).

5. Warranties

The Author(s) warrant and represent that:

a. (i) the Author(s) are the sole copyright owners or have been authorised by any additional copyright owner(s) to assign the rights defined in clause 2, (ii) the Article does not infringe any intellectual property rights (including without limitation copyright, database rights or trade mark rights) or other third party rights and no licence from or payments to a third party are required to publish the Article, (iii) the Article has not been previously published or licensed, (iv) if the Article contains material from other sources (e.g. illustrations, tables, text quotations), Author(s) have obtained written permissions to the extent necessary from the copyright holder(s), to license to the Assignee the same rights as set out in Clause 2 but on a non-exclusive basis and without the right to use any graphic elements on a stand-alone basis and have cited any such material correctly;
b. all of the facts contained in the Article are according to the current body of science true and accurate;
c. nothing in the Article is obscene, defamatory, violates any right of privacy or publicity, infringes any other human, personal or other rights of any person or entity or is otherwise unlawful and that informed consent to publish has been obtained for all research participants;
d. nothing in the Article infringes any duty of confidentiality which any of the Author(s) might owe to anyone else or violates any contract, express or implied, of any of the Author(s). All of the institutions in which work recorded in the Article was created or carried out have authorised and approved such research and publication; and

e. the signatory (the Author or the employer) who has signed this agreement has full right, power and authority to enter into this agreement on behalf of all of the Author(s).

6. Cooperation
The Author(s) shall cooperate fully with the Assignee in relation to any legal action that might arise from the publication of the Article, and the Author(s) shall give the Assignee access at reasonable times to any relevant accounts, documents and records within the power or control of the Author(s). The Author(s) agree that the distributing entity is intended to have the benefit of and shall have the right to enforce the terms of this agreement.

7. Author List
After signing, changes of authorship or the order of the authors listed will not be accepted unless formally approved in writing by the Assignee.

8. Edits & Corrections
The Author(s) agree(s) that the Assignee may retract the Article or publish a correction or other notice in relation to the Article if the Assignee considers in its reasonable opinion that such actions are appropriate from a legal, editorial or research integrity perspective.

This is an automated e-mail; please do not reply to this account. If you have any questions, please go to our help pages.

Thank you very much.

Kind regards,

Springer Author Services

Article Details

Journal title	Article title
Neurocritical Care	Genetic variability in the iron homeostasis pathway and patient outcomes after aneurysmal subarachnoid hemorrhage
DOI	**Corresponding Author**
10.1007/s12028-020-00961-z	Lacey Heinsberg
Copyright transferred to	**Transferred on**
Springer Science+Business Media, LLC, part of Springer Nature and Neurocritical Care Society	Tue Mar 17 19:53:51 CET 2020

Springer Nature Customer Service Center

Tiergartenstr. 15-17
69121 Heidelberg
Germany
phone: +49 6221 345 0
fax: +49 6221 345 4229
customerservice@springernature.com

Springer New York, LCC

233 Spring Street
New York, NY 10013
USA
phone: +1 212 460 1500 or 800-SPRINGER
(Weekdays 8:30am - 5:30pm ET)
fax: +1 212-460-1700
customerservice@springernature.com

Appendix E Manuscript 3: Evaluation of *APOE* Genotype and Ability to Perform Activities of Daily Living Following Aneurysmal Subarachnoid Hemorrhage

Biological Research for Nursing
2018, Vol. 20(2) 177-182

DOI: 10.1177/1099800417746504
journals.sagepub.com/home/brn
SAGE

Evaluation of *APOE* Genotype and Ability to Perform Activities of Daily Living Following Aneurysmal Subarachnoid Hemorrhage

Lacey Wright Heinsberg, BSN, BS, RN[1], Eleanor Turi, BSN, RN[2], Dianxu Ren, PhD[1], Elizabeth Crago, PhD, RN[1], Sheila Alexander, PhD, RN, FCCM[1], Ansley Grimes Stanfill, PhD, RN[3,4], and Yvette Perry Conley, PhD, FAAN[1,5]

Abstract

Survivors of aneurysmal subarachnoid hemorrhage (aSAH) often experience unfavorable functional outcomes that result in a reduced ability to independently perform activities of daily living. The apolipoprotein E gene (*APOE*) encodes for a protein known to facilitate lipid transport and aid in neuronal repair within the central nervous system and to moderate the inflammatory response, making functional variations in this gene likely candidate biomarkers to predict outcomes following aSAH. In the present work, we examined the relationship between *APOE* genotype and the ability to perform activities of daily living as measured by the Barthel Index (BI) score at 3 months ($n = 298$) and 12 months ($n = 288$) following aSAH. *APOE* genotypes were determined using polymerase chain reaction followed by restriction digestion and gel electrophoresis and treated as binary variables depending on the presence or absence of alleles E4 and E2. Multiple logistic regression was used to determine whether *APOE* genotype accounted for variability in BI score after controlling for age, sex, and severity of clinical condition as measured by the Hunt and Hess classification. No significant association was found between the presence of allele E4 and BI score at 3 ($p = .20$) or 12 months ($p = .29$) or between the presence of allele E2 and BI score at 3 ($p = .23$) or 12 months ($p = .86$) after controlling for covariates. The results of this study do not support a relationship between *APOE* genotype and the ability to perform activities of daily living after aSAH.

Keywords

subarachnoid hemorrhage, apolipoprotein E, genetics, polymorphism, Barthel Index score, activities of daily living

Aneurysmal subarachnoid hemorrhage (aSAH) is a devastating form of stroke with a disconcertingly young mean age of onset of around 55 years. Mortality rates for aSAH approach 45–50% within the first 30 days, and those who do survive are often left with severe functional, cognitive, and emotional deficits that make daily functioning difficult (van Gijn, Kerr, & Rinkel, 2007). The inability to perform activities of daily living leaves these survivors greatly dependent on caregivers and with a reduced quality of life (Al-Khindi, MacDonald, & Schweizer, 2010). While predictors of unfavorable outcomes after aSAH exist, including initial extent of bleed and the development of delayed cerebral ischemia (Teo et al., 2017), clinicians remain unable to comprehensively identify all patients at risk. This gap in knowledge impedes the development of personalized therapeutic interventions, rehabilitation provisions, and long-term care planning to maximize outcomes for this population.

Recent data suggest that inflammatory mechanisms and genetic variations may contribute to poor outcomes following aSAH (Chaudhry et al., 2017). Specifically, research has shown that variants in the apolipoprotein E gene (*APOE*) play an important role in moderating patient functional and neuropsychological outcomes following several forms of neurologic injury, including intracerebral hemorrhage (James, Blessing, Bennett, & Laskowitz, 2009), traumatic brain injury (Li et al., 2015), and subarachnoid hemorrhage (Lanterna et al.,

[1] School of Nursing, University of Pittsburgh, Pittsburgh, PA, USA
[2] Children's Hospital of Pittsburgh of UPMC, Pittsburgh, PA, USA
[3] College of Nursing, The University of Tennessee Health Science Center, Memphis, TN, USA
[4] College of Medicine, The University of Tennessee Health Science Center, Memphis, TN, USA
[5] Department of Human Genetics, University of Pittsburgh, Pittsburgh, PA, USA

Corresponding Author:
Lacey Wright Heinsberg, BSN, BS, RN, University of Pittsburgh School of Nursing, 440 Victoria Building, 3500 Victoria Street, Pittsburgh, PA 15261, USA.
Email: law145@pitt.edu

2007). Perhaps most recognized for its strong influences in mild cognitive impairment, Alzheimer's disease, and dementia (Viticchi et al., 2014), variation in *APOE* is widely accepted to play a strong role in cerebrovascular diseases. Although the exact mechanism of *APOE* involvement in the pathological processes following aSAH is unclear, the protein for which *APOE* encodes is well-known as a potent facilitator of lipid transport and neuronal repair, opponent of oxidative stress, and isoform-specific regulator of inflammatory responses (Colton, Brown, Czapiga, & Vitek, 2002) within the central nervous system.

Three major *APOE* alleles (E2, E3, and E4) are inherited as six common *APOE* genotypes (*APOE* 2/2, *APOE* 2/3, *APOE* 2/4, *APOE* 3/3, *APOE* 3/4, and *APOE* 4/4; Mahley, 1988). The antioxidant ability of *APOE* is isoform-specific, with E2 having the highest antioxidant potential. In contrast, E4 has the lowest antioxidant potential, suggesting the potential for patients with this allele to be less likely to release lipids for neuronal repair compared to those with other isoforms (Jolivalt et al., 2000). Although numerous studies have examined the role of *APOE* genotype in patient outcomes following aSAH, with mixed results, the relationship between *APOE* genotype and the ability to perform activities of daily living following aSAH specifically has received less attention. This gap in knowledge is important, as potential modification of care plans based on genotype has widespread nursing implications as the science moves toward personalized medicine.

Material and Method

Sample

We prospectively recruited participants for this exploratory genetic study as part of an ongoing National Institutes of Health–funded study ("the parent project") approved by the University of Pittsburgh Institutional Review Board. Participants were recruited from UPMC Presbyterian Hospital in Pittsburgh, PA, and were included in the parent project if they were between the ages of 18 and 75 years, had been diagnosed with aSAH verified with cerebral angiogram, and were able to read/speak English. Participants were excluded if they had a history of a debilitating neurological disorder or if their subarachnoid hemorrhage resulted from trauma, arteriovenous malformation, or mycotic aneurysm. Participants undergoing clinical data collection for the parent project were concurrently consented for the collection of a repository genetic sample. We obtained consent from a legal proxy if the participant was unable to provide personal consent. Participant demographics and clinical data, including age, sex, and severity of aSAH (as measured by the clinical grading system, the Hunt and Hess [HH] classification), were extracted from the medical record.

Genotype data were available for 524 participants. We excluded 116 participants from the present analysis due to missing outcome data at both 3 and 12 months. Given that allele frequencies differed by ancestry, we also excluded 62 non-Caucasian participants to control for population stratification, significantly reducing the chance of spurious findings based solely on the underlying structure of the population as opposed to an actual association with our locus of interest (Little et al., 2009). The non-Caucasian group was not large enough for the conduct of subgroup analyses. The final sample consisted of 346 Caucasian participants; sample size varied at cross-sectional outcome time points based on availability of outcome data (at 3 months, $n = 298$; at 12 months, $n = 288$).

Genotyping

We performed *APOE* genotyping using DNA extracted from either cerebrospinal fluid or whole blood collected at recruitment, as the DNA template will not vary depending on the tissue used. DNA was extracted using a simple salting-out procedure (Miller, Dykes, & Polesky, 1988), and genotypes were determined using polymerase chain reaction followed by restriction digestion and gel electrophoresis. These genotypes were compared to genetic ladders and to standards of known, sequenced genotypes. Individuals blinded to phenotypes double called all genotypes, and we resolved genotype discrepancies by retyping any disputed samples.

Outcome Phenotyping

A trained neuropsychiatric technician or a nurse trained in neuropsychiatric testing evaluated participants' ability to perform activities of daily living via face-to-face or telephone interview at 3 and 12 months following aSAH using the Barthel Index (BI). The BI score is a composite measure of the ability to perform activities of daily living including functions such as toileting, grooming, dressing, mobility, and transfer. The interviewer rated items on a scale ranging from 0 (*death or unable to perform*) to 2 (*can independently perform*) to produce a composite score ranging from 0 to 20, with lower scores indicating lesser ability to perform activities of daily living. The BI score has demonstrated excellent test–retest reliability (Cohen's $\kappa = .41–1.00$) and internal consistency (Cronbach $\alpha = .80–.93$) in stroke patients (Quinn, Langhorne, & Stott, 2011). Importantly, telephone administration has a reliability comparable to that of face-to-face administration (Della Pietra et al., 2011). Given that some participants were impaired, interviewers administered the BI questionnaire to either the participant or the primary caregiver.

Statistical Analysis

We carried out all statistical analyses using SPSS, Version 24 (IBM, Chicago, IL). First, we conducted a preliminary analysis to evaluate relationships between variables, compute detailed descriptive statistics, examine data graphically, and investigate sources of missing or extreme data. Given the distributions of genotypes within this sample, genotype subgroup analyses were not possible. We treated genotypes as binary variables and classified them first based on the presence or absence of at least one *APOE* E4 allele (*APOE* 2/4, *APOE* 3/4, and *APOE* 4/4 vs. *APOE* 2/2, *APOE* 2/3, and *APOE* 3/3) and again based

Table 1. Demographics and Univariate Analysis to Evaluate Relationships Between Variables Based on Presence or Absence of *APOE* Alleles E4 and E2.

		APOE E4			*APOE* E2		
	Sample	E4 Present (*APOE* 2/4, *APOE* 3/4, *APOE* 4/4)	E4 Absent (*APOE* 2/2, *APOE* 2/3, *APOE* 3/3)		E2 Present (*APOE* 2/2, *APOE* 2/3, *APOE* 2/4)	E2 Absent (*APOE* 3/3, *APOE* 3/4, *APOE* 4/4)	
Variable	$N = 346$	$n = 84$	$n = 262$	p Value	$n = 42$	$n = 304$	p Value
Age (years)				.45			.67
Mean (*SD*)	53.9 (11.3)	54.8 (12.2)	53.7 (11.0)		54.6 (10.9)	53.8 (11.4)	
Range	18–75	18–75	25–75		33–75	18–75	
Sex, female, *n* (%)	239 (69.1)	62 (25.9)	177 (74.1)	.28	30 (12.6)	209 (87.4)	.73
HH grade, *n* (%)				.62			.41
1–2	169 (48.8)	43 (25.4)	126 (74.6)		18 (10.7)	151 (89.3)	
3–5	177 (51.2)	41 (23.2)	136 (76.8)		24 (13.6)	153 (86.4)	

Note. HH = Hunt and Hess grade; *APOE* = apolipoprotein E gene; *SD* = standard deviation.

Table 2. Univariate Analysis to Evaluate Relationships Between Demographic Variables and HH Grades and Outcome, as Measured by the 3- and 12-Month BI Scores.

		3-Month BI Score			12-Month BI Score		
	Sample	Unfavorable	Favorable		Unfavorable	Favorable	
Variable	$N = 346$	$n = 75$	$n = 223$	p Value	$n = 59$	$n = 229$	p Value
Age (years)				.003			.10
Mean (*SD*)	53.9 (11.3)	56.9 (11.4)	52.8 (10.9)		56.4 (11.9)	54.3 (11.2)	
Range	18–75	18–75	24–75		18–75	24–75	
Sex, female, *n* (%)	239 (69.1)	54 (26.5)	150 (73.5)	.45	41 (20)	164 (80)	.75
HH grade, *n* (%)				<.001			<.001
1–2	169 (48.8)	14 (9.9)	128 (90.1)		9 (6.3)	134 (93.7)	
3–5	177 (51.2)	61 (39.1)	95 (60.9)		50 (34.5)	95 (65.5)	

Note. Unfavorable BI scores were in the range 0–18; favorable BI scores were in the range 19–20. HH = Hunt and Hess grade; BI = Barthel Index; *SD* = standard deviation.

on the presence or absence of at least one *APOE* E2 allele (*APOE* 2/2, *APOE* 2/3, *APOE* 2/4 vs. *APOE* 3/3, *APOE* 3/4, and *APOE* 4/4). Similarly, use of the full range of BI score values was not possible given the sample size and skewness of the data. Therefore, we dichotomized the BI score as unfavorable (0–18) versus favorable (19–20). We chose a range of $\geq$19 to indicate a favorable outcome because it is the optimum cut-off score for excellent recovery after stroke, corresponding to a modified Rankin Scale score of 1 (sensitivity 85.6%; specificity 91.7%; Uyttenboogaart, Stewart, Vroomen, De Keyser, & Luijckx, 2005). We dichotomized the covariate HH grade as low (1 or 2) or high (3, 4, or 5) based on previously reported differences in patient outcomes in these groups (Lantigua et al., 2015). We evaluated associations between genotype groups and BI score using logistic regression. In the final stage of analysis, our statistical modeling considered genotype groups separately to yield unadjusted regression coefficients, and we then expanded it to yield adjusted estimates controlling for age, sex, and HH grade. We considered results to be statistically significant if the p value was $\leq$.05.

Results

Participants in the present study had a mean age of 53.9 years and were 69.1% female. Age, sex, and HH grade for participants in this study did not differ significantly from those of the parent project. Participant demographics and HH grades are presented for the total sample and by allele group in Table 1. *APOE* genotype distribution showed *APOE* 3/3 as the most common (65.6%), followed by *APOE* 3/4 (21.4%), *APOE* 2/3 (10.1%), *APOE* 2/4 (2%), and *APOE* 4/4 (0.9%). The *APOE* 2/2 genotype did not occur in the present sample, but baseline demographic and clinical data did not differ between *APOE* groups.

Next, we conducted a preliminary analysis to evaluate the relationships between outcome, as measured by BI scores, and both demographics and HH grades. We present the results in Table 2. The percentage of unfavorable BI scores at 3 and 12 months was 25.2% and 20.5%, respectively. Increased age was associated with unfavorable BI score at 3 months (p = .003) but not at 12 months (p = .10). Higher HH grade was

Table 3. Results From Univariate and Multivariate Analyses Exploring Associations Between Presence of *APOE* Alleles E4 and E2 With 3- and 12-Month BI Score.

		APOE E4 Present (24.3%; *APOE* 2/4, *APOE* 3/4, *APOE* 4/4 vs. *APOE* 2/2, *APOE* 2/3, *APOE* 3/3)			*APOE* E2 Present (12.1%; *APOE* 2/2, *APOE* 2/3, *APOE* 2/4 vs. *APOE* 3/3, *APOE* 3/4, *APOE* 4/4)		
Variable	*n*	*OR*	95% CI	*p* Value	*OR*	95% CI	*p* Value
3-Month BI	298	.73	[0.38, 1.38]	.33	.71	[0.31, 1.62]	.42
12-Month BI	288	.66	[0.32, 1.36]	.26	.97	[0.40, 2.34]	.94
Adjusted estimate[a]							
3-Month BI	298	.64	[0.32, 1.27]	.20	.58	[0.24, 1.40]	.23
12-Month BI	288	.66	[0.31, 1.43]	.29	.92	[0.34, 2.36]	.86

Note. BI = Barthel Index; *APOE* = apolipoprotein E gene; CI = confidence interval; *OR* = odds ratio.
[a]Adjusted estimate controlling for age, sex, and Hunt and Hess grade.

associated with unfavorable BI score at 3 months ($p < .001$) and 12 months ($p < .001$). In contrast, sex was not associated with the ability to perform activities of daily living at 3 months ($p = .45$) or 12 months ($p = .75$).

Finally, we evaluated associations between *APOE* genotype and ability to perform activities of daily living and present the results in Table 3. We found no significant association between the presence of allele E4 and BI score at 3 months ($p = .20$) or 12 months ($p = .29$) after controlling for age, sex, and HH grade. Likewise, we found no significant association between the presence of allele E2 and BI score at 3 months ($p = .23$) or 12 months ($p = .86$) after controlling for age, sex, and HH grade.

Discussion

While *APOE* genotype has surfaced as an important predictor of select patient outcomes following traumatic brain injury (Li et al., 2015), intracerebral hemorrhage (James et al., 2009), and aSAH (Lanterna et al., 2007), the findings of the present ancillary genetic study do not support a relationship between *APOE* genotype and the ability to perform activities of daily living as measured by BI score following aSAH. Our work is in agreement with data from a study involving a combined population of ischemic and hemorrhagic stroke victims in which investigators found no relationship between *APOE* genotype and BI score (Wagle et al., 2010). Importantly, a similar analysis in a smaller sample size of aSAH patients corroborated the results of this study (Csajbok et al., 2016).

The lack of a definitive association between *APOE* genotype and patient outcomes suggests that any relationship that may exist is part of a complex network moderated by many factors. It is possible that the link between *APOE* genotype is more strongly linked to functional impairment via cognitive changes (Andrews, Das, Cherbuin, Anstey, & Easteal, 2016) and that a direct relationship between genotype and physical changes relevant to the ability to perform activities of daily living may not exist. Furthermore, the outcome measure that we selected for examination in this study provided a mean measure of ability to perform activities of daily living but did not examine additional measures of functional outcomes such as the Glasgow Outcome Scale (GOS), which researchers have found to be associated with *APOE* genotype following aSAH (Lanterna et al., 2007) and intracerebral hemorrhage (James et al., 2009). Although both instruments are well established measures of functional outcome, the BI Scale asks a multitude of questions specific to everyday activities, while the GOS is a gross measure of functional outcome with a smaller range of possible scores (1–5). While the scopes of both tools are important considerations for quality of life for aSAH patients, the results of the present study suggest that these outcomes could have very different biological underpinnings. Thus, it is possible that our results would be different if we had considered GOS score rather than BI score for our primary outcome measure. It is also possible that a self-report of one's ability to perform activities of daily living, like the BI, could be biased, with patients tending to overestimate their abilities (Shih, Rogers, Skidmore, Irrgang, & Holm, 2009). Lastly, in a review of 15 stroke trials, authors reported a remarkable difference in the choice of cutoff points for defining favorable BI scores compared with other measures of functional outcomes following stroke (Sulter, Steen, & De Keyser, 1999), which significantly hinders the comparison of the results of the present study with those of other studies examining the influence of *APOE* genotype on additional patient outcomes.

While the present study contributes to the existing literature finding no association between *APOE* genotype and the ability to perform activities of daily living following aSAH, there are several important limitations that we should acknowledge. First, the contribution of *APOE* genotype to recovery and the ability to perform activities of daily living in aSAH patients may have a small effect size that is not detectable in this sample. We selected a BI cutoff score of $\geq$19 to distinguish those with excellent recovery from those with dysfunction in ability to perform activities of daily living. Use of the full range of values, while not possible given this sample size, might show finer differences based on genotype. Our sample had many subjects without dysfunction in the ability to perform activities of daily living, which may have impaired our ability to identify these finer differences. Furthermore, our sample includes an insufficient number of homozygous risk alleles for subgroup analyses by all possible *APOE* genotypes separately. Specifically, there is underrepresentation of the two genotypes

reported to be most influential, *APOE* 2/2 and *APOE* 4/4. Given that isoform-specific proteins have different antioxidant ability (Jolivalt et al., 2000), it is likely that each isoform is uniquely relevant in the pathological changes that occur within the central nervous system following aSAH. It may be that one copy of the common *APOE* 3 allele is adequate to support neuronal recovery, and we were therefore unable to detect dose-specific effects. It is also possible that people with homozygous E4 alleles could experience a higher early mortality rate and are, therefore, less likely to be admitted to the neurovascular intensive care unit to begin with, thus biasing this study. The present study was further limited in that it included a self-reported Caucasian sample. Future efforts are needed to examine this relationship in patients with more diverse ancestries, an especially crucial requirement considering the nonrandom geographical distribution of the E4 allele as well as cultural factors important to aSAH, including diet (Huebbe & Rimbach, 2017). Lastly, a majority of the sample examined in this study was female (69.1%), which is consistent with the reported demographics of the aSAH patient population (Kongable et al., 1996). While we found no association between sex and BI score and did control for this in the final analysis, there may be associations that are important only when stratified by sex.

It is a strength of this study that we have a relatively large sample size as compared to other work in the area (Csajbok et al., 2016). Our findings also indicated that age and HH grade together are strong predictors of future ability to perform activities of daily living, which are consistent with the literature. Information such as this is critically needed given the heterogeneity of outcomes following aSAH. Despite the null genotypic findings of this study, there remains a need for further exploration of genotype and gene–gene interactions as predictors of outcome variability for these patients. Genetic biomarkers hold great promise for impacting nursing science in the future. By identifying aSAH patients at risk for unfavorable outcomes, genetic biomarkers can aid nurses in (1) targeting these patients with early personalized therapeutic interventions to mitigate unfavorable outcomes, (2) developing detailed discharge education materials to caution patients and families about potential or likely complications, and (3) care planning for the future to implement comprehensive rehabilitation plans as early as possible.

Authors' Note

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author Contribution

L. Heinsberg contributed to conception, design, analysis, and interpretation; drafted the manuscript; critically revised the manuscript; gave final approval; and agrees to be accountable for all aspects of work, ensuring integrity and accuracy. E. Turi contributed to conception, design, analysis, and interpretation; critically revised the manuscript; gave final approval; and agrees to be accountable for all aspects of work, ensuring integrity and accuracy. D. Ren contributed to analysis and interpretation; critically revised the manuscript; gave final approval; and agrees to be accountable for all aspects of work, ensuring integrity and accuracy. E. Crago contributed to acquisition and interpretation; critically revised the manuscript; gave final approval; and agrees to be accountable for all aspects of work, ensuring integrity and accuracy. S. Alexander contributed to acquisition and interpretation; critically revised the manuscript; gave final approval; and agrees to be accountable for all aspects of work, ensuring integrity and accuracy. A. Stanfill contributed to conception, design, and interpretation; critically revised the manuscript; gave final approval; and agrees to be accountable for all aspects of work, ensuring integrity and accuracy. Y. Conley contributed to conception, design, acquisition, and interpretation; critically revised the manuscript; gave final approval; and agrees to be accountable for all aspects of work, ensuring integrity and accuracy.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was supported by the National Institute of Nursing Research of the National Institutes of Health under award numbers F31NR017311, T32NR009759, R01NR004339, and R01NR013610.

References

Al-Khindi, T., MacDonald, R. L., & Schweizer, T. A. (2010). Cognitive and functional outcome after aneurysmal subarachnoid hemorrhage. *Stroke*, *41*, e519–e536. doi:10.1161/STROKEAHA.110.581975

Andrews, S. J., Das, D., Cherbuin, N., Anstey, K. J., & Easteal, S. (2016). Association of genetic risk factors with cognitive decline: The PATH through life project. *Neurobiology of Aging*, *41*, 150–158. doi:10.1016/j.neurobiolaging.2016.02.016

Chaudhry, S. R., Guresir, E., Vatter, H., Kinfe, T. M., Dietrich, D., Lamprecht, A., & Muhammad, S. (2017). Aneurysmal subarachnoid hemorrhage lead to systemic upregulation of IL-23/IL-17 inflammatory axis. *Cytokine*, *97*, 96–103. doi:10.1016/j.cyto.2017.05.025

Colton, C. A., Brown, C. M., Czapiga, M., & Vitek, M. P. (2002). Apolipoprotein-E allele-specific regulation of nitric oxide production. *Annals of the New York Academy of Sciences*, *962*, 212–225. doi:10.1111/j.1749-6632.2002.tb04070.x

Csajbok, L. Z., Nylén, K., Öst, M., Blennow, K., Zetterberg, H., Nellgård, P., & Nellgård, B. (2016). Apolipoprotein E polymorphism in aneurysmal subarachnoid haemorrhage in West Sweden. *Acta Neurologica Scandinavica*, *133*, 466–474. doi:10.1111/ane.12487

Della Pietra, G. L., Savio, K., Oddone, E., Reggiani, M., Monaco, F., & Leone, M. A. (2011). Validity and reliability of the Barthel Index administered by telephone. *Stroke*, *42*, 2077–2079. doi:10.1161/STROKEAHA.111.613521

Huebbe, P., & Rimbach, G. (2017). Evolution of human apolipoprotein E (APOE) isoforms: Gene structure, protein function and interaction with dietary factors. *Ageing Research Reviews*, *37*, 146–161. doi:10.1016/j.arr.2017.06.002

James, M. L., Blessing, R., Bennett, E., & Laskowitz, D. T. (2009). Apolipoprotein E modifies neurological outcome by affecting cerebral edema but not hematoma size after intracerebral hemorrhage

in humans. *Journal of Stroke and Cerebrovascular Diseases*, *18*, 144–149. doi:10.1016/j.jstrokecerebrovasdis.2008.09.012

Jolivalt, C., Leininger-Muller, B., Bertrand, P., Herber, R., Christen, Y., & Siest, G. (2000). Differential oxidation of apolipoprotein E isoforms and interaction with phospholipids. *Free Radical Biology and Medicine*, *28*, 129–140. doi:10.1016/S0891-5849(99)00232-4

Kongable, G. L., Lanzino, G., Germanson, T. P., Truskowski, L. L., Alves, W. M., Torner, J. C., & Kassell, N. F. (1996). Gender-related differences in aneurysmal subarachnoid hemorrhage. *Journal of Neurosurgery*, *84*, 43–48. doi:10.3171/jns.1996.84.1.0043

Lanterna, L. A., Ruigrok, Y., Alexander, S., Tang, J., Biroli, F., Dunn, L. T., & Poon, W. S. (2007). Meta-analysis of APOE genotype and subarachnoid hemorrhage: Clinical outcome and delayed ischemia. *Neurology*, *69*, 766–775. doi:10.1212/01.wnl.0000267640.03300.6b

Lantigua, H., Ortega-Gutierrez, S., Schmidt, J. M., Lee, K., Badjatia, N., Agarwal, S., … Mayer, S. A. (2015). Subarachnoid hemorrhage: Who dies, and why? *Critical Care*, *19*, 309. doi:10.1186/s13054-015-1036-0

Li, L., Bao, Y., He, S., Wang, G., Guan, Y., Ma, D., … Yang, J. (2015). The association between apolipoprotein E and functional outcome after traumatic brain injury. *Medicine*, *94*, 1–7. doi:10.1097/MD.0000000000002028

Little, J., Higgins, J. P. T., Ioannidis, J. P. A., Moher, D., Gagnon, F., von Elm, E., … Birkett, N. (2009). STrengthening the REporting of Genetic Association studies (STREGA)—An extension of the STROBE statement. *European Journal of Clinical Investigation*, *39*, 247–266. doi:10.1111/j.1365-2362.2009.02125.x

Mahley, R. (1988). Apolipoprotein E: Cholesterol transport protein with expanding role in cell biology. *Science*, *240*, 622–630. doi:10.1126/science.3283935

Miller, S. A., Dykes, D. D., & Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, *16*, 1215. doi:10.1093/nar/16.3.1215

Quinn, T. J., Langhorne, P., & Stott, D. J. (2011). Barthel Index for stroke trials: Development, properties, and application. *Stroke*, *42*, 1146–1151. doi:10.1161/STROKEAHA.110.598540

Shih, M. M., Rogers, J. C., Skidmore, E. R., Irrgang, J. J., & Holm, M. B. (2009). Measuring stroke survivors' functional status independence: Five perspectives. *American Journal of Occupational Therapy*, *63*, 600–608. doi:10.5014/ajot.63.5.600

Sulter, G., Steen, C., & De Keyser, J. (1999). Use of the Barthel Index and modified Rankin Scale in acute stroke trials. *Stroke*, *30*, 1538–1541. doi:10.1161/01.STR.30.8.1538

Teo, M., Guilfoyle, M. R., Turner, C., Kirkpatrick, P. J., Kirkpatrick, P. J., Turner, C. L., … Quinn, L. (2017). What factors determine treatment outcome in aneurysmal subarachnoid hemorrhage in the modern era? A post hoc STASH analysis. *World Neurosurgery*, *105*, 270–281. doi:10.1016/j.wneu.2017.05.005

Uyttenboogaart, M., Stewart, R. E., Vroomen, P. C. A. J., De Keyser, J., & Luijckx, G. J. (2005). Optimizing cutoff scores for the Barthel Index and the modified Rankin Scale for defining outcome in acute stroke trials. *Stroke*, *36*, 1984–1987. doi:10.1161/01.STR.0000177872.87960.61

van Gijn, J., Kerr, R. S., & Rinkel, G. J. E. (2007). Subarachnoid haemorrhage. *Lancet*, *369*, 306–318. doi:10.1016/S0140-6736(07)60153-6

Viticchi, G., Falsetti, L., Vernieri, F., Altamura, C., Altavilla, R., Luzzi, S., … Silvestrini, M. (2014). Apolipoprotein E genotype and cerebrovascular alterations can influence conversion to dementia in patients with mild cognitive impairment. *Journal of Alzheimer's Disease*, *41*, 401–410. doi:10.3233/JAD-132480

Wagle, J., Farner, L., Flekkøy, K., Wyller, T. B., Sandvik, L., Eiklid, K. L., … Engedal, K. (2010). Cognitive impairment and the role of the ApoE epsilon4-allele after stroke—A 13 months follow-up study. *International Journal of Geriatric Psychiatry*, *25*, 833–842. doi:10.1002/gps.2425

Appendix E.1 License Agreement for Manuscript 3

Dear Mrs. Lacey Heinsberg,

Your SAGE Publications request has been cancelled for the following reason:
Dear Lacey Heinsberg,

Thank you for your order. Under SAGE's journal author policy, you are permitted to include your own articles within your thesis in a publicly accessible platform, such as your institutional repository. You must provide a full citation and the permission does not include any third-party material. If you require formal written permission, please place an order via our Customer Permission Portal (https://jira.sagepub.com/servicedesk/customer/portal/9).

As we have replied via this comment, we will cancel this order. If we can be of further assistance, please contact us via our customer portal.

Best regards,
SAGE Rights Team.

You will not be charged for this order.

Appendix F Manuscript 4: Symptom Science: Advocating for Inclusion of Functional Genetic Polymorphisms

Article

Symptom Science: Advocating for Inclusion of Functional Genetic Polymorphisms

Biological Research for Nursing
2019, Vol. 21(4) 349-354

DOI: 10.1177/1099800419846407
journals.sagepub.com/home/brn
SAGE

Mitchell R. Knisely, PhD, RN[1], Megan Maserati, PhD[1], Lacey W. Heinsberg, BSN, RN, PhD[1], Lisa L. Shah, PhD, RN[1], Hongjin Li, MS, BSN, PhD[1], Yehui Zhu, MSN, PhD[1], Yumi Ma, MSN, RN, PhD[1], Letitia Y. Graves, MSN, RN, PhD[1], John D. Merriman, PhD, RN[2], and Yvette P. Conley, PhD, FAAN[1]

Abstract

Incorporating biologically based data into symptom science research can contribute substantially to understanding commonly experienced symptoms across chronic conditions. The purpose of this literature review was to identify functional polymorphisms associated with common symptoms (i.e., pain, sleep disturbance, fatigue, affective and cognitive symptoms) with the goal of identifying a parsimonious list of functional genetic polymorphisms with evidence to advocate for their inclusion in symptom science research. PubMed was searched to identify genes and functional polymorphisms associated with symptoms across chronic conditions, revealing eight functional genetic polymorphisms in seven different genes that showed evidence of association with at least three or more symptoms and/or symptom clusters: BDNF rs6265, COMT rs4680, FKBP5 rs3800373, IL-6 rs1800795, NFKB2 rs1056890, SLC6A4 5-HTTLPRþ rs25531, and TNFA rs1799964 and rs1800629. Of these genes, three represent protein biomarkers previously identified as common data elements for symptom science research (BDNF, IL-6, and TNFA), and the polymorphisms in these genes identified through the search are known to impact secretion or level of transcription of these protein biomarkers. Inclusion of genotype data for polymorphisms offers great potential to further advance scientific knowledge of the biological basis of individual symptoms and symptom clusters across studies. Additionally, these polymorphisms have the potential to be used as targets to optimize precision health through the identification of individuals at risk for poor symptom experiences as well as the development of symptom management interventions.

Keywords

symptom science, functional genetic polymorphism, biomarker

Symptom science focuses on characterizing symptoms and understanding associated biological and behavioral mechanisms that can support clinical applications that prevent or alleviate symptoms (Cashion & Grady, 2015). Researchers and funding agencies have made numerous calls to action to integrate omics-based approaches into symptom science research (Corwin et al., 2014; Miaskowski et al., 2017; National Institute of Nursing Research [NINR], 2016; Taylor & Barcelona de Mendoza, 2018). Notably, omics-based approaches offer the potential to aid in identifying risk factors and understanding the biological underpinnings of symptoms. Deciphering the biological underpinnings of individual symptoms or symptom clusters has the potential to optimize precision health approaches to symptom management through the identification of at-risk individuals and targets for pharmacological and non-pharmacological interventions.

Evidence supports shared biological pathways for and genomic influences on symptom development. A recent review identified common genes and biological pathways to be considered among five frequently experienced symptoms including sleep disturbance, cognitive impairment, fatigue, gastrointestinal distress, and pain (McCall et al., 2018). While this review identified candidate genes to be considered across studies of common symptoms, it did not identify specific genetic polymorphisms. Likewise, a recent position paper proposed several protein biomarkers as common data elements (CDEs) that may mediate or moderate symptom experiences (Page et al., 2018). Functional genetic polymorphisms are

[1] School of Nursing, University of Pittsburgh, Pittsburgh, PA, USA
[2] New York University Rory Meyers College of Nursing, New York, NY, USA

Corresponding Author:
Yvette P. Conley, PhD, FAAN, School of Nursing, University of Pittsburgh, 440 Victoria Building, 3500 Victoria Street, Pittsburgh, PA 15261, USA.
Email: yconley@pitt.edu

Table 1. PubMed Literature Search Terms.

Symptom	Search Terms
Anxiety	("anxiety"[MeSH Terms] OR "anxiety"[All Fields]) AND "functional"[All Fields] AND ("polymorphism, genetic"[MeSH Terms] OR {"polymorphism"[All Fields] AND "genetic"[All Fields]} OR "genetic polymorphism"[All Fields] OR "polymorphism"[All Fields])
Cognitive dysfunction	("cognitive dysfunction"[MeSH Terms] OR {"cognitive"[All Fields] AND "dysfunction"[All Fields]} OR "cognitive dysfunction"[All Fields] OR {"cognitive"[All Fields] AND "disturbance"[All Fields]} OR "cognitive disturbance"[All Fields]) AND "functional"[All Fields] AND ("polymorphism, genetic"[MeSH Terms] OR {"polymorphism"[All Fields] AND "genetic"[All Fields]} OR "genetic polymorphism"[All Fields] OR "polymorphisms"[All Fields])
Depression	("depressive disorder"[MeSH Terms] OR {"depressive"[All Fields] AND "disorder"[All Fields]} OR "depressive disorder"[All Fields] OR "depression"[All Fields] OR "depression"[MeSH Terms]) AND "functional"[All Fields] AND ("polymorphism, genetic"[MeSH Terms] OR {"polymorphism"[All Fields] AND "genetic"[All Fields]} OR "genetic polymorphism"[All Fields] OR "polymorphism"[All Fields])
Fatigue	("fatigue"[All Fields] OR "lack of energy"[All Fields]) AND "functional"[All Fields] AND ("polymorphism, genetic"[MeSH Terms] OR {"polymorphism"[All Fields] AND "genetic"[All Fields]} OR "genetic polymorphism"[All Fields] OR "polymorphism"[All Fields])
Pain	("pain"[MeSH Terms] OR "pain"[All Fields]) AND "functional"[All Fields] AND ("polymorphism, genetic"[MeSH Terms] OR {"polymorphism"[All Fields] AND "genetic"[All Fields]} OR "genetic polymorphism"[All Fields] OR "polymorphisms"[All Fields])
Sleep	("sleep"[MeSH Terms] OR "sleep"[All Fields]) AND "functional"[All Fields] AND ("polymorphism, genetic"[MeSH Terms] OR {"polymorphism"[All Fields] AND "genetic"[All Fields]} OR "genetic polymorphism"[All Fields] OR "polymorphism"[All Fields]) ("dyssomnias"[MeSH Terms] OR "dyssomnias"[All Fields] OR {"sleep"[All Fields] AND "disturbance"[All Fields]} OR "sleep disturbance"[All Fields]) AND functional[All Fields] AND ("polymorphism, genetic"[MeSH Terms] OR {"polymorphism"[All Fields] AND "genetic"[All Fields]} OR "genetic polymorphism"[All Fields] OR "polymorphism"[All Fields])
Positive affect	("positive affect"[All Fields] OR "positive mood"[All Fields] OR "psychological wellbeing"[All Fields] OR "euthymic"[All Fields] OR {"happiness"[MeSH Terms] OR "happiness"[All Fields]}) AND ("polymorphism, genetic"[MeSH Terms] OR {"polymorphism"[All Fields] AND "genetic"[All Fields]} OR "genetic polymorphism"[All Fields] OR "polymorphism"[All Fields])

variations in DNA for which there is evidence that they impact structure, function, or level of a gene product (Albert, 2011). Functional genetic polymorphisms that are associated with multiple symptoms have the potential to provide a minimal set of stable (across tissue types and over time) targets that can be an initial step in identifying additional biological CDEs for symptom science studies. Including polymorphisms in these studies could contribute substantially to building the knowledge base and addressing limitations in symptom science research (Corwin et al., 2017; Redeker et al., 2015).

The purpose of the present review of literature was to identify genes and specific targets to measure (i.e., functional polymorphisms) associated with five common symptoms (pain, sleep disturbance, fatigue, and affective and cognitive symptoms) across chronic conditions. Our goal was to identify a parsimonious list of functional polymorphisms for which there is replication of evidence of association with symptoms and to advocate for their utility to the symptom science research community.

Method

We conducted a structured search of the literature to identify functional genetic polymorphisms associated with symptoms. We selected the primary symptoms of relevance to the NINR-supported centers of excellence (Redeker et al., 2015). These symptoms included pain, sleep disturbance, fatigue, and affective (i.e., anxiety, depressive symptoms, positive affect) and cognitive symptoms. Search terms were specific to the symptom of interest and availability of associated literature (see Supplemental Material 1–7 for more information about symptom-specific searches methods). In brief, we queried in PubMed by combining several search terms, including *polymorphism* or *functional polymorphism* and the symptom of interest. Table 1 includes a comprehensive list of search terms used for each symptom. We completed all searches prior to December 18, 2017. Coauthors discussed and reviewed each other's search results, and an iterative process of group discussions guided additional searches and more detailed reviews. Articles selected evaluated associations between a genetic polymorphism and a symptom of interest and stating a significant finding(s) within the context of that study. We used a standardized table to guide data abstraction across all symptoms, and a second author verified abstracted data.

While our initial queries did not specifically include symptom clusters, some articles retrieved reported on symptom clusters containing one or more of the symptoms of interest. We included these articles in our final results. We excluded articles if they were reviews, editorials, preclinical studies, or written in a language other than English. Data extracted from selected articles included the author, year, gene, polymorphism(s), symptom phenotype, context (e.g., sample details), and

Table 2. Functional Gene Polymorphisms with Multisymptom Associations.

Gene	SNP	Anxiety	Cognitive Symptoms	Depressive Symptoms	Fatigue	Pain	Sleep	Positive Affect	Symptom Cluster
BDNF	rs6265	X	X	X	X	X	X		
COMT	rs4680	X	X	X		X	X	X	
FKBP5	rs3800373	X		X		X			
IL-6	rs1800795	X		X	X	X			
NFKB2	rs1056890				X		X		X[a,b]
SLC6A4	5-HTTLPR + rs25531	X	X	X	X	X	X	X	
TNFA	rs1799964	X		X		X			
	rs1800629		X		X	X	X		

Note. SNP = single nucleotide polymorphism.
[a]Mood-cognitive symptom cluster includes difficulty concentrating, feeling sad, worrying, itching, and feeling irritable. [b]Sickness behavior symptom cluster includes pain, lack of energy, feeling drowsy, difficulty sleeping, and sweats (Miaskowski et al., 2017).

relevant findings. Details for symptom-specific search strategies, Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagrams, and complete tables with extracted data from relevant articles are provided in Supplementary Material 1–7.

We synthesized the findings of the preliminary search described above in a single table displaying the associations between genetic polymorphisms and symptoms, with polymorphisms as the rows and each symptom as a column. Polymorphisms associated with two or more symptoms were then included in the second phase of our search to verify results and to identify any potential article that we had not captured in the initial search. In this phase, we searched PubMed using the specific polymorphism reference (rs) number and the symptom keywords used in the initial search. We reviewed results using the same criteria as listed above. The threshold for final inclusion was determined by evaluating all data and group discussion. As a result, polymorphisms included in the final list had evidence of association with three or more of the symptoms of interest and/or a symptom cluster that contained one or more of those symptoms. Including polymorphisms that were associated with three or more symptoms resulted in the most comprehensive yet parsimonious list.

Results

Table 2 includes a synthesis of this review highlighting eight functional genetic polymorphisms that showed evidence of association with three or more symptoms and/or with symptom clusters. The solute carrier family 6 member 4 (*SLC6A4*) 5-HTTLPR+rs25531 polymorphism was associated with all seven symptoms. Brain-derived neurotrophic factor (*BDNF*) rs6265 had evidence of association with all symptoms except positive affect. Catechol-O-methyltransferase (*COMT*) rs4680 had associations with all symptoms except fatigue. Another two polymorphisms were associated with four separate symptoms: Interleukin 6 (*IL-6*) rs1800795 was associated with anxiety and depressive symptoms, pain, and fatigue, whereas tumor necrosis factor alpha (*TNFA*) rs1800629 had associations with pain, sleep, fatigue, and cognitive symptoms.

Both FK506 binding protein 5 (*FKBP5*) rs3800373 and *TNFA* rs1799964 showed evidence of association with three separate symptoms: anxiety, depression, and pain. Furthermore, nuclear factor kappa B subunit 2 (*NFKB2*) rs1056890 had evidence of associations with both sleep disturbance and fatigue as well as two symptom clusters (i.e., mood-cognitive or sickness behavior symptom clusters). Table 3 provides a summary of the putative functional impact of each polymorphism. Symptom-specific findings and associated references are provided in the Supplementary Materials.

Discussion

This review is the first to synthesize the literature to identify associations between functional genetic polymorphisms and common symptoms experienced across chronic conditions. The review identified eight polymorphisms in seven genes that were associated with at least three symptoms or symptom clusters, providing a starting place for those contemplating the inclusion of genetic polymorphisms in their symptom research trajectories. Consistently collecting and analyzing genetic polymorphisms across studies can provide further insight into the genomic contributions to variability in patients' experiences with common symptoms and symptom clusters across disease processes.

The review independently identified three genes, *BDNF*, *IL6*, and *TNFA*, for which the resulting protein has been previously identified as a CDE for symptom science research (Page et al., 2018). Interestingly, the functional polymorphisms identified for these genes impact either the secretion of the protein or the level of transcription of the gene (Bull et al., 2009; Chen et al., 2004; Kroeger et al., 1997; Paul-Samojedny et al., 2010; Wilson et al., 1997). Independently identifying these genes that have already been implicated as CDEs at the protein level increases the evidence in support of including the polymorphisms in these genes in research focused on symptoms.

The final list of polymorphisms is consistent with several publications underscoring the putative function of inflammatory and immune processes and neuroendocrine processes

Table 3. Description and Function of Polymorphisms.

Gene	Polymorphism (Alternative Name If Applicable)	Polymorphism Function
BDNF	rs6265 (val66met)	This missense variant results in an amino acid substitution (i.e., valine to methionine) at codon 66. This change impairs the secretion of BDNF in the nervous system (Chen et al., 2004; Egan et al., 2003).
COMT	rs4680 (val158met)	This missense variant results in an amino acid substitution (i.e., valine to methionine) at position 158 of *COMT*. This alters the structure of the COMT enzyme and reduces its activity, resulting in higher dopamine levels (Stein, Newman, Savitz, & Ramesar, 2006).
FKBP5	rs3800373	This variant is located in the 3 prime untranslated region and likely alters the stability and half-life of the mRNA and modulates glucocorticoid signaling and hypothalamic–pituitary–adrenal axis function. This variant has been associated with greater *FKBP5* induction by cortisol and decreased glucocorticoid-receptor sensitivity (Fudalej et al., 2015; Tatro et al., 2009).
IL-6	rs1800795 (-174 G/C)	This variant is located in the promoter region, which is a region essential for inducing transcription of *IL-6*. rs1800795 has been associated with differential gene expression and decreased plasma levels of IL-6 during immune activation (Bull et al., 2009; Paul-Samojedny et al., 2010).
NFKB2	rs1056890	This variant is located in the 3 prime untranslated region, which is a region known to bind miRNAs and regulate protein translation (Ma, Becker Buscaglia, Barker, & Li, 2011; Tian et al., 2018).
SLC6A4	5-HTTLPR + rs25531	The HTTLPR variant is a 43-base-pair variable-number tandem repeat (VNTR) polymorphism, and rs25531 is a single nucleotide polymorphism (SNP) located in the promoter region of *SLC6A4*. The VNTR determines if the alleles are long or short, and the SNP further divides the long allele into La and Lg. The La allele produces significantly more 5-HTT mRNA and protein and results in increased expression and serotonin transporters in the cell membrane. The short allele results in lower levels of serotonin (Wendland, Martin, Kruse, Lesch, & Murphy, 2006).
TNFA	rs1799964 (-1031 T/C)	This variant is located in the promoter region and influences gene expression. The C allele is correlated with increased serum TNFA levels (Nourian et al., 2017; Sandoval-Pinto et al., 2016).
	rs1800629 (-308 G/A)	This variant is located in the promoter region. The A allele has been associated with an increase in the binding of nuclear factors and heightened transcription of the gene (Kroeger, Carville, & Abraham, 1997; Wilson, Symons, McDowell, McDevitt, & Duff, 1997).

Note. miRNAs = microRNAs; mRNA = messenger RNA.

(e.g., regulation of the hypothalamic–pituitary–adrenal axis) in the experiences of symptoms or symptom clusters (McCall et al., 2018; Miaskowski & Aouizerat, 2012; Miaskowski et al., 2017; Page et al., 2018). Of the polymorphisms identified in this review, five are in genes that play an important role in inflammation and immune regulation. For example, *IL-6* encodes a protein that is involved in the regulation of inflammation and maturation of the lymphocytes including T cells and B cells. Likewise, *NFKB2* is integral in central activation of the inflammatory system through transcription activation and repression of several genes. *TNFA* encodes a pro-inflammatory cytokine that is involved in cell proliferation, differentiation, apoptosis, lipid metabolism, and coagulation. *FKBP5* is integral in immunoregulation (National Library of Medicine, 2018; Online Mendelian Inheritance in Man OMIM®, 2018).

The remaining polymorphisms are in genes involved in neural development and/or neurotransmission. BDNF is a nerve growth factor involved in neuroplasticity and regulation of synapse transmission. *COMT* encodes for an enzyme that is responsible for the metabolism of several neurotransmitters such as epinephrine, norepinephrine, and dopamine. Lastly, *SLC6A4* regulates serotonergic signaling and transport in the central nervous system (National Library of Medicine, 2018; Online Mendelian Inheritance in Man OMIM®, 2018). These findings further support the shared biological underpinnings of symptoms and symptom clusters.

While there are many types of biomarkers (e.g., proteins, epigenetic markers) that may be considered useful for symptom science research, genetic polymorphisms have several strengths. They are stable over time and are not impacted by factors such as age, sex, comorbidities, medication regimens, or other clinical interventions. Additionally, it does not matter what tissue type (e.g., whole blood, serum, saliva) is used for DNA extraction, and current technologies allow for the measurement of these polymorphisms with precision, sensitivity, and specificity quickly enough to have clinical utility. In contrast, other biological markers such as proteins (e.g., IL-6, TNF-α, CRP) are dynamic and can fluctuate based on a number of clinical factors, gene expression, or tissue type (Gry, Oksvold, Ponten, & Uhlen, 2010). Therefore, the use of genetic polymorphisms as biomarkers in symptom research increases the potential for their future clinical utility across populations, ages, and conditions.

While not all studies currently have relevant research questions or resources (e.g., infrastructure, finances) for genetic data collection and analyses, we encourage researchers to store biological samples. The investment involved in recruitment and phenotyping for symptom-based studies warrants securing the possibility to address genetic polymorphisms in the future. Stored samples could be used to expand current research or provide for replication of findings from other studies. Additionally, more symptom science research that includes genetic

polymorphisms would allow for the mega-analyses and meta-analyses essential to the evidence base required for translation to clinical utility.

The results of the present review should be interpreted with consideration of some limitations. It was not designed to identify all polymorphisms that could explain the complex phenotype of symptoms. Our search was limited to functional polymorphisms, and findings could be biased toward polymorphisms of candidate genes chosen in symptom-related studies. The final list of polymorphisms represents only the current literature at the time of the search. As more evidence emerges, additional polymorphisms will likely be identified and more support for the polymorphisms identified in the present review will likely come to light. Although we conducted this search in a systematic fashion, our literature search was limited to PubMed and was not exhaustive. It is possible that we excluded relevant studies outside of our search criteria. There was notable variability in phenotyping of symptoms in the included studies, which could have impacted the findings of associations between the genetic polymorphisms and symptoms of interest.

Conclusion

In the present review, we identified specific and stable functional genetic polymorphisms we recommend for inclusion in symptom science research. Using genetic polymorphisms as biomarkers has the potential to provide greater understanding of the biological basis of individual symptoms and symptom clusters and stable biomarkers related to symptom development. These genetic polymorphisms could also be used to identify individuals at risk for poor symptom experiences and target for precision health interventions.

Acknowledgment

The authors would like to thank Teresa Plummer for her assistance in the development of this article.

Author Contributions

All authors contributed to the conceptualization, data collection from the literature, interpretation of the literature, and writing of this manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study has been supported by National Institutes of Health (T32NR009759, K99NR015473).

ORCID iD

Lacey W. Heinsberg https://orcid.org/0000-0002-7690-5485
John D. Merriman https://orcid.org/0000-0002-7113-5389

Supplemental Material

Supplemental Material for this article is available online.

References

Albert, P. R. (2011). What is a functional genetic polymorphism? Defining classes of functionality. *Journal of Psychiatry & Neuroscience*, *36*, 363–365. doi:10.1503/jpn.110137

Bull, S. J., Huezo-Diaz, P., Binder, E. B., Cubells, J. F., Ranjith, G., Maddock, C., … Pariante, C. M. (2009). Functional polymorphisms in the interleukin-6 and serotonin transporter genes, and depression and fatigue induced by interferon-alpha and ribavirin treatment. *Molecular Psychiatry*, *14*, 1095–1104. doi:10.1038/mp.2008.48

Cashion, A. K., & Grady, P. A. (2015). The National Institutes of Health/National Institute of Nursing Research intramural research program and the development of the National Institutes of Health symptom science model. *Nursing Outlook*, *63*, 484–487. doi:10.1016/j.outlook.2015.03.001

Chen, Z. Y., Patel, P. D., Sant, G., Meng, C. X., Teng, K. K., Hempstead, B. L., & Lee, F. S. (2004). Variant brain-derived neurotrophic factor (BDNF) (Met66) alters the intracellular trafficking and activity-dependent secretion of wild-type BDNF in neurosecretory cells and cortical neurons. *Journal of Neuroscience*, *24*, 4401–4411. doi:10.1523/jneurosci.0348-04.2004

Corwin, E. J., Berg, J. A., Armstrong, T. S., DeVito Dabbs, A., Lee, K. A., Meek, P., & Redeker, N. (2014). Envisioning the future in symptom science. *Nursing Outlook*, *62*, 346–351. doi:10.1016/j.outlook.2014.06.006

Corwin, E. J., Moore, S. M., Plotsky, A., Heitkemper, M. M., Dorsey, S. G., Waldrop-Valverde, D., … Grady, P. A. (2017). Feasibility of combining common data elements across studies to test a hypothesis. *Journal of Nursing Scholarship*, *49*, 249–258. doi:10.1111/jnu.12287

Egan, M. F., Kojima, M., Callicott, J. H., Goldberg, T. E., Kolachana, B. S., Bertolino, A., … Weinberger, D. R. (2003). The BDNF val66met polymorphism affects activity-dependent secretion of BDNF and human memory and hippocampal function. *Cell*, *112*, 257–269.

Fudalej, S., Kopera, M., Wolynczyk-Gmaj, D., Fudalej, M., Krajewski, P., Wasilewska, K., … Ploski, R. (2015). Association between FKBP5 functional polymorphisms and completed suicide. *Neuropsychobiology*, *72*, 126–131. doi:10.1159/000441659

Gry, M., Oksvold, P., Ponten, F., & Uhlen, M. (2010). Tissue-specific protein expression in human cells, tissues and organs. *Journal of Proteomics & Bioinformatics*, *3*, 286–293. doi:10.4172/jpb.1000153

Kroeger, K. M., Carville, K. S., & Abraham, L. J. (1997). The -308 tumor necrosis factor-alpha promoter polymorphism effects transcription. *Molecular Immunology*, *34*, 391–399.

Ma, X., Becker Buscaglia, L. E., Barker, J. R., & Li, Y. (2011). MicroRNAs in NF-kappa B signaling. *Journal of Molecular Cell Biology*, *3*, 159–166. doi:10.1093/jmcb/mjr007

McCall, M. K., Stanfill, A. G., Skrovanek, E., Pforr, J. R., Wesmiller, S. W., & Conley, Y. P. (2018). Symptom science: Omics supports common biological underpinnings across symptoms. *Biological Research for Nursing*, *20*, 183–191. doi:10.1177/1099800417751069

Miaskowski, C., & Aouizerat, B. E. (2012). Biomarkers: Symptoms, survivorship, and quality of life. *Seminars in Oncology Nursing*, *28*, 129–138. doi:10.1016/j.soncn.2012.03.008

Miaskowski, C., Barsevick, A., Berger, A., Casagrande, R., Grady, P., Jacobsen, P., … Matocha, M. (2017). Advancing symptom science through symptom cluster research: Expert panel proceedings and recommendations. *Journal of the National Cancer Institute*, *109*. doi:10.1093/jnci/djw253

National Institute of Nursing Research. (2016). The NINR strategic plan: Advancing science, improving life. Retrieved from https://www.ninr.nih.gov/sites/www.ninr.nih.gov/files/NINR_StratPlan2016_reduced.pdf

National Library of Medicine. (2018). Genetics home reference. Retrieved from https://ghr.nlm.nih.gov/

Nourian, M., Chaleshi, V., Pishkar, L., Azimzadeh, P., Baradaran Ghavami, S., Balaii, H., … Zali, M. R. (2017). Evaluation of tumor necrosis factor (TNF)-alpha mRNA expression level and the rs1799964 polymorphism of the TNF-alpha gene in peripheral mononuclear cells of patients with inflammatory bowel diseases. *Biomedical Reports*, *6*, 698–702. doi:10.3892/br.2017.908

Online Mendelian Inheritance in Man OMIM®. (2018). Retrieved from https://www.omim.org/

Page, G. G., Corwin, E. J., Dorsey, S. G., Redeker, N. S., McCloskey, D. J., Austin, J. K., … Grady, P. (2018). Biomarkers as common data elements for symptom and self-management science. *Journal of Nursing Scholarship*, *50*, 273–286. doi:10.1111/jnu.12378

Paul-Samojedny, M., Kowalczyk, M., Suchanek, R., Owczarek, A., Fila-Danilow, A., Szczygiel, A., & Kowalski, J. (2010). Functional polymorphism in the interleukin-6 and interleukin-10 genes in patients with paranoid schizophrenia—A case-control study. *Journal of Molecular Neuroscience*, *42*, 112–119. doi:10.1007/s12031-010-9365-6

Redeker, N. S., Anderson, R., Bakken, S., Corwin, E., Docherty, S., Dorsey, S. G., … Grady, P. (2015). Advancing symptom science through use of common data elements. *Journal of Nursing Scholarship*, *47*, 379–388. doi:10.1111/jnu.12155

Sandoval-Pinto, E., Padilla-Gutierrez, J. R., Valdes-Alvarado, E., Garcia-Gonzalez, I. J., Valdez-Haro, A., Munoz-Valle, J. F., … Valle, Y. (2016). Association of the -1031T>C polymorphism and soluble TNF-alpha levels with acute coronary syndrome. *Cytokine*, *78*, 37–43. doi:10.1016/j.cyto.2015.11.014

Stein, D. J., Newman, T. K., Savitz, J., & Ramesar, R. (2006). Warriors versus worriers: The role of COMT gene variants. *CNS Spectrums*, *11*, 745–748.

Tatro, E. T., Everall, I. P., Masliah, E., Hult, B. J., Lucero, G., Chana, G., … Achim, C. L. (2009). Differential expression of immunophilins FKBP51 and FKBP52 in the frontal cortex of HIV-infected patients with major depressive disorder. *Journal of Neuroimmune Pharmacology*, *4*, 218–226. doi:10.1007/s11481-009-9146-6

Taylor, J. Y., & Barcelona de Mendoza, V. (2018). Improving -omics-based research and precision health in minority populations: Recommendations for nurse scientists. *Journal of Nursing Scholarship*, *50*, 11–19. doi:10.1111/jnu.12358

Tian, T., Wang, J., Huang, P., Li, J., Yu, R., Fan, H., … Yue, M. (2018). Genetic variations in NF-κB were associated with the susceptibility to hepatitis C virus infection among Chinese high-risk population. *Scientific Reports*, *8*, 104. doi:10.1038/s41598-017-18463-y

Wendland, J. R., Martin, B. J., Kruse, M. R., Lesch, K. P., & Murphy, D. L. (2006). Simultaneous genotyping of four functional loci of human SLC6A4, with a reappraisal of 5-HTTLPR and rs25531. *Molecular Psychiatry*, *11*, 224–226. doi:10.1038/sj.mp.4001789

Wilson, A. G., Symons, J. A., McDowell, T. L., McDevitt, H. O., & Duff, G. W. (1997). Effects of a polymorphism in the human tumor necrosis factor alpha promoter on transcriptional activation. *Proceedings of the National Academy of Science of the United States of America*, *94*, 3195–3199.

Appendix F.1 License Agreement for Manuscript 4

SAGE

Symptom Science: Advocating for Inclusion of Functional Genetic Polymorphisms

Author: Mitchell R. Knisely, Megan Maserati, Lacey W. Heinsberg, et al
Publication: Biological Research For Nursing
Publisher: SAGE Publications
Date: 07/01/2019

Dear Mrs. Lacey Heinsberg,

Your SAGE Publications request has been cancelled for the following reason:
Dear Lacey Heinsberg,

Thank you for your order. Under SAGE's journal author policy, you are permitted to include your own articles within your thesis in a publicly accessible platform, such as your institutional repository. You must provide a full citation and the permission does not include any third-party material. If you require formal written permission, please place an order via our Customer Permission Portal (https://jira.sagepub.com/servicedesk/customer/portal/9).

As we have replied via this comment, we will cancel this order. If we can be of further assistance, please contact us via our customer portal.

Best regards,
SAGE Rights Team.

You will not be charged for this order.

Appendix G Manuscript 5: Omics for Nurse Scientists Conducting Environmental Health Research

CHAPTER 3

Omics for Nurse Scientists Conducting Environmental Health Research

Lacey W. Heinsberg and Yvette P. Conley

ABSTRACT

Nurse scientists are ideally positioned to perform environmental health research and it is critical that the role of omics in the complex relationships between environmental exposures and an individual's unique physiology in human health outcomes be appreciated. Importantly, omics can offer nurse scientists a tool to measure exposure, demonstrate molecular phenotypic changes associated with exposure, and potentially uncover mechanisms of exposure-related disease or negative health outcomes. The purpose of this summary is to serve as an overview of omics methodologies for nurse scientists conducting environmental health research and provides future directions of this work as well as exemplar funding opportunities that demonstrate the growing need and interest in this area. The intersection of nursing and exposure science will accelerate the work in environmental health and bring forth translation of research findings into clinical and community practice. Importantly, this information can better help us understand the variation in response to the environment and support environmental health policy change at the local, state, and federal level to improve community health and well-being.

http://dx.doi.org/10.1891/0739-6686.38.35

INTRODUCTION

Environmental exposures have been shown to adversely impact health throughout the life span and nurses have witnessed exposure-related, poor health outcomes firsthand (Landrigan et al., 2016; Polivka & Chaudry, 2018). Public health is at a greater risk than ever before due to environmental exposures such as climate change, chemicals, processed foods, and more (U.S. Global Change Research Program [USGCRP], 2016). In many areas, the acceleration of these changes has made it difficult to maintain relevant federal regulations to protect community health and well-being (Maffini, Neltner, & Vogel, 2017). Because of this, there is a critical need for rigorous environmental health research and nurse scientists are ideally positioned to perform this work because they (a) are trusted healthcare professionals fully invested in public health, (b) are able to understand and translate scientific literature to the public, and (c) are interested in a range of phenotypes including the symptom experience, mental and physical well-being, health across the life span including end-of-life, and patient outcomes from injury (versus only "disease" states). All of these qualities make nurse scientists excellent knowledge brokers capable of bringing about delivery of health promotion and wellness as well as environmental, social, and economic return to communities (Thompson & Schwartz Barcott, 2019). This blended value has the potential to transform exposure and nursing science, clinical practice, and healthcare, especially during this time of declining planetary and human health.

In the era of the human genome project, exposure science received less attention than warranted in a range of disciplines. A promise of the human genome project was the transformation of healthcare through an understanding of the underlying causes of disease pathology and acceleration of precision health. While our knowledge of the mechanisms of many diseases has been greatly extended due to the project, genetics has been found to account for a much smaller percentage of disease susceptibility than originally thought. Notably, the remaining contributions appear to be environmental exposures and the subsequent interactions with an individual's unique genetics and physiology (Rappaport, 2016). Thus, to identify modifiable exposures and comprehensively understand the associated mechanisms of disease, there is a critical need for environmental health research using omics methodologies. This type of work has the potential to accelerate environmental health research and achieve broader health goals for the general public, including vulnerable populations who might not receive direct and immediate benefit from precision health (Senier, Brown, Shostak, & Hanna, 2017).

"Omics" refers to the range of methods used to investigate the functions and relationships of the molecular makeup of organisms (e.g., genomics, epigenomics, transcriptomics; Horgan & Kenny, 2011). An excellent resource for foundational information related to omics for the nurse scientist can be found in

the primer series published in *Biological Research for Nursing* (Alexander, 2017; Dorman, Schmella, & Wesmiller, 2017). Importantly, omics can offer nurse scientists a tool to (a) provide surrogate measures for environmental exposures (i.e., biological monitoring of exposure), (b) provide a more comprehensive view of disease processes by uncovering the combined role of genetics and environmental exposures in human health outcomes, and (c) offer supporting molecular evidence to demonstrate the phenotypic effects of exposure and/or temporal pathology of disease or poor health outcomes (i.e., mechanism) as well as potential molecular impacts of proposed interventions (e.g., modification or removal of a specific environmental exposure). We believe this information could supplement and strengthen nursing research to support environmental health policy change at the local, state, and federal level to improve community health and well-being.

The theoretical foundation of this overview considers the complex relationships between environmental exposures (defined here as a range of factors such as climate, diet, chemicals, social dynamics, physical activity, human behavior, etc.), human genomics and molecular response (i.e., omics), and human health outcomes across the life span (Figure 3.1).

Figure 3.1 depicts the potential ways these complex relationships could be examined in research studies, even in cases where data do not exist from each category. The purpose of this summary is to serve as an outline of omics amenable to environmental health nursing research by providing a brief overview of omics methodologies with examples of their use from the environmental health research literature, as well as potential future directions and funding opportunities for nurse scientists.

OVERVIEW OF OMICS METHODOLOGY

Genomics

Overview

Genomics is the study of the order and function of the DNA sequence of an organism and, in general, includes either a targeted evaluation looking at a specific gene or genes as well as a genome-wide approach (Dorman et al., 2017). Unlike the other omics methodologies discussed in this chapter, the DNA sequence is unique in that it is stable across time and tissue type (e.g., blood, saliva), so genotyping generally only needs to occur once in a lifetime. The most common (and most commonly studied) type of genetic variation is the single nucleotide polymorphism (SNP), a variation in one base of the DNA sequence (Horgan & Kenny, 2011). Research has demonstrated many associations between SNPs and human health outcomes in a variety of populations and this measure of biological

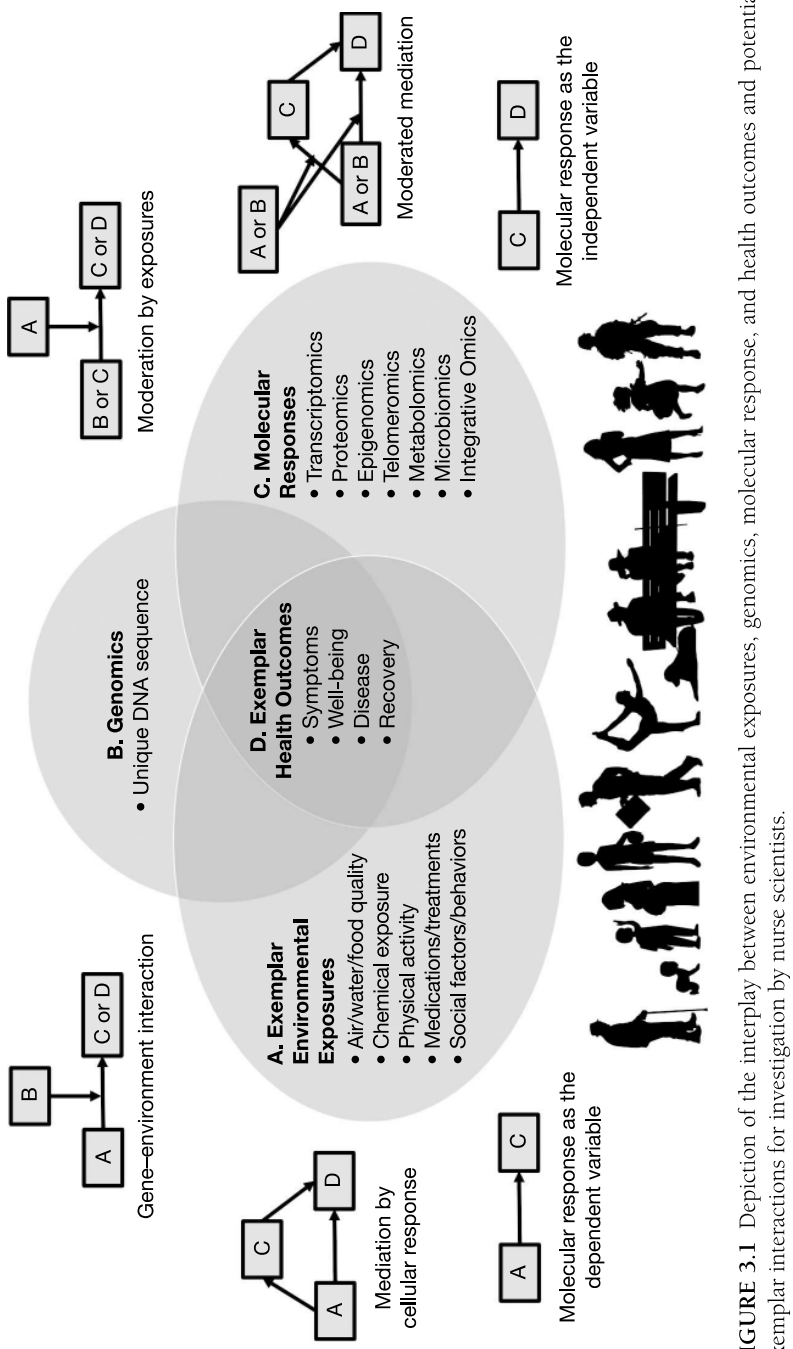


FIGURE 3.1 Depiction of the interplay between environmental exposures, genomics, molecular response, and health outcomes and potential exemplar interactions for investigation by nurse scientists.

variation between people offers merit as a tool in environmental health nursing research. In the following, we have highlighted an exemplar research question from the environmental health literature to demonstrate the potential of genomics in this important area of research.

An Exemplar From the Literature: Gene–Environment Interaction—Body Mass Index and Hypertension

Individuals may have different susceptibilities to human health outcomes depending on their age, sex, unique environmental exposures, genetics, or combination of these factors. Genetic variation can be used in environmental health nursing research to study the gene–environment (GxE) interaction. Here we have highlighted an exemplar from the environmental health nursing literature of a study by Taylor, Sun, Hunt, and Kardia (2010). This study examined the relationship between the interaction of genetic variation and a potential surrogate measure of a modifiable environmental exposure, body mass index (BMI), in the development of hypertension in two generations of African American women (parent cohort, $n = 868$; offspring cohort, $n = 322$). In this candidate gene association study, variation in one SNP from the *CAPN13* gene had a gene–BMI interaction on hypertension in the offspring cohort that remained significant after adjustment for multiple testing (data not available in the parent cohort). In this study, hypertension in the offspring cohort depended on both genotype and BMI. The conclusions of this study were that early life genetic screening could alert nurses to those in need of modifiable environmental exposures (e.g., altered diet, physical activity) to prevent or delay hypertension and improve outcomes in this population (Taylor et al., 2010). This work could be extended to an array of research questions in environmental health nursing research.

Transcriptomics

Overview

The genomic code discussed earlier is the basis for other omic mechanisms and makes up genes that help control functions within the human body. The first step in expression of our genes is the transcription of DNA into RNA (Dorman et al., 2017). Transcriptomics is the study of the expression of RNA in cells and tissues (Horgan & Kenny, 2011). Like genomic studies discussed earlier, transcriptomics can be studied using a targeted candidate gene approach or a genome-wide approach. However, transcriptome biomarker levels vary based on the tissue type and are not stable over time, setting it apart from the stable DNA sequence–based genomics mentioned earlier. Because the transcriptome can vary across time and tissues, great attention must be offered to tissue selection and timing of sample collection to best assess the phenotype being investigated. Serial sampling and using a within participant change over time approach can be considered.

Additionally, given the instability of RNA, collection and storage of specimens for assessment of RNA needs to be considered to reserve the snapshot representing the time of specimen collection. Transcriptome biomarkers of RNA expression have been shown to be associated with various cancers and chronic inflammation and can provide important evidence of health and molecular consequences of interaction with environmental exposures. In the following, we provide an exemplar from the environmental health research of the utility of transcriptomics.

An Exemplar From the Environmental Health Literature: Polychlorinated Biphenyls and Transcriptome Response

During their manufacturing and use, human-made chemicals (e.g., pesticides, consumer cosmetic ingredients, drug metabolites) are released into the environment and often bioaccumulate within people and wildlife. One such class of chemicals is polychlorinated biphenyls (PCBs), which were used in electrical equipment and mechanical fluids prior to 1977. Although these chemicals were banned by the U.S. Environmental Protection Agency, they remain present in the environment today and can enter and accumulate in the body through inhalation, consuming contaminated food, or through skin contact (Grimm et al., 2015). In a population-based case control study of lymphoma by Espín-Pérez et al. (2019), the relationship between PCB exposure and genome-wide gene expression profiles was evaluated (n = 649). Notably, this study found that exposure to PCBs was associated with sex- and white blood cell type specific gene expression changes. In male participants, gene expression of cancer-related pathways was associated with PCB exposure while in female participants, gene expression of immune system-related pathways was associated with PCB exposure. This study supports previous research in which males demonstrated a greater incidence of cancer compared with females in association with exposure to PCB (Espín-Pérez et al., 2019). Not only can a study using transcriptomics offer supporting evidence of molecular changes associated with human health outcomes but can also offer preliminary insights into the mechanisms of exposure-related phenotypes, further strengthening environmental health nursing research.

Proteomics

Overview

The second step in gene expression is the translation of RNA to the proteins that carry out functions throughout the body (Dorman et al., 2017). Proteomics is the study of all the proteins present in a specific cell or tissue type (Horgan & Kenny, 2011). The goal of this type of work is to identify and quantify the types of proteins in a sample and investigate the relationships with environmental exposures and health outcomes. Investigating proteins requires methodological considerations similar to RNA work given that protein expression can differ by

tissue, differ over time, and stability can vary across different proteins. Although proteins are based on RNA transcripts discussed earlier, post-transcriptional or post-translational modifications (discussed later) can occur causing discordance between RNA and protein levels. Studying both transcriptomics and proteomics can be valuable tools in untangling the effects of environmental exposures in nursing research.

An Exemplar From the Environmental Health Literature: Air Pollution and Proteomic Response

Evidence suggests that air pollution is associated with a range of health issues including poor cardiorespiratory health, neurodevelopment defects, and more (Payne-Sturges et al., 2019). An exemplar study from the environmental health literature using proteomics to investigate the impact of air pollution on human health outcomes was conducted by Kumarathasan et al. (2018). The goal of this study was to identify the association between air pollution near a steel plant and human proteomic biomarkers. In brief, in a randomized cross-over study of healthy volunteers (n = 52), participants spent 5 days near a steel plant ("Bayview Site," approximately one half a mile away from continuously operating coke ovens) or on a college campus ("College Site," approximately three miles away from continuously operating coke ovens) or were fitted with a personal air filtration mask only at the Bayview Site ("Bayview-Mask"). A 9-day washout period was used between site rotations. For each portion of the study, participants were tested for proteomic markers of oxidative stress, inflammation, and endothelial dysfunction. The researchers also collected a series of air pollution measurements and health outcomes. The findings of this study demonstrated that air pollution was higher near the steel plant and that participants at the Bayview Site had significantly increased pro-inflammatory cytokines and endothelins compared with the College Site and the Bayview-Mask. Health consequences associated with these changes included increased heart rate, blood pressure, and C-reactive protein (Kumarathasan et al., 2018). The findings of this novel, natural experiment suggested that air pollution near the steel plant influenced vascular and inflammatory mechanisms in humans. This type of study design using omics methodology could strengthen environmental health nursing research to help nurses advocate for stricter air quality regulations in vulnerable populations.

Epigenomics

Overview

Epigenomics is the study of modifications to the genome that impact gene regulation (i.e., turn genes on and off) without impacting the sequence of the DNA (Fessele & Wright, 2018). These processes are responsible for normal growth and development, recovery from injury, and have been shown to be dysregulated

in disease states (Jin & Liu, 2018) making them excellent biological tools and biomarkers for environmental health nursing research. Of note, epigenetic mechanisms can be responsible for the disconnect in the correlation between RNA and protein levels discussed earlier (i.e., the cause of post-transcriptional and/or post-translational modifications). Like genomic and transcriptomic studies discussed earlier, epigenomics can be studied using a targeted candidate gene approach or a genome-wide approach. Likewise, similar to RNA and protein biomarkers, epigenomic signatures can differ between tissues and can be dynamic, therefore, similar attention should be given to selecting the most appropriate tissue and time points to represent the phenotype under investigation.

The most commonly studied epigenetic modifications include DNA methylation, histone modification, and RNA interference (i.e., microRNAs; Fessele & Wright, 2018). DNA methylation is the most well-characterized epigenomic mechanism and involves the addition of a methyl group (clusters of hydrocarbons) to the DNA molecule. When located in the promoter region of a gene, DNA methylation typically acts to repress gene transcription (i.e., turns the gene off). Next, histone modification is a modification to histone proteins that takes place following translation and includes acetylation, methylation, phosphorylation, ubiquitylation, sumoylation, adenosine diphosphate (ADP)-ribosylation, deamination, and proline isomerization (the most well-studied being acetylation and methylation) and impacts gene expression by altering the chromatin structure (Fessele & Wright, 2018; Jin & Liu, 2018). Finally, microRNAs affect gene expression at the post-transcriptional level by impeding translation of proteins or destabilizing transcripts (Fessele & Wright, 2018; Vasilatou, Papageorgiou, Dimitriadis, & Pappa, 2013). These types of epigenomic measurements can be used as a tool in environmental health nursing research as potential biomarkers of environmental exposures and disease processes. In the following, we provide an exemplar study from the environmental health nursing research using this biological tool.

An Exemplar From the Literature: Social Epigenomics and Stress

Social epigenomics is an emerging area of research in chronic disease development and progression. Individuals living in vulnerable and disadvantaged neighborhoods are often exposed to air and noise pollution, food deserts, and psychosocial stress and have associated differential DNA methylation (Cunliffe, 2016). Here, we have highlighted an exemplar research study from the environmental health nursing literature by Wright et al. (2017). In this cross-sectional, observational study, researchers investigated the association between maternal parenting stress and epigenome-wide DNA methylation of saliva samples among African American mother–child pairs (n = 74 dyads). This study found that variation in DNA methylation of *PARP-1*, a gene associated with stress response

regulation, was associated with reported maternal parenting stress. The DNA methylation patterns in children reflected those of their mothers. Although none of the methylation sites associated with maternal stress were found to be significant in children after correction for multiple testing, four sites were marginally differentially methylated among children (Wright et al., 2017). The results of this study suggest that the social environment impacts the molecular response and may support nursing research by providing important evidence in similar studies.

Telomeromics

Overview

Telomeromics is the study of telomeres, which are the "caps" at the ends of DNA strands that offer protection to chromosomes. Telomeres are targets for epigenetic changes discussed earlier, but also have epigenetic regulatory elements that impact telomerase and cellular aging. Over time, telomeres shorten during cellular replication and are eventually unable to function properly and cause dysfunction in cells. Because of this, telomere length is strongly associated with age and has been shown to be prematurely shortened by environmental exposures (Venkatesan, Khaw, & Hande, 2017). Like RNA and protein biomarkers, telomeres can differ between tissues and can be dynamic, therefore, similar attention should be given to selecting the most appropriate tissue and time points to represent the phenotype under investigation. In the following, we provide an exemplar from the environmental health nursing literature demonstrating the use of telomere length as a biological tool.

An Exemplar From the Environmental Health Literature: Neighborhood Quality and Telomere Length

Living in disadvantaged neighborhoods where air quality may be poor, or residents may be exposed to other types of physical and psychological stress, has been shown to have significant negative health impacts including asthma, cardiorespiratory symptoms, and cancers (Chitewere, Shim, Barker, & Yen, 2017). In a study by Park, Verhoeven, Cuijpers, Reynolds, and Penninx (2015), researchers investigated the relationship between perceived neighborhood quality and molecular health as measured by telomere length (n = 2,902). This study established that, compared to those who perceived neighborhood quality to be good, the average telomere length in perceived "moderate" and "poor" quality neighborhoods were 69 and 174 base pairs shorter, respectively (which translates to 8.7 and 11.9 years older in chronological age; Park et al., 2015). The role of telomere length in environmental health nursing research is promising in that it offers a biological measurement of premature molecular aging, which could lead to pathology associated with a range of diseases or poor health outcomes.

Metabolomics

Overview

Metabolomics is the global measure of all the products of metabolism within the body and include small molecules (i.e., metabolites) such as inorganics, peptides, lipids, carbohydrates, and more (Li, Dunlop, Jones, & Corwin, 2016). Given the recent expansion of bioinformatics technologies, metabolomics does an excellent job of reflecting the interplay between genomics and environmental exposures and the subsequent impact on cellular function within the human body and should be considered as a potential biological tool in nursing research. A comprehensive review of the metabolomics field, methodological considerations, and implications for nursing science including information surrounding the metabolomics of preterm birth as a paradigm for exposure research can be found elsewhere (Li et al., 2016). In the following, we provide an additional exemplar of the use of metabolomics in environmental health research.

An Exemplar From the Environmental Health Literature: Arsenic Metabolism

Despite research demonstrating the negative health impacts of heavy metals, exposures via occupation or drinking water remains prevalent throughout the world today. In excess, heavy metals can cause oxidative damage to tissues and negative health impacts such as kidney injury, diabetes, and some cancers (Rehman, Fatima, Waheed, & Akash, 2018). Ample environmental health research has demonstrated variation in the metabolism of one such heavy metal—arsenic. For example, in a candidate gene study of 149 male copper mill workers and 52 healthy controls from southwestern Poland conducted by Janasik et al. (2015), researchers measured arsenic metabolites in urine and compared levels of the mill workers and healthy controls. The team found that genetic variation in two genes (*AS3MT* and *GSTO2*) were associated with higher arsenic concentration and species of arsenic in urine, even after controlling for inhaled arsenic concentrations (Janasik et al., 2015). In addition, in a sample of research participants from Bangladesh ($n > 4{,}800$), researchers identified a relationship between a genetic variant in a novel gene, *FTCD*, and increased urinary arsenic metabolites. Moreover, participants with the low-efficiency metabolism allele had an increased arsenic-induces skin lesion risk. The researchers found that together *FTCD* and *AS3MT* risk SNPs explain about 10% of variation in arsenic metabolism in this sample (Pierce et al., 2019a, b). Using metabolomics as a biological monitoring tool may further strengthen environmental health nursing research in the future by offering evidence of molecular changes and contributing to the robustness of findings.

Microbiomics

Overview

The microbiome includes all of the microorganisms in a given environment (e.g., gut, skin, nasal, vaginal), which has an influence on the host during wellness homeostasis and disease (Thursby & Juge, 2017). A substantial amount of research has demonstrated its associations with human health. Not only is there evidence that many exposures can cause perturbation of the microbiome, including pharmacologic agents, age, genetics, sex, food, metals, and pesticides, but that the microbiome plays a complex role in the body's metabolism of environmental chemicals (Claus, Guillou, & Ellero-Simatos, 2016). The microbiome offers great potential as a biological measure of the human response to environmental exposures as well as an important mediator of the relationship between environmental exposures and human health outcomes. In the following, we provide an exemplar from the environmental health literature highlighting the potential of the microbiome as a tool in environmental health nursing research. A guide to key methodological factors for nurse scientists designing a microbiome study can be found elsewhere (Maki et al., 2019).

An Exemplar From the Environmental Health Literature: Pesticide Exposure, Microbiome Response, and Outcomes

The expansion and industrialization of agriculture has increased the chemical burden on planetary and human health. Negative health consequences from pesticide exposure range from mild symptoms of headaches to a possible role in the etiology of Parkinson's disease (Nicolopoulou-Stamati, Maipas, Kotampasi, Stamatis, & Hens, 2016). Although little research exists investigating the impacts of pesticide exposure on the microbiome (or the potential mediating role of the microbiome) in human health outcomes, there has been some important work done in animal models. In a study by Aitbali et al. (2018), the authors investigated the relationship between exposure to glyphosate-based herbicides (a pesticide designed to kill invasive plant pests) and the gut microbiome of mice and subsequent effects on behavioral functions following controlled exposure. As part of this study, the researchers assigned mice to three groups including an acute exposure control group and subchronic and chronic exposure groups. After exposure to the herbicide, mice underwent longitudinal behavioral testing and the gut microbiome was investigated. In this study, the exposure groups demonstrated behavior alterations in mice while the control group showed no difference in behaviors from baseline. Moreover, in the subchronic and chronic exposure groups, there was significant alteration in the abundance and diversity of key gut microbes (Aitbali et al., 2018). This study supports the link between herbicide

exposure and poor health outcomes using the microbiome as ancillary evidence and merits further investigation in humans.

Beyond this example, researchers have recently identified differences in the vaginal microbiome that were significantly associated with preterm birth, particularly in women of African ancestry (Fettweis et al., 2019). While the authors did not investigate the potentially modifiable exposures that may have contributed to the differences in the microbiome among these women, this study highlights a significant opportunity for future directions and similar work by nurse scientists to fill in this environmental health knowledge gap and bring about clinical change in this important area.

Integrative Omics (Multi-Omics)

Overview

Integrative omics (i.e., multi-omics) includes the analysis of an array of data collected using the methodologies discussed earlier to investigate joint biological effects in organisms (e.g., investigation of genomics, transcriptomics, and proteomics simultaneously; Karczewski & Snyder, 2018). Using integrative omics, there is a higher likelihood of being able to demonstrate causality in exposure science and uncover the mechanism of disease compared with a single-omics approach. A range of omics could be integrated depending on the research question of interest. In the following, we provide an exemplar of the use of integrative omics from the environmental health research literature.

An Exemplar From the Environmental Health Literature: Phthalate Exposure and Alteration of the Placental Methylome and Transcriptome

Endocrine disruptors are chemicals that interfere with the human hormone system and are associated with cancer, kidney damage, birth defects, and more (Kabir, Rahman, & Rahman, 2015). Phthalates are known endocrine disruptors and are found in hundreds of consumer products including plastics, cleaning products, and personal care products. The mechanism of prenatal exposure to phthalates on molecular health is largely unknown. In a study by Grindler et al. (2018), the authors examined the relationship of maternal phthalate exposure on the DNA methylome and transcriptome of the human placenta in the first trimester (n = 49). Maternal urine samples were collected in the first trimester, urinary phthalate was measured, and placental villous tissue was extracted for collection of genome-wide gene expression and DNA methylation data. The group used an integrative analysis and identified 39 genes with significantly altered DNA methylation and gene expression in the group with high phthalate exposure (Grindler et al., 2018). Although this study did not connect phthalate exposure with long-term health effects, its use of integrative omics as a measure

of the molecular changes of exposure is novel and could be an excellent tool for environmental health nursing research.

FUTURE DIRECTIONS FOR NURSE SCIENTISTS

In this chapter we have provided an overview of omics methodologies amenable to environmental health nursing research along with some exemplars from the literature. Given their well-rounded view of the whole person, nurse scientists are well positioned to perform this type of work and can bring a unique perspective by examining the role of environmental exposures in health outcomes through a nursing lens. Moreover, the National Institute of Nursing Research (NINR) advocates for the inclusion of omics in nursing research as demonstrated in the Strategic Plan (NINR, 2016). The methods discussed earlier could be extended beyond the current literature to novel areas of research that are in line with the mission and strategic plan of the NINR (e.g., symptom science, health promotion and development, self-management of chronic conditions, end-of-life and palliative care). In the following, we have provided some contemporary research questions that nurse scientists could answer using both omics and environmental exposure research methods.

- What early life environmental exposures impact wellness across the life span? Are there genetic moderators (i.e., GxE interactions) that are important in these relationships?
- Are there modifiable environmental contributors to molecular responses (i.e., omics) associated with symptom occurrence, type, severity, duration, and response to treatment in specific phenotypes? Are there molecular responses that could indicate the importance of a modifiable environmental exposure associated with a change in symptom trajectories in specific phenotypes?
- What are the relationships between environmental exposures and symptom clusters in specific phenotypes? What pathways are differentiated in symptom clusters in specific phenotypes? Can environmental exposures associated with these mechanistic changes be identified and measured?
- How do environmental exposures impact the ability to self-manage symptoms in chronic disease? What modifiable environmental exposures can be changed to improve the symptom experience in the self-management of chronic disease, and can any supporting evidence of molecular response to intervention be identified (i.e., removal or modification of an environmental exposure)?
- What is the relationship between environmental exposures and associated biological factors in the trajectory of disease or recovery from disease or injury?

- How do environmental exposures impact the symptom burden in the end of life? Are there environmental exposures that are associated with molecular changes and symptom burden that can be modified to improve end-of-life and palliative care?

Notably, there are several funding opportunities to support nurse scientists conducting environmental health research using omics (Table 3.1).

In Table 3.1, we have highlighted some examples of these funding opportunities that demonstrate the need and growing interest in environmental health research. Nurse scientists could capitalize on these announcements and answer novel research questions in environmental health research.

TABLE 3.1
Exemplar Funding Opportunities to Support Nurse Scientists Conducting Environmental Health Research Using Omics

Funding Opportunity Description	Funding Opportunity Number	Funding Agency
Telomeres in Wellness and Disease: A Biobehavioral Approach	PA-19-074	NINR
Environmental Exposures and Health: Exploration of Nontraditional Settings	PA-18-142	NINR; NIEHS
Research to Action: Assessing and Addressing Community Exposures to Environmental Contaminants	PA-18-260	NINR; NIEHS
Addressing Chronic Wound Trajectories Through Social Genomics Research	PA-17-492	NINR
Environmental Influences on Aging: Effects of Extreme Weather and Disaster Events on Aging Processes	PAR-19-249	NIA; NIEHS; OBSSR
Transition to Independent Environmental Health Research Career Award	PAR-19-225	NIEHS

(*Continued*)

TABLE 3.1
Exemplar Funding Opportunities to Support Nurse Scientists Conducting Environmental Health Research Using Omics (Continued)

Funding Opportunity Description	Funding Opportunity Number	Funding Agency
Advancing Translational and Clinical Probiotic/Prebiotic and Human Microbiome Research	PA-18-902	NCI; NCCIH; ODS
Outstanding New Environmental Scientist Award	RFA-ES-18-001	NIEHS
Revolutionizing Innovative, Visionary Environmental Health Research	RFA-ES-19-008	NIEHS
Oscillatory Patterns of Gene Expression in Aging and Alzheimer's Disease	RFA-AG-20-040	NIA
Mechanism for Time-Sensitive Opportunity in Environmental Health Sciences	RFA-ES-16-005	NIEHS
Functional Genetics, Epigenetics, and Noncoding RNAs in Substance Use Disorders	PA-17-155	NIDA
Social Epigenomics Research Focused on Minority Health and Health Disparities	PAR-16-355	NIMHD; NCI; NIA
Humans, Disasters, and the Built Environment	PD-19-1638	NSF

NIH = National Institutes of Health; NINR = National Institute of Nursing Research of the NIH; NIEHS = National Institute of Environmental Health Sciences of the NIH; NIA = National Institute of Aging of the NIH; OBSSR = Office of Behavioral and Social Sciences Research; NCI = National Cancer Institute of the NIH; NCCIH = National Center for Complementary and Integrative Health; ODS = Office of Dietary Supplements; NIDA = National Institute on Drug Abuse; NIMHD = National Institute on Minority Health and Health Disparities; NSF = National Science Foundation.

CONCLUSION

Nurse scientists are well positioned to perform environmental health research and the use of omics methodologies may further strengthen this work. In this summary, we have discussed ways in which omics can support environmental health nursing research as a biological tool to measure exposure, demonstrate molecular changes associated with exposure, and potentially uncover mechanisms of exposure-related disease. Based on the examples discussed earlier, we conclude that neither omics nor the environment can be ignored in rigorous environmental health research. Given the increasing number of funding opportunities and affordability of this type of research, in conjunction with advances in validated technologies, bioinformatics, and statistical approaches, nurse scientists can use these biological tools to uncover an integrated view of environmental exposures, omics, and human health. This intersection of nursing and exposure science will accelerate the work in this field and bring forth translation of research findings into clinical and community practice. Understanding these complex phenomena offers the potential to yield immeasurable contributions to healthcare and community wellness through controlling adverse environmental exposures, developing predictive models of symptoms and disease, improving management of symptoms and disease by mitigating the effect of environmental exposures, and ultimately reducing the symptom and disease burden around the globe.

ACKNOWLEDGMENTS

National Institute of Nursing Research F31NR017311.

REFERENCES

Aitbali, Y., Ba-M'hamed, S., Elhidar, N., Nafis, A., Soraa, N., & Bennis, M. (2018). Glyphosate based-herbicide exposure affects gut microbiota, anxiety and depression-like behaviors in mice. *Neurotoxicology and Teratology*, *67*, 44–49. doi:10.1016/j.ntt.2018.04.002

Alexander, S. A. (2017). Announcing the primer in genetics and genomics series. *Biological Research For Nursing*, *19*(1), 5–6. doi:10.1177/1099800416677094

Chitewere, T., Shim, J. K., Barker, J. C., & Yen, I. H. (2017). How neighborhoods influence health: Lessons to be learned from the application of political ecology. *Health and Place*, *45*, 117–123. doi:10.1016/j.healthplace.2017.03.009

Claus, S. P., Guillou, H., & Ellero-Simatos, S. (2016). The gut microbiota: A major player in the toxicity of environmental pollutants? *NPJ Biofilms and Microbiomes*, *2*, 16003. Nature Publishing Group. doi:10.1038/npjbiofilms.2016.3

Cunliffe, V. T. (2016). The epigenetic impacts of social stress: How does social adversity become biologically embedded? *Epigenomics*, *8*, 1653–1669. Future Science Group. https://doi.org/10.2217/epi-2016-0075

Dorman, J. S., Schmella, M. J., & Wesmiller, S. W. (2017). Primer in genetics and genomics, article 1: DNA, genes, and chromosomes. *Biological Research for Nursing*, *19*(1), 7–17. doi:10.1177/1099800416678321

Espín-Pérez, A., Hebels, D. G. A. J., Kiviranta, H., Rantakokko, P., Georgiadis, P., Botsivali, M., & de Kok, T. M. C. M. (2019). Identification of sex-specific transcriptome responses to polychlorinated biphenyls (PCBs. *Scientific Reports*, *9*(1), 746. doi:10.1038/s41598-018-37449-y

Fessele, K. L., & Wright, F. (2018). Primer in genetics and genomics, Article 6: Basics of epigenetic control. *Biological Research for Nursing*, *20*(1), 103–110. doi:10.1177/1099800417742967

Fettweis, J. M., Serrano, M. G., Brooks, J. P., Edwards, D. J., Girerd, P. H., Parikh, H. I., & Buck, G. A. (2019). The vaginal microbiome and preterm birth. *Nature Medicine*, *25*(6), 1012–1021. doi:10.1038/s41591-019-0450-2

Grimm, F. A., Hu, D., Kania-Korwel, I., Lehmler, H. J., Ludewig, G., Hornbuckle, K. C., & Robertson, L. W. (2015). Metabolism and metabolites of polychlorinated biphenyls. *Critical Reviews in Toxicology*, *45*, 245–272. NIH Public Access. doi:10.3109/10408444.2014.999365

Grindler, N. M., Vanderlinden, L., Karthikraj, R., Kannan, K., Teal, S., Polotsky, A. J., & Jansson, T. (2018). Exposure to phthalate, an endocrine disrupting chemical, alters the first trimester placental methylome and transcriptome in women. *Scientific Reports*, *8*(1), 6086. doi:10.1038/s41598-018-24505-w

Horgan, R. P., & Kenny, L. C. (2011). "Omic" technologies: Genomics, transcriptomics, proteomics and metabolomics. *Obstetrician & Gynaecologist*, *13*(3), 189–195. doi:10.1576/toag.13.3.189.27672

Janasik, B., Reszka, E., Stanislawska, M., Wieczorek, E., Fendler, W., & Wasowicz, W. (2015). Biological monitoring and the influence of genetic polymorphism of As3MT and GSTs on distribution of urinary arsenic species in occupational exposure workers. *International Archives of Occupational and Environmental Health*, *88*(6), 807–818. doi:10.1007/s00420-014-1009-7

Jin, Z., & Liu, Y. (2018). DNA methylation in human diseases. *Genes and Diseases*, *5*, 1–8. Chongqing Medical University. doi:10.1016/j.gendis.2018.01.002

Kabir, E. R., Rahman, M. S., & Rahman, I. (2015). A review on endocrine disruptors and their possible impacts on human health. *Environmental Toxicology and Pharmacology*, *40*, 241–258. doi:10.1016/j.etap.2015.06.009

Karczewski, K. J., & Snyder, M. P. (2018). Integrative omics for health and disease. *Nature Reviews Genetics*, *19*, 299–310. Nature Publishing Group. doi:10.1038/nrg.2018.4

Kumarathasan, P., Vincent, R., Blais, E., Bielecki, A., Guénette, J., Filiatreault, A., & Dales, R. (2018). Cardiovascular and inflammatory mechanisms in healthy humans exposed to air pollution in the vicinity of a steel mill. *Particle and Fibre Toxicology*, *15*(1), 34. doi:10.1186/s12989-018-0270-4

Landrigan, P. J., Sly, J. L., Ruchirawat, M., Silva, E. R., Huo, X., Diaz-Barriga, F., & Sly, P. D. (2016). Health consequences of environmental exposures: Changing global patterns of exposure and disease. *Annals of Global Health*, *82*, 10–19. No longer published by Elsevier. doi:10.1016/j.aogh.2016.01.005

Li, S., Dunlop, A. L., Jones, D. P., & Corwin, E. J. (2016). High-resolution metabolomics: Review of the field and implications for nursing science and the study of preterm birth. *Biological Research for Nursing*, *18*(1), 12–22. doi:10.1177/1099800415595463

Maffini, M. V., Neltner, T. G., & Vogel, S. (2017). We are what we eat: Regulatory gaps in the United States that put our health at risk. *PLOS Biology*, *15*(12), e2003578. doi:10.1371/journal.pbio.2003578

Maki, K. A., Diallo, A. F., Lockwood, M. B., Franks, A. T., Green, S. J., & Joseph, P. V. (2019). Considerations when designing a microbiome study: Implications for nursing science. *Biological Research for Nursing*, *21*(2), 125–141. doi:10.1177/1099800418811639

National Institute of Nursing Research. (2016). *The NINR strategic plan: Advancing science, improving lives*. Retrieved from https://www.ninr.nih.gov/sites/www.ninr.nih.gov/files/NINR_StratPlan2016_reduced.pdf

Nicolopoulou-Stamati, P., Maipas, S., Kotampasi, C., Stamatis, P., & Hens, L. (2016). Chemical pesticides and human health: The urgent need for a new concept in agriculture. *Frontiers in Public Health*, *4*, 148. doi:10.3389/fpubh.2016.00148

Park, M., Verhoeven, J. E., Cuijpers, P., Reynolds, C. F., & Penninx, B. W. J. H. (2015). Where you live may make you old: The association between perceived poor neighborhood quality and leukocyte telomere length. *PLOS ONE*, *10*(6), e0128460. doi:10.1371/journal.pone.0128460

Payne-Sturges, D. C., Marty, M. A., Perera, F., Miller, M. D., Swanson, M., Ellickson, K., & Hertz-Picciotto, I. (2019). Healthy air, healthy brains: Advancing air pollution policy to protect children's health. *American Journal of Public Health*, *109*(4), 550–554. doi:10.2105/AJPH.2018.304902

Pierce, B. L., Tong, L., Dean, S., Argos, M., Jasmine, F., Rakibuz-Zaman, M., & Ahsan, H. (2019a). A missense variant in FTCD is associated with arsenic metabolism and toxicity phenotypes in Bangladesh. *PLOS Genetics*, *15*(3), e1007984. doi:10.1371/journal.pgen.1007984

Pierce, B. L., Tong, L., Dean, S., Argos, M., Jasmine, F., Rakibuz-Zaman, M., & Ahsan, H. (2019b). Correction: A missense variant in FTCD is associated with arsenic metabolism and toxicity phenotypes in Bangladesh. *PLOS Genetics*, *15*(5), e1008172. doi:10.1371/journal.pgen.1008172

Polivka, B. J., & Chaudry, R. V. (2018). A scoping review of environmental health nursing research. *Public Health Nursing*, *35*(1), 10–17. doi:10.1111/phn.12373

Rappaport, S. M. (2016). Genetic factors are not the major causes of chronic diseases. *PLOS ONE*, *11*(4), e0154387. doi:10.1371/journal.pone.0154387

Rehman, K., Fatima, F., Waheed, I., & Akash, M. S. H. (2018). Prevalence of exposure of heavy metals and their impact on health consequences. *Journal of Cellular Biochemistry*, *119*(1), 157–184. doi:10.1002/jcb.26234

Senier, L., Brown, P., Shostak, S., & Hanna, B. (2017). The socio-exposome: Advancing exposure science and environmental justice in a postgenomic era. *Environmental Sociology*, *3*(2), 107–121. doi:10.1080/23251042.2016.1220848

Taylor, J. Y., Sun, Y. V., Hunt, S. C., & Kardia, S. L. R. (2010). Gene-environment interaction for hypertension among african american women across generations. *Biological Research for Nursing*, *12*(2), 149–155. doi:10.1177/1099800410371225

Thompson, M. R., & Schwartz Barcott, D. (2019). The role of the nurse scientist as a knowledge broker. *Journal of Nursing Scholarship*, *51*(1), 26–39. doi:10.1111/jnu.12439

Thursby, E., & Juge, N. (2017). Introduction to the human gut microbiota. *The Biochemical Journal*, *474*(11), 1823–1836. doi:10.1042/BCJ20160510

U.S. Global Change Research Program. (2016). *The impacts of climate change on human health in the United States: A scientific assessment*. Washington, DC: Author.

Vasilatou, D., Papageorgiou, S. G., Dimitriadis, G., & Pappa, V. (2013). Epigenetic alterations and microRNAs: New players in the pathogenesis of myelodysplastic syndromes. *Epigenetics*, *8*, 561–570. Taylor & Francis. doi:10.4161/epi.24897

Venkatesan, S., Khaw, A., & Hande, M. (2017). Telomere biology—Insights into an intriguing phenomenon. *Cells*, *6*(2), 15. doi:10.3390/cells6020015

Wright, M. L., Huang, Y., Hui, Q., Newhall, K., Crusto, C., Sun, Y. V., & Taylor, J. Y. (2017). Parenting stress and DNA methylation among African Americans in the InterGEN study. *Journal of Clinical and Translational Science*, *1*(6), 328–333. doi:10.1017/cts.2018.3

Appendix G.1 License Agreement for Manuscript 5

Copyright Clearance Center Marketplace™

Springer Publishing Company, Inc. - License Terms and Conditions

This is a License Agreement between Lacey W. Heinsberg / University of Pittsburgh ("You") and Springer Publishing Company, Inc. ("Publisher") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer Publishing Company, Inc., and the CCC terms and conditions.

All payments must be made in full to CCC.

Order Date	03-Apr-2020	**Type of Use**	Republish in a thesis/dissertation
Order license ID	1026519-2		
ISSN	1944-4028	**Publisher**	Springer Pub. Co
		Portion	Chapter/article

LICENSED CONTENT

Publication Title	Annual review of nursing research	**Rightsholder**	Springer Publishing Company, Inc.
Date	01/01/1983	**Publication Type**	Journal
Language	English	**URL**	http://www.springerpub.com/prod.aspx?prod_id=01267
Country	United States of America		

REQUEST DETAILS

Portion Type	Chapter/article	**Rights Requested**	Main product
Page range(s)	35-52	**Distribution**	Worldwide
Total number of pages	18	**Translation**	Original language of publication
Format (select all that apply)	Print, Electronic	**Copies for the disabled?**	Yes
Who will republish the content?	Academic institution	**Minor editing privileges?**	No
Duration of Use	Life of current edition	**Incidental promotional use?**	No
Lifetime Unit Quantity	Up to 250,000	**Currency**	USD

NEW WORK DETAILS

Title	Multi-Omics of the Iron Homeostasis Pathway in Patient Outcomes after Aneurysmal Subarachnoid Hemorrhage	**Institution name**	University of Pittsburgh
		Expected presentation date	2020-05-27
Instructor name	Lacey W. Heinsberg		

ADDITIONAL DETAILS

Order reference number	N/A

Appendix H Manuscript 6: Psychometric Properties of the Patient Assessment of Own Functioning Inventory Following Aneurysmal Subarachnoid Hemorrhage

[Journal Name], Volume 28, Number 2, 2020

Psychometric Properties of the Patient Assessment of Own Functioning Inventory Following Aneurysmal Subarachnoid Hemorrhage

Lacey W. Heinsberg, BSN, RN
Catherine M. Bender, PhD, RN, FAAN
Jeffrey M. Rohay, PhD, MSIS
Paula R. Sherwood, PhD, RN, CNRN, FAAN
Elizabeth A. Crago, PhD, RN
University of Pittsburgh, Pittsburgh, PA

Background and Purpose: The purpose of this study was to examine the psychometric properties of the Patient Assessment of Own Functioning Inventory (PAOFI) following aneurysmal subarachnoid hemorrhage (aSAH). **Methods:** The PAOFI was completed by 182 participants 3 months after verified aSAH. Exploratory factor analysis was used to evaluate the underlying factor structure of the PAOFI and reliability and concurrent validity were evaluated for each subscale. **Results:** A three-factor structure accounted for 58.9% of the extracted variance. Cronbach's alpha coefficients for extracted factors ranged from .867 to .924. The PAOFI subscales demonstrated concurrent validity with neuropsychological tests measuring similar constructs. **Conclusion:** There is evidence of reliability and validity of the PAOFI following aSAH. Further studies are needed to confirm these results.

Keywords: Stroke; Patient reported outcome measures; Psychometrics; Neuropsychological tests

Aneurysmal subarachnoid hemorrhage (aSAH) is a devastating type of stroke. While mortality rates have decreased in recent years, it remains a debilitating disease with a large fraction of survivors unable to return to functional independence (Shukla, 2017). Notably, cognitive impairment including memory loss, difficulty learning, and reduced ability to make decisions impacts survivors (Campbell et al., 2016), which is especially threatening to patient outcomes and potential productive years of life given the average age of aSAH onset of around 54 years (Ziemba-Davis et al., 2014).

http://dx.doi.org/10.1891/JNM-D-18-00117

BACKGROUND AND CONCEPTUAL FRAMEWORK

If cognitive decline after aSAH could be detected early through patient screening, the overall impact would be significant. Not only could this reduce patient and caregiver burden by identifying those in need of early therapeutic interventions but could also potentially reduce the financial burden of long-term skilled nursing care on our healthcare system. Currently, neuropsychological (NP) testing, a performance-based assessment of function, is a common standard of evaluation of cognitive consequences of disease and helps to identify areas that prohibit return to normal function after aSAH. However, there are important limitations of NP testing that should be acknowledged. First, NP test batteries are resource intensive. The substantial time commitment from patients and staff to complete NP test batteries makes them less practical in the clinical setting. Second, specific test batteries, time courses, and cutoff scores are frequently ambiguous between studies and generally underreported within the aSAH population making comparisons between studies difficult (Shukla, 2017).

In contrast, self-report cognitive measures are relatively quick to administer and may offer great clinical utility as a screening tool after aSAH. One such tool is the Patient Assessment of Own Functioning Inventory (PAOFI). The PAOFI is a tool intended to measure a continuum of self-reported functioning including subscales of Memory, Language and Communication, Use of Hands, Sensory-Perceptual, and Higher Level Cognitive and Intellectual Functions. The conceptual foundation and structure of the PAOFI have been clearly summarized elsewhere (Bell, Terhorst, & Bender, 2013; Chelune & Lehman, 1986) and the five subscales and detailed items of the PAOFI have been outlined in detail in Supplemental Table 1. This tool has been utilized in a range of populations including persons with human aids virus (Saylor et al., 2017), breast cancer (Merriman et al., 2017), and patients receiving dialysis (Song et al., 2015). However, comprehensive examination of the psychometric properties of this tool has only been reported in a handful of studies outlined in Supplemental Table 2, including a study of subjects referred for NP evaluation

TABLE 1. Comparison of Sample Characteristics Between Groups

	Included in Analyses	**Excluded (Outliers)**	**Did Not Complete PAOFI**
Variable	***n* = 182 (35.5%)**	***n* = 22 (4.3%)**	***n* = 308 (60.2%)**
Age (years)			
Mean	52.8	52.6	53.2
SD	10.1	11.1	11.8
Range	24–75	26–68	25–75
Sex			
Female, *n (%)*	133 (73.1)	17 (77.3)	221 (71.8)
Male, *n (%)*	49 (26.9)	5 (22.7)	87 (28.2)

(*Continued*)

TABLE 1. Comparison of Sample Characteristics Between Groups (*Continued*)

	Included in Analyses	Excluded (Outliers)	Did Not Complete PAOFI
Variable	***n* = 182 (35.5%)**	***n* = 22 (4.3%)**	***n* = 308 (60.2%)**
Race			
White, *n (%)*	167 (91.8)	15 (68.2)	256 (83.1)
African American, *n (%)*	11 (6.0)	6 (27.3)	40 (13.0)
Other, *n (%)*	4 (2.2)	1 (4.5)	12 (3.9)
Marital Status			
Single, *n (%)*	41 (22.5)	3 (13.7)	75 (24.4)
Married, *n (%)*	115 (63.2)	15 (68.3)	171 (55.5)
Divorced, *n (%)*	17 (9.3)	2 (9.0)	39 (12.7)
Other, *n (%)*	9 (5)	2 (9)	23 (7.4)
Years of education			
≤12 years, *n (%)*	82 (45.1)	14 (63.6)	43 (14.0)
13–16 years, *n (%)*	76 (41.8)	5 (22.7)	28 (9.0)
>16 years, *n (%)*	13 (7.1)	2 (9.1)	7 (2.3)
Unknown, *n (%)*	11 (6.0)	1 (4.6)	230 (74.7)
HH Grade			
1, *n (%)*	21 (11.5)	1 (4.5)	22 (7.1)
2, *n (%)*	81 (44.5)	10 (45.5)	103 (33.4)
3, *n (%)*	57 (31.3)	7 (31.8)	110 (35.7)
4, *n (%)*	19 (10.4)	4 (18.2)	47 (15.3)
5, *n (%)*	4 (2.3)	0 (0)	26 (8.4)
Aneurysm Location			
Anterior, *n (%)*	148 (81.3)	17 (77.3)	204 (66.2)
Posterior, *n (%)*	34 (18.7)	5 (22.7)	104 (33.8)
Intervention			
Surgical, *n (%)*	57 (31.3)	12 (54.5)	93 (30.2)
Coil, *n (%)*	125 (68.7)	10 (45.5)	213 (69.2)
None, *n (%)*	0 (0)	0 (0)	2 (0.6)

(*Continued*)

TABLE 1. Comparison of Sample Characteristics Between Groups (*Continued*)

	Included in Analyses	Excluded (Outliers)	Did Not Complete PAOFI
Variable	***n* = 182 (35.5%)**	***n* = 22 (4.3%)**	***n* = 308 (60.2%)**
mRS at 3 months			
Favorable (0–2), *n (%)*	150 (82.4)	13 (59.1)	63 (20.5)
Unfavorable (3–6), *n (%)*	21 (11.5)	6 (27.3)	95 (30.8)
Unknown, *n (%)*	11 (0.06)	3 (13.6)	150 (48.7)

Note. Table 1 presents results of a comparison of demographics and clinical characteristics for participants included in the analysis, participants labeled as outliers and ultimately excluded from the analysis, and participants who did not complete the PAOFI; HH = Hunt and Hess grade; mRS = Modified Rankin Scale; PAOFI = Patient Assessment of Own Functioning Inventory; SD = standard deviation.

TABLE 2. Internal Consistency of Original PAOFI Subscales Within aSAH Cohort

PAOFI Subscale	Cronbach's alpha	Number of Items
I	.876	9
II[a]	.870	8
III	.422	2
IV	.303	3
V	.923	9
Entire PAOFI[a]	.950	31

Note. Table 2 presents Cronbach's alpha for each subscale as an evaluation of the internal consistency of the original PAOFI subscales within the current sample (*n* = 182); aSAH = aneurysmal subarachnoid hemorrhage; PAOFI = Patient Assessment of Own Functioning Inventory.
[a]Item 13a excluded from analysis.

(Chelune & Lehman, 1986), postmenopausal women with breast cancer (Bell et al., 2013), and survivors of breast cancer following surgical intervention or chemotherapy (Van Dyk, Ganz, Ercoli, Petersen, & Crespi, 2016). An additional study Richardson-Vejlgaard, Dawes, Heaton, and Bell (2009) evaluated the validity of the PAOFI in a sample of veterans recruited from a daily substance abuse treatment program and a sample of nonclinical controls (Richardson-Vejlgaard et al., 2009).

Psychometric evaluation of the PAOFI following aSAH or other types of stroke has not been performed. The purpose of this study was to fill this gap in knowledge by evaluating the psychometric properties of the PAOFI after aSAH. Given the time and resource limitations of NP test batteries, if the PAOFI is found to be reliable after aSAH it may offer great potential as a clinically relevant and simple method to help nurses screen patients to identify those in need of follow up with more comprehensive NP testing. Furthermore, if reliable after aSAH, this tool may offer substantial insight for care planning and rehabilitation by identifying self-perceptions of function and difficult to detect deficits in patients with the capacity to complete the PAOFI.

METHODS

Participants

This cross-sectional study was a secondary data analysis of the parent study which included participants recruited through ongoing prospective studies at (location and grant numbers masked for blind review) approved by the [location masked for blind review) Institutional Review Board. Participants were recruited to the parent study from (location masked for blind review] Hospital neurovascular intensive care unit and informed consent was obtained from the participant or their legal representative. Prospective participants were eligible for the parent study if they were recently diagnosed (≤5 days) with subarachnoid hemorrhage verified by cerebral angiogram, aged 18 to 75 years, and could read and speak English. Prospective participants were not eligible for the parent study if they suffered subarachnoid hemorrhage due to trauma, mycotic aneurysm, arteriovenous malformation, or unknown cause or if they had a history of a severe neurological disorder.

A schematic explanation of the sample size for this secondary data analysis is represented in Figure 1. The parent study attempted to collect PAOFI data on 512 participants. Of these, 297 did not complete the PAOFI questionnaire for reasons including death (n = 62), cognitively unable (n = 33), refusal (n = 33), or because we were unable to arrange a visit (n = 169). In addition, data screening revealed 11 participants with incomplete data. Data for 22 participants were found to be influential outliers resulting in a final sample

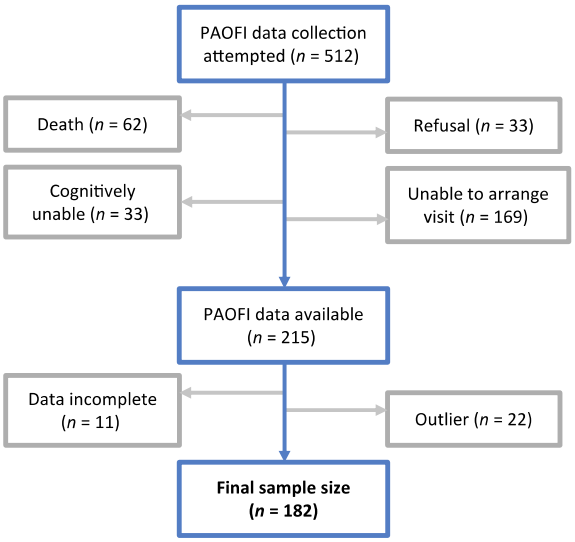


Figure 1. Flowchart depicting explanation of sample size for psychometric evaluation of the PAOFI.

Note. Figure 1 depicts an explanation of the sample size for the current study. The parent project of this ancillary study attempted to collect PAOFI data on 512 participants. Of these, 297 did not complete the PAOFI questionnaire for assorted reasons including death (n = 62), cognitively unable (n = 33), refusal (n = 33), or because we were unable to contact the participant, arrange a visit, or an unknown reason (n = 169). In addition, data screening revealed 11 participants with incomplete data and an additional 22 were excluded as outliers. The final sample size consisted of 182 participants. PAOFI = Patient Assessment of Own Functioning Inventory.

size of $n = 182$. Outcomes were assessed at 3 months following aSAH by a nurse or technician trained in NP testing and included a battery of NP tests and functional outcome measures.

Measures

Demographic and Clinical Data. Demographic and clinical data were extracted from the medical record including age, sex, aneurysm location, intervention, and Hunt and Hess (HH) clinical grading scale. The HH scale is a grading tool specifically used to evaluate the clinical severity of subarachnoid hemorrhage containing 5 grades where grade I includes mild symptoms of headache progressing through grade V which indicates a patient is comatose (Hunt & Hess, 1968). Grading was completed by attending physicians at the time of admission. Participants' marital status, race, and education level were self-reported by the participants or their proxy and included in the current study because these factors have been shown to be associated with self-reported outcomes (Mohd Zulkifly, Ghazali, Che Din, & Subramaniam, 2016).

Patient Assessment of Own Functioning Inventory. Participants' self-reported assessment of their ability to function was measured at 3 months following aSAH using the PAOFI. The PAOFI is a self-report measure of participants' perceived ability to function and perform everyday tasks. Each item is rated on a scale of *0 (almost never)* to *5 (almost always)*. This tool produces 5 subscale scores including Memory, Language and Communication, Use of Hands, Sensory-Perceptual, and Higher Level Cognitive and Intellectual Functions that sum to one total score. Lower PAOFI scores indicate higher levels of functioning. A summary of the PAOFI questionnaire can be found in Supplemental Table 1 (Chelune & Lehman, 1986).

Modified Rankin Scale (mRS). The mRS incorporates cognitive and functional deficits resulting from neurological injury using a composite score from *0 (no symptoms)* to *6 (death)* and has well established validity after stroke (Banks & Marotta, 2007). The mRS was used as a global measure of functional status to examine differences between participants included in the present study and those labeled as outliers, to help interpret study results, and to determine the study generalizability.

Wechsler Memory Scale III (WMS-III), Logical Memory Subset. The WMS-III has demonstrated clinical validity (Price, Tulsky, Millis, & Weiss, 2002) to assesses memory and retention using a subset of tests which include two stories. During this test, participants are instructed to read two stories (Story A and B) and are asked to repeat the stories from memory immediately, and again after 30 minutes (Wechsler, 1997). Higher scores indicate better performance as participants are credited for general story themes and for each correctly recalled detail.

Controlled Word Association Test. The controlled word association test (COWA) assesses language construct in addition to executive function and cognitive flexibility and has well established validity (Tombaugh, Kozak, & Rees, 1999). The COWA assesses the participant's ability to generate a spontaneous list of words beginning with a designated letter within 1 minute (Benton & Hamsher, 1989). Higher scores indicate better performance as participants are credited for each word generated.

Stroop Color Word Test (Stroop). The Stroop measures executive functioning and has been validated in patients with neurologic injury (Guise, Thompson, Greve, Bianchini, & West, 2014). In the current study we report the Stroop Color/Word trial which measures the relative speed of participants' ability to name colors from words of colors printed with

an incongruently colored ink (Golden, 2002). In this test, higher scores indicate better performance and are adjusted for age.

Trail Making Test B (Trails B). The Trails B is a validated measure of executive functioning as well as mental flexibility and attention (Sánchez-Cubillo et al., 2009). In the Trails B test, participants are presented with a sheet of paper with circled numbers and letters and are instructed to draw lines connecting the circles in order, alternating between numbers and letters as quickly as possible (Reitan, 1955). The total time to complete the trial is recorded with lower scores indicating better performance.

Data Analysis

Data Screening. All data analyses were conducted using SPSS, version 24 (IBM, Chicago, IL, USA). Data were screened statistically and graphically to evaluate missing data and item response distribution. Mahalanobis distance was generated to identify potentially influential multivariate outliers and the Shapiro–Wilk test of normality was used to evaluate univariate distribution of item responses. Characteristics of participants identified as outliers were compared with those of participants included in the study to evaluate their potential influence. PAOFI item clusters were identified using inter-item correlations, sample size adequacy was evaluated using the Kaiser–Meyer–Olkin (KMO) test, and multicollinearity was examined using Bartlett's test of sphericity. Internal consistency of the original five subscales was evaluated using Cronbach's alpha statistic (Cronbach, 1951).

Factor Analysis. To evaluate the factor structure of the PAOFI, exploratory factor analysis (EFA) using principal axis factoring (PAF) was used because of the general floor effect and non-normal distribution of PAOFI item responses (Harman, 1976). Consistent with the methodology used by Bell et al., 2013, varimax rotation with Kaiser normalization was selected to simplify interpretation of results and to evaluate the uncorrelated individual contributions of factors. First, PAF using varimax rotation was attempted on all five subscales using a five-factor structure. In our sample, two subscales (Use of Hands [items 18–19] and Sensory-Perceptual [items 20–22]) had low validity causing illogical and noninterpretable factor loadings. Thus, similar to the methodology used by Richardson-Vejlgaard et al. (2009), these motor and sensory subscales were removed and only the remaining three subscales were included in the final analysis (Memory [items 1–9], Language and Communication [items 10–17, excluding 13a], and Higher Level Cognitive and Intellectual Functions [items 23–31]). PAF was repeated using a three-factor structure. Following PAF, components were retained based on visual evaluation of the scree plot, eigenvalues >1, and if at least 50% of the variance in responses was explained by the extracted factors (Pett, Lackey, & Sullivan, 2003). The threshold for loading and cross-loading was set at ≥ 0.4 (Raubenheimer, 2004). With 26 items in this analysis, the sample size of 182 participants was sufficient to support EFA given the recommended participant-to-item ratio of one-to-five (Bryant & Yarnold, 1994).

Reliability and Concurrent Validity. The reliability of the proposed 3-factor structure was evaluated. Cronbach's alpha statistic was computed for each subscale with an alpha of ≥.7 considered to be reliable in this sample (DeVellis, 2012). The final item composition, item-total correlations, and inter-item correlations were inspected. A standard measure comparative to the PAOFI was not available in the existing data; therefore, we evaluated the correlation of the proposed 3-factor structure with NP measures that evaluate similar constructs. Attempts to normalize the PAOFI data failed, preventing the evaluation of Pearson's correlations. Thus, Spearman's correlation was used to evaluate the

relationships between variables. We evaluated correlations between the PAOFI Memory subscale and the WMS-III Logical Memory Subset (Story A and Story B Immediate and Delayed Recall Unit Scores); the PAOFI Language and Communications subscale with the COWA; and the PAOFI Higher Level Cognitive Functions subscale with both the Stroop and Trails B tests. Correlations with p-values < .05 were considered statistically significant.

RESULTS

Data Screening

Data were screened statistically and graphically for normality, missingness, and outliers. As previously stated, data screening revealed 11 participants with incomplete data. Given the small sample size and diffuse cognitive deficits following aSAH, values for missing data were not imputed to avoid biasing study results. In the current study, data for 22 participants were found to be influential outliers based on a Mahalanobis distance >61.098.

To understand the differences between participants found to be outliers based on Mahalanobis distance, characteristics between groups (outliers in the current study [n = 22] vs. participants passing data screening [n = 182] vs. parent study participants who did not complete the PAOFI [n = 308]) were compared and these results are presented in Table 1. In the outlier group, 27.3% of participants self-reported their race as African American compared with only 6% in the group included in the present analysis. Furthermore, within the outlier group, 54.5% had a surgical intervention following aSAH (versus embolization or coil) compared with only 31.3% of those included in the present analysis and 30.2% of those unable to complete the PAOFI. Lastly, 82.4% of the participants included in the current study had a favorable mRS score of 0, 1, or 2 at 3 months indicating a majority of participants were relatively independent. In the outlier group and those unable to complete the PAOFI, 59.1% and 20.5% respectively had a favorable mRS score at 3 months. No further key differences were identified based on age, sex, marital status, HH grade, or aneurysm location. Data were analyzed with and without influential outliers and results were found to be discordant. Given the disparity in aSAH outcomes and differences in participant characteristics identified between outliers and those included in the study (in particular, a majority 82.4% with favorable mRS scores), it was ultimately decided to exclude the outliers from the final analysis.

Graphical evaluation of item responses suggested a non-normal distribution of data with a substantial floor effect (favoring lower PAOFI scores and suggesting better outcomes). Non-normality was confirmed with the Shapiro-Wilk test of normality ($p < 0.001$). Within the Language and Communication subscale, almost all participants answered “almost never” to item 13 (When you speak are your words indistinct or improperly pronounced?). Thus, the follow up question, item 13a (How often do people have difficulty understanding what you are trying to say?), was dropped because it did not apply to a majority of participants.

The final sample of 182 participants were 91.8% White, which is consistent with the expected racial distribution based on local demographics, and 73.1% female with a mean age of 52.8 ± 10.1 years, which is consistent with sex and age distributions within the aSAH population (Ziemba-Davis et al., 2014). Given the potential impact of marital status

on self-assessment measures, we examined the current marital status of participants (Mohd Zulkifly et al., 2016). In the current sample, 63.2% of participants were married while the remaining were either single, divorced, or widowed. Finally, HH grade is often predictive of outcomes after aSAH with lower grades associated with better outcomes (Hunt & Hess, 1968). Of note, over half (56%) of participants who completed the PAOFI had a HH grade of 1 or 2, indicating good clinical status upon arrival to the hospital.

Factor Analysis

Inter-item correlations were examined and ranged from 0.199 to 0.684. The KMO test statistic was equal to 0.931 and Bartlett's test of sphericity was significant (<0.001), thus the item correlation matrix was determined to be factorable (Pett et al., 2003). The internal consistency of the original subscales was evaluated using Cronbach's alpha and the results are presented in Table 2. Cronbach's alpha for the original subscales ranged from low (0.303, Sensory-Perceptual subscale) to high (0.923, Higher Level Cognitive and Intellectual Functions). After excluding the Use of Hands and Sensory-Perceptual subscales as previously described, PAF using varimax rotation and a three-factor structure identified three factors with eigenvalues >1 that accounted for 58.9% of the extracted variance. Proposed factor loadings are presented in Table 3. All items in this population loaded on the original structure of the PAOFI with the exception of item 9 (How often do you forget things you are supposed to do or have agreed to do?). Item 9 was originally included as part of the Memory subscale, but in our analysis, this item loaded under Higher Level Cognitive and Intellectual Functions. Item communalities ranged from 0.251 to 0.780 after extraction. The three extracted factors were labeled according to the original structure of the PAOFI including Factor 1 (Higher Level Cognitive and Intellectual Functions), Factor 2 (Language and Communication), and Factor 3 (Memory).

Reliability and Concurrent Validity

The results of the evaluation of reliability are presented in Table 3. Internal consistency of factors was evaluated using the Cronbach's alpha coefficient; Factors 1, 2, and 3 were found to have a Cronbach's alpha of .924, .879, and .867 respectively. The Cronbach's alpha for the finalized structure was .953. The results of concurrent validity evaluation examining the correlation between the PAOFI subscales with NP tests measuring similar constructs are presented in Table 4. The PAOFI Memory subscale showed a small to moderate correlation with the WMS-III Story A Immediate Recall Unit Score (r^2 = –0.174, p = .021), Story A Delayed Recall Unit Score (r^2 = –0.228, p = .002), Story B Immediate Recall Unit Score (r^2 = –0.218, p = .004), and the Story B Delayed Recall Unit Score (r^2 = –0.246, p = .001) indicating lower scores on the PAOFI were correlated with higher scores on the WMS Story tests. Similarly, the PAOFI Higher Level Cognitive and Intellectual Functions subscale was correlated with both the Stroop test (r^2 = –0.269, p < .001) and the Trails B test (r^2 = 0.200, p = .008) indicating lower scores on the PAOFI were correlated with higher scores on the Stroop test and lower scores on the Trails B test. In contrast, no significant correlation was found between the PAOFI Language and Communication subscale and the COWA (r^2 = –0.128, p = .089).

TABLE 3. Factor Loadings, Percent of Variance Explained From Exploratory Factor Analysis Using Principal Axis Factoring With Varimax Rotation and Kaiser Normalization, and Internal Consistency of Proposed Factors

PAOFI Item	Original Scale	Question	Factor 1	Factor 2	Factor 3
Memory					
1	Memory	How often do you forget something that has been told to you within the last day or two?			0.679
2	Memory	How often do you forget events which have occurred in the last day or two?			0.672
3	Memory	How often do you forget people whom you met in the last day or two?			0.599
4	Memory	How often do you forget things that you knew a year or more ago?			0.412
5	Memory	How often do you forget people whom you knew or met a year or more ago?			0.514
6	Memory	How often do you lose track of time, or do things either earlier or later than they are usually done or are supposed to be done?			0.571
7	Memory	How often do you fail to finish something you start because you forgot that you were doing it? (Include such things as forgetting to put out cigarettes, turn off the stove, etc.)			0.498

(*Continued*)

TABLE 3. Factor Loadings, Percent of Variance Explained From Exploratory Factor Analysis Using Principal Axis Factoring With Varimax Rotation and Kaiser Normalization, and Internal Consistency of Proposed Factors (*Continued*)

PAOFI Item	Original Scale	Question	Factor 1	Factor 2	Factor 3
8	Memory	How often do you fail to complete a task that you start because you have forgotten how to do one or more aspects of it?			0.419
Language and communication					
10	Language and communication	How often do you have difficulties understanding what is said to you?		0.546	
11	Language and communication	How often do you have difficulties recognizing or identifying printed words?		0.624	
12	Language and communication	How often do you have difficulty understanding reading material which at one time you could have understood?		0.56	
13	Language and communication	When you speak, are your words indistinct or improperly pronounced?		0.413	
14	Language and communication	How often do you have difficulty thinking of the names of things?		0.464	
15	Language and communication	How often do you have difficulty thinking of the words (other than names) for what you want to say?		0.575	

(*Continued*)

TABLE 3. Factor Loadings, Percent of Variance Explained From Exploratory Factor Analysis Using Principal Axis Factoring With Varimax Rotation and Kaiser Normalization, and Internal Consistency of Proposed Factors (*Continued*)

PAOFI Item	Original Scale	Question	Factor 1	Factor 2	Factor 3
16	Language and communication	When you write things, how often do you have difficulty forming the letters correctly?		0.448	
17	Language and communication	Do you have more difficulty spelling, or make more errors in spelling, than you used to?		0.569	
Higher level cognitive and intellectual functions					
9	Memory	How often do you forget things that you are supposed to do or have agreed to do(such as putting gas in the car, paying bills, taking care of errands, etc.)?	0.433		
23	Higher level cognitive and intellectual functions	How often do your thoughts seem confused or illogical?	0.600		
24	Higher level cognitive and intellectual functions	How often do you become distracted from what you are doing or saying by insignificant things which at one time you would have been able to ignore?	0.654		
25	Higher level cognitive and intellectual functions	How often do you become confused about (or make a mistake about) where you are?	0.427		
26	Higher level cognitive and intellectual functions	How often do you have difficulty finding your way about?	0.632		

(*Continued*)

TABLE 3. Factor Loadings, Percent of Variance Explained From Exploratory Factor Analysis Using Principal Axis Factoring With Varimax Rotation and Kaiser Normalization, and Internal Consistency of Proposed Factors (*Continued*)

PAOFI Item	Original Scale	Question	Factor 1	Factor 2	Factor 3
27	Higher level cognitive and intellectual functions	Do you have more difficulty now than you used to in calculating or working with numbers (including managing finances, paying bills, etc.)?	0.63		
28	Higher level cognitive and intellectual functions	Do you have more difficulty now than you used to in planning or organizing activities (example: deciding what to do and how it should be done)?	0.767		
29	Higher level cognitive and intellectual functions	Do you have more difficulty now than you used to in solving problems that come up around the house, at your job, and so on? (In other words, when something new has to be accomplished, or some new difficulty comes up, do you have more trouble figuring out what should be done and how to do it)?	0.805		
30	Higher level cognitive and intellectual functions	Do you have more difficulty now than you used to in following directions to get somewhere?	0.626		

(*Continued*)

TABLE 3. Factor Loadings, Percent of Variance Explained From Exploratory Factor Analysis Using Principal Axis Factoring With Varimax Rotation and Kaiser Normalization, and Internal Consistency of Proposed Factors (*Continued*)

PAOFI Item	Original Scale	Question	Factor 1	Factor 2	Factor 3
31	Higher level cognitive and intellectual functions	Do you have more difficulty now than you used to in following instructions concerning how to do things?	0.727		
Total variance explained (58.9%)			47.5%	6.5%	4.9%
Cronbach's alpha (.953)			0.924	0.879	0.867

Note. Table 3 presents the proposed factor loadings, percent of variance explained from exploratory factor analysis, and internal consistency of each proposed factor.

DISCUSSION

The purpose of this study was to evaluate the psychometric properties of the PAOFI after aSAH. To our knowledge, it is the first comprehensive examination of this tool in this population. While a self-assessment questionnaire such as the PAOFI is potentially very useful as a screening tool to help identify patients in need of more comprehensive NP testing after aSAH, the results of this analysis suggest that the original five-factor structure is unreliable and that use of the PAOFI may need to be limited to the Memory, Language and Communication, and Higher Level Cognitive and Intellectual Functions subscales in this population. After excluding the two subscales measuring functional and sensorimotor domains, we identified a reliable factor structure almost identical to that of the original PAOFI. Notably, Cronbach's alpha for each identified factor, as well as the final combined factor structure, were excellent with all values greater than .85 exceeding the limit for reliability of .7 (DeVellis, 2012).

In the present study, the use of all five subscales from the original PAOFI resulted in illogical factor loadings within this population. The low validity of these subscales, as measured by Cronbach's alpha, and the resulting noninterpretable factor structure produced by the factor analysis could be explained by extreme variability in physical functional outcomes after aSAH (including sensory and motor ability; Lantigua et al., 2015). While a majority of items in the PAOFI apply to cognitive function, five items apply to physical and sensorimotor functioning and are further divided into left and right subgroups (two items ask about right hand functioning and two ask about left hand functioning). It is possible that in this small sample, responses regarding ability to function were too variable between participants based on the location of injury in the brain preventing the identification of the underlying structure of these two subscales. Unfortunately, we were unable to evaluate the impact of brain aneurysm location on the factor structure of the PAOFI given the sample size of 182 participants. It is important that the influence of brain aneurysm location be examined in future studies of larger sample size.

TABLE 4. Evaluation of Concurrent Validity (Correlations of PAOFI Subscales With Neuropsychological Tests Measuring Similar Constructs)

PAOFI Subscale		Neuropsychological Test			
		Wechsler Memory Scale III, Logical Memory Subset			
		Story A immediate recall unit score	*Story A delayed recall unit score*	*Story B immediate recall unit score*	*Story B delayed recall unit score*
Memory subscale (n = 176)	Spearman correlation	–0.174	–0.228	–0.218	–0.246
	p-value (2-tailed)	0.021	0.002	0.004	0.001
		The Controlled Word Association Test			
		Summed F+A+S words			
Language and communication subscale (n = 177)	Spearman correlation	–0.128			
	p-value (2-Tailed)	0.089			
		Stroop Color Word Test			
		Stroop color/word T-score			
Higher level cognitive and intellectual functions subscale (n = 165)	Spearman correlation	–0.269			
	p-value (2-tailed)	< 0.001			
		Trail Making Test B			
		Trails B score			
Higher level cognitive and intellectual functions subscale (n = 177)	Spearman correlation	0.2			
	p-value (2-tailed)	0.008			

Note. Table 4 reports results from the evaluation of concurrent validity showing Spearman correlations of the proposed PAOFI subscales with neuropsychological tests measuring similar constructs; PAOFI = Patient Assessment of Own Functioning Inventory; *p*-values < .05 are considered significant.

Of importance, 9 of the 26 items cross-loaded on more than one factor (original PAOFI questionnaire summarized in Supplemental Table 1; results of EFA and new factor loadings summarized in Table 3). All items, with the exception of one, loaded highest on their respective original subscales and were subsequently designated to a matching factor (i.e., new "Memory" factor was labeled based on the original "Memory" subscale). Specifically, item 8, which poses the question, "How often do you fail to complete a task that you start because you have forgotten how to do one or more aspects of it?" loaded on all three factors with the highest loading value on the Higher Level Cognitive and Intellectual Functions subscale. In contrast with the other items, it was ultimately decided that this item should be included with the new Memory factor (vs. the factor it loaded highest on) because its conceptual fit was more appropriate. Additional cross-loading items included items 5, 7, 12, 14, 15, 17, 23, and 30. A potential explanation for this is that ambiguous items may be interpreted differently by participants or be involved in more than one cognitive domain. Cross loading is an important consideration because it potentially impacts conclusions about the tool's content validity and replicability of study findings.

Outside of the current study, evaluation of the psychometric properties of the PAOFI using factor analysis has been limited to three studies and are outlined in Supplemental Table 2 (Bell et al., 2013; Chelune & Lehman, 1986; Van Dyk et al., 2016). Although the results of this study identified different factor loadings than that of Bell et al. (2013) and Van Dyk et al. (2016), it is important to consider the inherent differences between the populations examined. Following aSAH, survivors are most often met with deficits in memory, language, and executive function (Al-Khindi, MacDonald, & Schweizer, 2010) and, even in a study of aSAH participants with good outcomes, cognitive impairment was detected in an overwhelming 70% of participants (Ravnik et al., 2006). The complex and diffuse disturbances in cognitive function after aSAH may be a potential explanation as to why the responses to items in this population seem to have different factor loadings than those previously examined such as postmenopausal women with breast cancer.

In a study by Richardson-Vejlgaard et al. (2009), the authors evaluated the validity of the PAOFI in a sample of veterans recruited from a daily substance abuse treatment program (Richardson-Vejlgaard et al., 2009). Similar to the current study, the authors removed the Use of Hands subscale and the Sensory-Perceptual subscale prior to evaluation. While a factor analysis of the PAOFI was not performed in this study, the authors did report Pearson correlations between PAOFI scores and factor scores of an objective NP test battery. Interestingly, in contrast to the current study, examination of these relationships failed to demonstrate concurrent validity. Of note, existing research, although mixed, highlights discrepancies between subjective symptomatology questionnaires and objective NP testing in other populations (Srisurapanont, Suttajit, Eurviriyanukul, & Varnado, 2017). An important contrast to Richardson-Vejlgaard et al. (2009) in the current study is the correlation between two out of three PAOFI subscales with NP tests measuring similar constructs as outlined in Table 4. Specifically, the PAOFI Memory subscale was correlated with the WMS-III test and the Higher Level Cognitive and Intellectual Functions subscale was correlated with both the Stroop and Trails B tests. Although the correlations identified were only small to moderate, the directionality of the correlations were as expected, and the *p*-values in these analyses were strong, suggesting these results are not just a product of sample size and noise.

This study is significant in that it is the first to evaluate the factor structure, reliability, and concurrent validity of a self-assessment tool following aSAH, however there are several important limitations that should be considered. First, as with many patient outcome

studies, the study participant must possess the physical and cognitive ability to complete the self-assessment questionnaire, therefore, results are inherently biased toward those with better outcomes after aSAH (Shukla, 2017). In the current study, over half of the participants presented with an admission HH grade of 1 to 2 (moderate to severe headache without focal symptoms), which has been associated with better outcomes. In addition, the majority of participants had a favorable mRS score of 0 (no symptoms), 1 (no significant disability), or 2 (slight disability but unable to perform all activities) at 3 months after aSAH indicating a majority of participants were relatively independent. However, cognitive impairment is a common problem even in aSAH patients with good outcomes (Ravnik et al., 2006) and it is specifically this group that the PAOFI might be most useful for.

Second, the sample size of this study prohibited subgroup psychometric evaluation of the PAOFI to determine if the underlying structure is different based on participant factors such as aneurysm location, medical intervention, or HH grade. It is possible that differential factor loadings could occur based on aneurysm and brain damage location, potentially altering the utility of the PAOFI. Because the current sample included a range of clinical severity and aneurysm locations, the results of this study are preliminary, reflecting the general population of aSAH patients with favorable outcomes (compared to those who have significant neurological impairment) and may not capture potential finer differences in the underlying factor structure between subgroups of aSAH patients. Likewise, there was a disparity in the distribution of men and women. This sample consisted of approximately 73% female participants and while this number is similar to aSAH demographics (Ziemba-Davis et al., 2014), previous studies have shown that women are more likely to self-report physical and cognitive deficits (Barsky, Peekna, & Borus, 2001).

Third, while the parent project for this secondary data analysis collected a comprehensive battery of NP measures, we were limited to the available data. The NP tests included in this study may not be measuring the same constructs as the PAOFI subscales. For example, the COWA, the only available NP measure related to language in the parent study, did not correlate with the PAOFI Language and Communication subscale suggesting these tools measure different constructs. While the directionality and small to moderate correlations between both the PAOFI Memory subscale and the Higher Level Cognitive and Intellectual Functions subscale with selected NP measures may support the findings of this study, they should be interpreted with caution as a content validity analysis comparing these tests was not conducted. While the PAOFI potentially offers great clinical utility as a screening tool after aSAH in patients with generally good outcomes, it is important to note that the information gained by NP testing and self-assessment tools yield different, yet valuable assessments of patient outcomes following aSAH.

Finally, it should be noted that 10.8% of the original sample (n = 22) were labeled outliers based on Mahalanobis distance and therefore excluded from the present analysis. There were several distinctions between the outliers and those included in the present study. Participants in the outlier group had a higher percentage of surgical aneurysm repair. Interestingly, it has been shown that surgical intervention after aSAH results in significantly greater disability in the short term, but no difference in the long term (Taheri et al., 2015). It is possible that if we had evaluated the psychometric properties of the PAOFI at 12 months after aSAH as opposed to at 3 months, recovery could have stabilized potentially reducing the number of outliers within our sample. Furthermore, the outlier group had a higher percentage of participants who self-reported their race as African American compared with the group included in the analysis. It has been shown that patients who self-report as African American have an increased risk of death after aSAH (Rivero Rodríguez

et al., 2015). Given that the present sample was approximately 92% White, consistent with demographics of the area in which participants were recruited, we were unable to explore this finding further, which impacts the generalizability of the study. It is important that this study be replicated in a larger group of more diverse races in the future to further evaluate the underlying factor structure. Finally, available education data in the overall cohort was largely incomplete interfering with meaningful evaluation between groups. While education could have an influence on the results of this study, we were unable to evaluate this due to missing data.

RELEVANCE TO NURSING PRACTICE

The results of this study demonstrate that the PAOFI may be a valid and reliable test in aSAH patients with generally good outcomes, however, the factor structure identified through this EFA should be confirmed in an independent test sample in the future. These findings, if corroborated in a larger sample, could impact current nursing practice. A major limitation of standard NP test batteries is that they are often unavailable in the clinical setting due to time and resource limitations. This study indicates the PAOFI may provide insight into patient functioning and offers potential as a screening tool to help nurses to quickly identify patients with cognitive decline who are in need of personalized interventions to improve their care following aSAH. Moreover, while the inability of the PAOFI to differentiate domain-specific dysfunction limits the ability to utilize it in developing a patient-centered rehabilitation program, its overall reliability and validity support its use as a clinically relevant and simple method to screen for those patients who need follow up with more comprehensive NP testing. A screening tool such as this within the aSAH population has the potential to decrease the number of patients who are referred for objective testing, lessening the burden on the healthcare system.

REFERENCES

Al-Khindi, T., MacDonald, R. L., & Schweizer, T. A. (2010). Cognitive and functional outcome after aneurysmal subarachnoid hemorrhage. *Stroke*, *41*(8), e519–e536. https://doi.org/10.1161/STROKEAHA.110.581975

Banks, J. L., & Marotta, C. A. (2007, March 1). Outcomes validity and reliability of the modified rankin scale: Implications for stroke clinical trials - A literature review and synthesis. *Stroke*, *38*, 1091–1096. https://doi.org/10.1161/01.STR.0000258355.23810.c6

Barsky, A. J., Peekna, H. M., & Borus, J. F. (2001, April). Somatic symptom reporting in women and men. *Journal of General Internal Medicine*, *16*, 266–275. https://doi.org/10.1046/j.1525-1497.2001.016004266.x

Bell, M. J., Terhorst, L., & Bender, C. M. (2013). Psychometric analysis of the Patient Assessment of Own Functioning Inventory in women with breast cancer. *Journal of Nursing Measurement*, *21*(2), 320–334. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24053059

Benton, A., & Hamsher, K. (1989). *Controlled Oral Word Association Test*. Iowa City: AJA Associates.

Bryant, F., & Yarnold, P. (1994). Principal components, and exploratory and confirmatory factor analysis. In L. Grimm & P. Yarnold (Eds.), *Reading and understanding multivariate statistics*. Washington, DC: American Psychological Association. Retrieved

from https://www.researchgate.net/publication/261062471_Principal_components_and_exploratory_and_confirmatory_factor_analysis

Campbell, N. L., Zhan, J., Tu, W., Weber, Z., Ambeuhl, R., McKay, C., & McElwee, N. (2016). Self-reported medication adherence barriers among ambulatory older adults with mild cognitive impairment. *Pharmacotherapy*, *36*(2), 196–202. https://doi.org/10.1002/phar.1702

Chelune, G., & Lehman, R. (1986). Neuropsychological and personality correlates of patients' complaints of disability. In G. Goldstein & R. E. Tarter (Eds.), *Advances in clinical neuropsychology* (pp. 95–118). New York, NY: Plenum Press.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

DeVellis, R. F. (2012). L. Bickman & D. J. Rog (Eds.), *Scale development: Theory and applications* (L. Bickman & D. J. Rog, eds., 3rd ed.). Los Angeles, CA: SAGE. Retrieved from https://books.google.es/books/about/Scale_Development.html?id=vmwBHYuchfAC&redir_esc=y L. Bickman & D. J. Rog, eds.,

Golden, C. (2002). *Stroop color and word test*. : Luts; PAR/Psychological Assessment Resources.

Guise, B. J., Thompson, M. D., Greve, K. W., Bianchini, K. J., & West, L. (2014). Assessment of performance validity in the stroop color and word test in mild traumatic brain injury patients: A criterion-groups validation design. *Journal of Neuropsychology*, *8*(1), 20–33. https://doi.org/10.1111/jnp.12002

Harman, H. (1976). *Modern factor analysis*. Chicago, IL : University of Chicago. Retrieved from https://books.google.com/books?hl=en&lr=&id=e-vMN68C3M4C&oi=fnd&pg=PR11&dq=modern+factor+analysis+harmon+1976&ots=t5HkFqeR_A&sig=GSM4xTXgtItq0iHd-3Zh8vhhgBw#v=onepage&q=modern factor analysis harmon 1976&f=false

Hunt, W. E., & Hess, R. M. (1968). Surgical risk as related to time of intervention in the repair of intracranial aneurysms. *Journal of Neurosurgery*, *28*(1), 14–20. https://doi.org/10.3171/jns.1968.28.1.0014

Lantigua, H., Ortega-Gutierrez, S., Schmidt, J. M., Lee, K., Badjatia, N., Agarwal, S., . . . Mayer, S. A. (2015). Subarachnoid hemorrhage: Who dies, and why? *Critical Care*, *19*(1), 309. https://doi.org/10.1186/s13054-015-1036-0

Merriman, J. D., Sereika, S. M., Brufsky, A. M., McAuliffe, P. F., McGuire, K. P., Myers, J. S., . . . Bender, C. M. (2017). Trajectories of self-reported cognitive function in postmenopausal women during adjuvant systemic therapy for breast cancer. *Psycho-Oncology*, *26*(1), 44–52. https://doi.org/10.1002/pon.4009

Mohd Zulkifly, M. F., Ghazali, S. E., Che Din, N., & Subramaniam, P. (2016). The influence of demographic, clinical, psychological and functional determinants on post-stroke cognitive impairment at day care stroke center, Malaysia. *Malaysian Journal of Medical Sciences*, *23*(2), 53–64. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/27547115

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research*. Thousand Oaks, CA: Sage Pub.

Price, L., Tulsky, D., Millis, S., & Weiss, L. (2002). Redefining the factor structure of the Wechsler Memory Scale-III: Confirmatory factor analysis with cross-validation. *Journal of Clinical and Experimental Neuropsychology*, *24*(5), 574–585. https://doi.org/10.1076/jcen.24.5.574.1013

Raubenheimer, J. E. (2004). An item selection procedure to maximise scale reliability and validity. *SA Journal of Industrial Psychology*, *30*(4), 59–64. https://doi.org/10.4102/sajip.v30i4.168

Ravnik, J., Starovasnik, B., Sesok, S., Pirtosek, Z., Svigelj, V., Bunc, G., & Bosnjak, R. (2006). Long-term cognitive deficits in patients with good outcomes after aneurysmal subarachnoid hemorrhage from anterior communicating artery. *Croatian Medical Journal*, *47*(2), 253–263. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/16625690

Reitan, R. M. (1955). The relation of the trail making test to organic brain damage. *Journal of Consulting Psychology*, *19*(5), 393–394. https://doi.org/10.1037/h0044509

Richardson-Vejlgaard, R., Dawes, S., Heaton, R. K., & Bell, M. D. (2009). Validity of cognitive complaints in substance-abusing patients and non-clinical controls: The Patient's Assessment of Own Functioning Inventory (PAOFI). *Psychiatry Research*, *169*(1), 70–74. https://doi.org/10.1016/j.psychres.2008.06.018

Rivero Rodríguez, D., Scherle Matamoros, C., Cúe, L. F., Miranda Hernández, J. L., Pernas Sánchez, Y., & Pérez Nellar, J. (2015). Predictor's of mortality in patients with aneurysmal subarachnoid hemorrhage and rebleeding. *Neurology Research International*, *2015*, 545407. https://doi.org/10.1155/2015/545407

Sánchez-Cubillo, I., Periáñez, J. A., Adrover-Roig, D., Rodríguez-Sánchez, J. M., Ríos-Lago, M., Tirapu, J., & Barceló, F. (2009, May 30). Construct validity of the trail making test: Role of task-switching, working memory, inhibition/interference control, and visuomotor abilities. *Journal of the International Neuropsychological Society*, *15*, 438–450. https://doi.org/10.1017/S1355617709090626

Saylor, D., Nakigozi, G., Nakasujja, N., Robertson, K., Gray, R. H., Wawer, M. J., & Sacktor, N. (2017). Peripheral neuropathy in HIV-infected and uninfected patients in Rakai, Uganda. *Neurology*, *89*, 485–491. https://doi.org/10.1212/WNL.0000000000004136

Shukla, D. P. (2017). Outcome and rehabilitation of patients following aneurysmal subarachnoid haemorrhage. *Journal of Neuroanaesthesiology and Critical Care*, *4*(4), 65. https://doi.org/10.4103/2348-0548.199952

Song, M. K., Ward, S. E., Bair, E., Weiner, L. J., Bridgman, J. C., Hladik, G. A., & Gilet, C. A. (2015). Patient-reported cognitive functioning and daily functioning in chronic dialysis patients. *Hemodialysis International*, *19*(1), 90–99. https://doi.org/10.1111/hdi.12202

Srisurapanont, M., Suttajit, S., Eurviriyanukul, K., & Varnado, P. (2017). Discrepancy between objective and subjective cognition in adults with major depressive disorder. *Scientific Reports*, *7*(1), 3901. https://doi.org/10.1038/s41598-017-04353-w

Taheri, Z.., Harirchian, M. H., Ghanaati, H., Khoshnevisan, A., Salamati, P., Miri, M., . . . Rahimi-Movaghar, V. (2015). Comparison of endovascular coiling and surgical clipping for the treatment of intracranial aneurysms: A prospective study. *Iranian Journal of Neurology*, *14*(1), 22–28. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/25874053

Tombaugh, T. N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology*, *14*(2), 167–177. https://doi.org/10.1016/S0887-6177(97)00095-4

Van Dyk, K., Ganz, P. A., Ercoli, L., Petersen, L., & Crespi, C. M. (2016). Measuring cognitive complaints in breast cancer survivors: Psychometric properties of the patient's assessment of own functioning inventory. *Supportive Care in Cancer*, *24*(12), 4939–4949. https://doi.org/10.1007/s00520-016-3352-6

Wechsler, D. (1997). *Wechsler Memory Scale – Third edition manual*. The Psychological Corporation.

Ziemba-Davis, M., Bohnstedt, B. N., Payner, T. D., Leipzig, T. J., Palmer, E., & Cohen-Gadol, A. A. (2014). Incidence, epidemiology, and treatment of aneurysmal subarachnoid hemorrhage in 12 midwest communities. *Journal of Stroke and Cerebrovascular Diseases*, *23*(5), 1073–1082. https://doi.org/10.1016/j.jstrokecerebrovasdis.2013.09.010

Disclosure. The authors have no relevant financial interest or affiliations with any commercial interests related to the subjects discussed within this article.

Acknowledgments. Research reported in this publication was supported by the National Institute of Nursing Research and the National Heart, Lung, and Blood Institute of the National Institutes of Health under Award Numbers NR004339, HL074316, T32NR009759, and F31NR017311. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Funding. The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Foundation for the National Institutes of Health, NR004339.

Ethical Conduct of Research. Institutional Review Board approval at the University of Pittsburgh has been obtained (IRB approval number 021039) and we have adhered to ethical considerations in the protection of all human subjects involved.

Correspondence regarding this article should be directed to Lacey W. Heinsberg, BSN, RN, University of Pittsburgh, Department of Health Promotion and Development, School of Nursing, 440 Victoria Building, 3500 Victoria Street, Pittsburgh, PA 15261. E-mail: law145@pitt.edu

Supplemental Table 1. The Patient Assessment of Own Functioning Inventory Questionnaire

Item	Question
Subscale I: Memory	
1	How often do you forget something that has been told to you within the last day or two?
2	How often do you forget events which have occurred in the last day or two?
3	How often do you forget people whom you met in the last day or two?
4	How often do you forget things you knew a year of more ago?
5	How often do you forget people you knew or met a year or more ago?
6	How often do you lose track of time, or do things either earlier or later than they are usually done or are supposed to be done?
7	How often do you fail to finish something you start because you forgot that you were doing it? (Include such things as forgetting to put out cigarettes, turn off the stove, etc.)
8	How often do you fail to complete a task because you have forgotten how to do one or more aspects of it?
9	How often do you forget things you are supposed to do or have agreed to do (such as putting gas in the car, paying bills, taking care of errands, etc.)?
Subscale II: Language and communication	
10	How often do you have difficulties understanding what is said to you?
11	How often do you have difficulties recognizing or identifying printed words?
12	How often do you have difficulty understanding reading material which at one time you could have understood?
13	When you speak are your words indistinct or improperly pronounced?
13a	How often do people have difficulty understanding what words you are trying to say?
14	How often do you have difficulty thinking of the names of things?
15	How often do you have difficulty thinking of the words (other than names) for what you want to say?
16	When you write things, how often do you have difficulty forming the letters correctly?
17	Do you have more difficulty spelling, or make more errors in spelling, than you used to?

(*Continued*)

Supplemental Table 1. The Patient Assessment of Own Functioning Inventory Questionnaire (*Continued*)

Item	Question
Subscale III: Use of hands	
18	How often do you have difficulty performing tasks with your right hand (including such things as writing, dressing, carrying, lifting, sports, cooking, etc.)?
19	How often do you have difficulty performing tasks with your left hand?
Subscale IV: Sensory-Perceptual	
20	How often do you have difficulty feeling things with your right hand?
21	How often do you have difficulty feeling things with your left hand?
22	Lately, do you have more difficulty than you used to in seeing all of what you are looking at, or all of what is in front of you (in other words, are some areas of your vision less clear or less direct than others)?
Subscale V: Higher level cognitive and intellectual functions	
23	How often do your thoughts seem confused or illogical?
24	How often do you become distracted from what you are doing or saying by insignificant things which at one time you would have been able to ignore?
25	How often do you become confused about (or make a mistake about) where you are?
26	How often do you have difficulty finding your way about?
27	Do you have more difficulty now than you used to in calculating or working with numbers (including managing finances, paying bills, etc.)?
28	Do you have more difficulty now than you used to in planning or organizing activities (example: deciding what to do and how it should be done)?
29	Do you have more difficulty now than you used to in solving problems that come up around the house, at your job, and so on? (In other words, when something new has to be accomplished, or some new difficulty comes up, do you have more trouble figuring out what should be done and how to do it?)
30	Do you have more difficulty now than you used to in following directions to get somewhere?
31	Do you have more difficulty now than you used to in following instructions concerning how to do things?
Instructions. Please answer each of the following questions by filling in the circle that corresponds to the response which most accurately describes the way you have been recently. Choose only one answer for each question.	
Likert scale. 0 = almost never, 1 = very infrequently, 2 = once in a while, 3 = fairly often, 4 = very often, 5 = almost always.	

Note. Supplemental Table 1 includes a list of the PAOFI item questions; PAOFI = Patient Assessment of Own Functioning Inventory.

SUPPLEMENTAL TABLE 2. Past Evaluations of the Psychometric Properties of the Patient Assessment of Own Functioning Inventory

Study	Population	Number of factors identified with Eigenvalues > 1	Percent of extracted variance	Labels for extracted factors
Chelune and Lehman (1986)	Patients referred for neuropsychological evaluation (n = 598); Nonclinical controls (n = 105)	5	57.5%	Memory 1; Memory 2; Sensorimotor; Language and Communication; and Higher Level Cognitive and Intellectual Functioning
Bell et al. (2013)	Postmenopausal women with early stage breast cancer prior to adjuvant therapy (n = 259)	5	54%	Memory 1; Memory 2; Sensorimotor; Language and Communication; and Higher Level Cognitive and Intellectual Functioning
Van Dyk et al. (2016)	Breast cancer survivors after primary surgery, chemotherapy, or radiation (n = 189); Healthy controls (n = 63)	5	60%	Higher Level Cognition; Memory 1; Memory 2; Language Production; Motor/Sensory Perceptual

Note. Supplemental Table 2 summarizes past evaluations of the psychometric properties of the Patient Assessment of Own Functioning Inventory.

Appendix H.1 License Agreement for Manuscript 6

Copyright Clearance Center Marketplace™

Springer Publishing Company, Inc. - License Terms and Conditions

This is a License Agreement between Lacey W. Heinsberg / University of Pittsburgh ("You") and Springer Publishing Company, Inc. ("Publisher") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer Publishing Company, Inc., and the CCC terms and conditions.

All payments must be made in full to CCC.

Order Date	30-Apr-2020	**Type of Use**	Republish in a thesis/dissertation
Order license ID	1031902-1	**Publisher**	Springer Publishing Company
ISSN	1945-7049	**Portion**	Chapter/article

LICENSED CONTENT

Publication Title	Journal of nursing measurement	**Rightsholder**	Springer Publishing Company, Inc.
Date	01/01/1998	**Publication Type**	e-Journal
Language	English	**URL**	http://www.springerpub.com/journal.aspx?jid=1061-3749
Country	United States of America		

REQUEST DETAILS

Portion Type	Chapter/article	**Rights Requested**	Main product and other compilations/derivative products
Page range(s)	1-24, Epub ahead of print	**Distribution**	Worldwide
Total number of pages	24	**Translation**	Original language of publication
Format (select all that apply)	Print, Electronic	**Copies for the disabled?**	Yes
Who will republish the content?	Academic institution	**Minor editing privileges?**	No
Duration of Use	Life of current and all future editions	**Incidental promotional use?**	No
Lifetime Unit Quantity	Up to 250,000	**Currency**	USD

NEW WORK DETAILS

Title	Multi-Omics of the Iron Homeostasis Pathway in Patient Outcomes after Aneurysmal Subarachnoid Hemorrhage	**Institution name**	University of Pittsburgh
		Expected presentation date	2020-05-27
Instructor name	Lacey W. Heinsberg		

ADDITIONAL DETAILS

Order reference number	N/A	**The requesting person / organization to appear on the license**	Lacey W. Heinsberg / University of Pittsburgh

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Psychometric Properties of the Patient Assessment of Own Functioning Inventory Following Aneurysmal Subarachnoid Hemorrhage	**Title of the article/chapter the portion is from**	N/A
		Author of portion(s)	Lacey W. Heinsberg, Catherine M. Bender, Jeffrey M. Rohay, Paula R. Sherwood, Elizabeth A. Crago
Editor of portion(s)	Janice Hinkle		
Volume of serial or monograph	28 (2)	**Issue, if republishing an article from a serial**	N/A
Page or page range of portion	Unknown, currently Epub Ahead of Print	**Publication date of portion**	2020-04-27

CCC Republication Terms and Conditions

1. Description of Service; Defined Terms. This Republication License enables the User to obtain licenses for republication of one or more copyrighted works as described in detail on the relevant Order Confirmation (the "Work(s)"). Copyright Clearance Center, Inc. ("CCC") grants licenses through the Service on behalf of the rightsholder identified on the Order Confirmation (the "Rightsholder"). "Republication", as used herein, generally means the inclusion of a Work, in whole or in part, in a new work or works, also as described on the Order Confirmation. "User", as used herein, means the person or entity making such republication.

2. The terms set forth in the relevant Order Confirmation, and any terms set by the Rightsholder with respect to a particular Work, govern the terms of use of Works in connection with the Service. By using the Service, the person transacting for a republication license on behalf of the User represents and warrants that he/she/it (a) has been duly authorized by the User to accept, and hereby does accept, all such terms and conditions on behalf of User, and (b) shall inform User of all such terms and conditions. In the event such person is a "freelancer" or other third party independent of User and CCC, such party shall be deemed jointly a "User" for purposes of these terms and conditions. In any event, User shall be deemed to have accepted and agreed to all such terms and conditions if User republishes the Work in any fashion.

3. Scope of License; Limitations and Obligations.

    3.1. All Works and all rights therein, including copyright rights, remain the sole and exclusive property of the Rightsholder. The license created by the exchange of an Order Confirmation (and/or any invoice) and payment by User of the full amount set forth on that document includes only those rights expressly set forth in the Order Confirmation and in these terms and conditions, and conveys no other rights in the Work(s) to User. All rights not expressly granted are hereby reserved.

    3.2. General Payment Terms: You may pay by credit card or through an account with us payable at the end of the month. If you and we agree that you may establish a standing account with CCC, then the following terms apply: Remit Payment to: Copyright Clearance Center, 29118 Network Place, Chicago, IL 60673-1291. Payments Due: Invoices are payable upon their delivery to you (or upon our notice to you that they are available to you for downloading). After 30 days, outstanding amounts will be subject to a service charge of 1-1/2% per month or, if less, the maximum rate allowed by applicable law. Unless otherwise specifically set forth in the Order Confirmation or in a separate written agreement signed by CCC, invoices are due and payable on "net 30" terms. While User may exercise the rights licensed immediately upon issuance of the Order Confirmation, the license is automatically revoked and is null and void, as if it had never been

issued, if complete payment for the license is not received on a timely basis either from User directly or through a payment agent, such as a credit card company.

3.3. Unless otherwise provided in the Order Confirmation, any grant of rights to User (i) is "one-time" (including the editions and product family specified in the license), (ii) is non-exclusive and non-transferable and (iii) is subject to any and all limitations and restrictions (such as, but not limited to, limitations on duration of use or circulation) included in the Order Confirmation or invoice and/or in these terms and conditions. Upon completion of the licensed use, User shall either secure a new permission for further use of the Work(s) or immediately cease any new use of the Work(s) and shall render inaccessible (such as by deleting or by removing or severing links or other locators) any further copies of the Work (except for copies printed on paper in accordance with this license and still in User's stock at the end of such period).

3.4. In the event that the material for which a republication license is sought includes third party materials (such as photographs, illustrations, graphs, inserts and similar materials) which are identified in such material as having been used by permission, User is responsible for identifying, and seeking separate licenses (under this Service or otherwise) for, any of such third party materials; without a separate license, such third party materials may not be used.

3.5. Use of proper copyright notice for a Work is required as a condition of any license granted under the Service. Unless otherwise provided in the Order Confirmation, a proper copyright notice will read substantially as follows: "Republished with permission of [Rightsholder's name], from [Work's title, author, volume, edition number and year of copyright]; permission conveyed through Copyright Clearance Center, Inc. " Such notice must be provided in a reasonably legible font size and must be placed either immediately adjacent to the Work as used (for example, as part of a by-line or footnote but not as a separate electronic link) or in the place where substantially all other credits or notices for the new work containing the republished Work are located. Failure to include the required notice results in loss to the Rightsholder and CCC, and the User shall be liable to pay liquidated damages for each such failure equal to twice the use fee specified in the Order Confirmation, in addition to the use fee itself and any other fees and charges specified.

3.6. User may only make alterations to the Work if and as expressly set forth in the Order Confirmation. No Work may be used in any way that is defamatory, violates the rights of third parties (including such third parties' rights of copyright, privacy, publicity, or other tangible or intangible property), or is otherwise illegal, sexually explicit or obscene. In addition, User may not conjoin a Work with any other material that may result in damage to the reputation of the Rightsholder. User agrees to inform CCC if it becomes aware of any infringement of any rights in a Work and to cooperate with any reasonable request of CCC or the Rightsholder in connection therewith.

4. Indemnity. User hereby indemnifies and agrees to defend the Rightsholder and CCC, and their respective employees and directors, against all claims, liability, damages, costs and expenses, including legal fees and expenses, arising out of any use of a Work beyond the scope of the rights granted herein, or any use of a Work which has been altered in any unauthorized way by User, including claims of defamation or infringement of rights of copyright, publicity, privacy or other tangible or intangible property.

5. Limitation of Liability. UNDER NO CIRCUMSTANCES WILL CCC OR THE RIGHTSHOLDER BE LIABLE FOR ANY DIRECT, INDIRECT, CONSEQUENTIAL OR INCIDENTAL DAMAGES (INCLUDING WITHOUT LIMITATION DAMAGES FOR LOSS OF BUSINESS PROFITS OR INFORMATION, OR FOR BUSINESS INTERRUPTION) ARISING OUT OF THE USE OR INABILITY TO USE A WORK, EVEN IF ONE OF THEM HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. In any event, the total liability of the Rightsholder and CCC (including their respective employees and directors) shall not exceed the total amount actually paid by User for this license. User assumes full liability for the actions and omissions of its principals, employees, agents, affiliates, successors and assigns.

6. Limited Warranties. THE WORK(S) AND RIGHT(S) ARE PROVIDED "AS IS". CCC HAS THE RIGHT TO GRANT TO USER THE RIGHTS GRANTED IN THE ORDER CONFIRMATION DOCUMENT. CCC AND THE RIGHTSHOLDER DISCLAIM ALL OTHER WARRANTIES RELATING TO THE WORK(S) AND RIGHT(S), EITHER EXPRESS OR IMPLIED, INCLUDING

WITHOUT LIMITATION IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. ADDITIONAL RIGHTS MAY BE REQUIRED TO USE ILLUSTRATIONS, GRAPHS, PHOTOGRAPHS, ABSTRACTS, INSERTS OR OTHER PORTIONS OF THE WORK (AS OPPOSED TO THE ENTIRE WORK) IN A MANNER CONTEMPLATED BY USER; USER UNDERSTANDS AND AGREES THAT NEITHER CCC NOR THE RIGHTSHOLDER MAY HAVE SUCH ADDITIONAL RIGHTS TO GRANT.

7. Effect of Breach. Any failure by User to pay any amount when due, or any use by User of a Work beyond the scope of the license set forth in the Order Confirmation and/or these terms and conditions, shall be a material breach of the license created by the Order Confirmation and these terms and conditions. Any breach not cured within 30 days of written notice thereof shall result in immediate termination of such license without further notice. Any unauthorized (but licensable) use of a Work that is terminated immediately upon notice thereof may be liquidated by payment of the Rightsholder's ordinary license price therefor; any unauthorized (and unlicensable) use that is not terminated immediately for any reason (including, for example, because materials containing the Work cannot reasonably be recalled) will be subject to all remedies available at law or in equity, but in no event to a payment of less than three times the Rightsholder's ordinary license price for the most closely analogous licensable use plus Rightsholder's and/or CCC's costs and expenses incurred in collecting such payment.

8. Miscellaneous.

    8.1. User acknowledges that CCC may, from time to time, make changes or additions to the Service or to these terms and conditions, and CCC reserves the right to send notice to the User by electronic mail or otherwise for the purposes of notifying User of such changes or additions; provided that any such changes or additions shall not apply to permissions already secured and paid for.

    8.2. Use of User-related information collected through the Service is governed by CCC's privacy policy, available online here:https://marketplace.copyright.com/rs-ui-web/mp/privacy-policy

    8.3. The licensing transaction described in the Order Confirmation is personal to User. Therefore, User may not assign or transfer to any other person (whether a natural person or an organization of any kind) the license created by the Order Confirmation and these terms and conditions or any rights granted hereunder; provided, however, that User may assign such license in its entirety on written notice to CCC in the event of a transfer of all or substantially all of User's rights in the new material which includes the Work(s) licensed under this Service.

    8.4. No amendment or waiver of any terms is binding unless set forth in writing and signed by the parties. The Rightsholder and CCC hereby object to any terms contained in any writing prepared by the User or its principals, employees, agents or affiliates and purporting to govern or otherwise relate to the licensing transaction described in the Order Confirmation, which terms are in any way inconsistent with any terms set forth in the Order Confirmation and/or in these terms and conditions or CCC's standard operating procedures, whether such writing is prepared prior to, simultaneously with or subsequent to the Order Confirmation, and whether such writing appears on a copy of the Order Confirmation or in a separate instrument.

    8.5. The licensing transaction described in the Order Confirmation document shall be governed by and construed under the law of the State of New York, USA, without regard to the principles thereof of conflicts of law. Any case, controversy, suit, action, or proceeding arising out of, in connection with, or related to such licensing transaction shall be brought, at CCC's sole discretion, in any federal or state court located in the County of New York, State of New York, USA, or in any federal or state court whose geographical jurisdiction covers the location of the Rightsholder set forth in the Order Confirmation. The parties expressly submit to the personal jurisdiction and venue of each such federal or state court.If you have any comments or questions about the Service or Copyright Clearance Center, please contact us at 978-750-8400 or send an e-mail to support@copyright.com.

v 1.1

Copyright Clearance Center Marketplace™

Springer Publishing Company, Inc. - License Terms and Conditions

This is a License Agreement between Lacey W. Heinsberg/ University of Pittsburgh ("You") and Springer Publishing Company, Inc. ("Publisher") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer Publishing Company, Inc., and the CCC terms and conditions.

All payments must be made in full to CCC.

Order Date	30-Apr-2020	**Type of Use**	Republish in a thesis/dissertation
Order license ID	1031902-3	**Publisher**	SPRINGER PUBLISHING COMPANY
ISSN	1061-3749	**Portion**	Chapter/article

LICENSED CONTENT

Publication Title	Journal of nursing measurement	**Country**	United States of America
Date	01/01/1993	**Rightsholder**	Springer Publishing Company, Inc.
Language	English	**Publication Type**	Journal

REQUEST DETAILS

Portion Type	Chapter/article	**Rights Requested**	Main product and other compilations/derivative products
Page range(s)	1-24, Epub Ahead of print	**Distribution**	Worldwide
Total number of pages	24	**Translation**	Original language of publication
Format (select all that apply)	Print, Electronic	**Copies for the disabled?**	Yes
Who will republish the content?	Academic institution	**Minor editing privileges?**	No
Duration of Use	Life of current and all future editions	**Incidental promotional use?**	No
Lifetime Unit Quantity	Up to 250,000	**Currency**	USD

NEW WORK DETAILS

Title	Psychometric Properties of the Patient Assessment of Own Functioning Inventory Following Aneurysmal Subarachnoid Hemorrhage	**Institution name**	University of Pittsburgh
		Expected presentation date	2020-05-27
Instructor name	Lacey W Heinsberg		

ADDITIONAL DETAILS

Order reference number	N/A	**The requesting person / organization to appear on the license**	Lacey W. Heinsberg/ University of Pittsburgh

REUSE CONTENT DETAILS

Title, description or numeric reference of the portion(s)	Psychometric Properties of the Patient Assessment of Own Functioning Inventory Following Aneurysmal Subarachnoid Hemorrhage	**Title of the article/chapter the portion is from**	N/A
		Author of portion(s)	Lacey W. Heinsberg, Catherine M. Bender, Jeffrey M. Rohay, Paula R. Sherwood, Elizabeth A. Crago
Editor of portion(s)	Janice Hinkle		
Volume of serial or monograph	28	**Issue, if republishing an article from a serial**	2
Page or page range of portion	1-24, unknown, currently epub ahead of print	**Publication date of portion**	2020-04-27

CCC Republication Terms and Conditions

1. Description of Service; Defined Terms. This Republication License enables the User to obtain licenses for republication of one or more copyrighted works as described in detail on the relevant Order Confirmation (the "Work(s)"). Copyright Clearance Center, Inc. ("CCC") grants licenses through the Service on behalf of the rightsholder identified on the Order Confirmation (the "Rightsholder"). "Republication", as used herein, generally means the inclusion of a Work, in whole or in part, in a new work or works, also as described on the Order Confirmation. "User", as used herein, means the person or entity making such republication.

2. The terms set forth in the relevant Order Confirmation, and any terms set by the Rightsholder with respect to a particular Work, govern the terms of use of Works in connection with the Service. By using the Service, the person transacting for a republication license on behalf of the User represents and warrants that he/she/it (a) has been duly authorized by the User to accept, and hereby does accept, all such terms and conditions on behalf of User, and (b) shall inform User of all such terms and conditions. In the event such person is a "freelancer" or other third party independent of User and CCC, such party shall be deemed jointly a "User" for purposes of these terms and conditions. In any event, User shall be deemed to have accepted and agreed to all such terms and conditions if User republishes the Work in any fashion.

3. Scope of License; Limitations and Obligations.

   3.1. All Works and all rights therein, including copyright rights, remain the sole and exclusive property of the Rightsholder. The license created by the exchange of an Order Confirmation (and/or any invoice) and payment by User of the full amount set forth on that document includes only those rights expressly set forth in the Order Confirmation and in these terms and conditions, and conveys no other rights in the Work(s) to User. All rights not expressly granted are hereby reserved.

   3.2. General Payment Terms: You may pay by credit card or through an account with us payable at the end of the month. If you and we agree that you may establish a standing account with CCC, then the following terms apply: Remit Payment to: Copyright Clearance Center, 29118 Network Place, Chicago, IL 60673-1291. Payments Due: Invoices are payable upon their delivery to you (or upon our notice to you that they are available to you for downloading). After 30 days, outstanding amounts will be subject to a service charge of 1-1/2% per month or, if less, the maximum rate allowed by applicable law. Unless otherwise specifically set forth in the Order Confirmation or in a separate written agreement signed by CCC, invoices are due and payable on "net 30" terms. While User may exercise the rights licensed immediately upon issuance of the Order Confirmation, the license is automatically revoked and is null and void, as if it had never been

issued, if complete payment for the license is not received on a timely basis either from User directly or through a payment agent, such as a credit card company.

3.3. Unless otherwise provided in the Order Confirmation, any grant of rights to User (i) is "one-time" (including the editions and product family specified in the license), (ii) is non-exclusive and non-transferable and (iii) is subject to any and all limitations and restrictions (such as, but not limited to, limitations on duration of use or circulation) included in the Order Confirmation or invoice and/or in these terms and conditions. Upon completion of the licensed use, User shall either secure a new permission for further use of the Work(s) or immediately cease any new use of the Work(s) and shall render inaccessible (such as by deleting or by removing or severing links or other locators) any further copies of the Work (except for copies printed on paper in accordance with this license and still in User's stock at the end of such period).

3.4. In the event that the material for which a republication license is sought includes third party materials (such as photographs, illustrations, graphs, inserts and similar materials) which are identified in such material as having been used by permission, User is responsible for identifying, and seeking separate licenses (under this Service or otherwise) for, any of such third party materials; without a separate license, such third party materials may not be used.

3.5. Use of proper copyright notice for a Work is required as a condition of any license granted under the Service. Unless otherwise provided in the Order Confirmation, a proper copyright notice will read substantially as follows: "Republished with permission of [Rightsholder's name], from [Work's title, author, volume, edition number and year of copyright]; permission conveyed through Copyright Clearance Center, Inc. " Such notice must be provided in a reasonably legible font size and must be placed either immediately adjacent to the Work as used (for example, as part of a by-line or footnote but not as a separate electronic link) or in the place where substantially all other credits or notices for the new work containing the republished Work are located. Failure to include the required notice results in loss to the Rightsholder and CCC, and the User shall be liable to pay liquidated damages for each such failure equal to twice the use fee specified in the Order Confirmation, in addition to the use fee itself and any other fees and charges specified.

3.6. User may only make alterations to the Work if and as expressly set forth in the Order Confirmation. No Work may be used in any way that is defamatory, violates the rights of third parties (including such third parties' rights of copyright, privacy, publicity, or other tangible or intangible property), or is otherwise illegal, sexually explicit or obscene. In addition, User may not conjoin a Work with any other material that may result in damage to the reputation of the Rightsholder. User agrees to inform CCC if it becomes aware of any infringement of any rights in a Work and to cooperate with any reasonable request of CCC or the Rightsholder in connection therewith.

4. Indemnity. User hereby indemnifies and agrees to defend the Rightsholder and CCC, and their respective employees and directors, against all claims, liability, damages, costs and expenses, including legal fees and expenses, arising out of any use of a Work beyond the scope of the rights granted herein, or any use of a Work which has been altered in any unauthorized way by User, including claims of defamation or infringement of rights of copyright, publicity, privacy or other tangible or intangible property.

5. Limitation of Liability. UNDER NO CIRCUMSTANCES WILL CCC OR THE RIGHTSHOLDER BE LIABLE FOR ANY DIRECT, INDIRECT, CONSEQUENTIAL OR INCIDENTAL DAMAGES (INCLUDING WITHOUT LIMITATION DAMAGES FOR LOSS OF BUSINESS PROFITS OR INFORMATION, OR FOR BUSINESS INTERRUPTION) ARISING OUT OF THE USE OR INABILITY TO USE A WORK, EVEN IF ONE OF THEM HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. In any event, the total liability of the Rightsholder and CCC (including their respective employees and directors) shall not exceed the total amount actually paid by User for this license. User assumes full liability for the actions and omissions of its principals, employees, agents, affiliates, successors and assigns.

6. Limited Warranties. THE WORK(S) AND RIGHT(S) ARE PROVIDED "AS IS". CCC HAS THE RIGHT TO GRANT TO USER THE RIGHTS GRANTED IN THE ORDER CONFIRMATION DOCUMENT. CCC AND THE RIGHTSHOLDER DISCLAIM ALL OTHER WARRANTIES RELATING TO THE WORK(S) AND RIGHT(S), EITHER EXPRESS OR IMPLIED, INCLUDING

WITHOUT LIMITATION IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. ADDITIONAL RIGHTS MAY BE REQUIRED TO USE ILLUSTRATIONS, GRAPHS, PHOTOGRAPHS, ABSTRACTS, INSERTS OR OTHER PORTIONS OF THE WORK (AS OPPOSED TO THE ENTIRE WORK) IN A MANNER CONTEMPLATED BY USER; USER UNDERSTANDS AND AGREES THAT NEITHER CCC NOR THE RIGHTSHOLDER MAY HAVE SUCH ADDITIONAL RIGHTS TO GRANT.

7. Effect of Breach. Any failure by User to pay any amount when due, or any use by User of a Work beyond the scope of the license set forth in the Order Confirmation and/or these terms and conditions, shall be a material breach of the license created by the Order Confirmation and these terms and conditions. Any breach not cured within 30 days of written notice thereof shall result in immediate termination of such license without further notice. Any unauthorized (but licensable) use of a Work that is terminated immediately upon notice thereof may be liquidated by payment of the Rightsholder's ordinary license price therefor; any unauthorized (and unlicensable) use that is not terminated immediately for any reason (including, for example, because materials containing the Work cannot reasonably be recalled) will be subject to all remedies available at law or in equity, but in no event to a payment of less than three times the Rightsholder's ordinary license price for the most closely analogous licensable use plus Rightsholder's and/or CCC's costs and expenses incurred in collecting such payment.

8. Miscellaneous.

    8.1. User acknowledges that CCC may, from time to time, make changes or additions to the Service or to these terms and conditions, and CCC reserves the right to send notice to the User by electronic mail or otherwise for the purposes of notifying User of such changes or additions; provided that any such changes or additions shall not apply to permissions already secured and paid for.

    8.2. Use of User-related information collected through the Service is governed by CCC's privacy policy, available online here:https://marketplace.copyright.com/rs-ui-web/mp/privacy-policy

    8.3. The licensing transaction described in the Order Confirmation is personal to User. Therefore, User may not assign or transfer to any other person (whether a natural person or an organization of any kind) the license created by the Order Confirmation and these terms and conditions or any rights granted hereunder; provided, however, that User may assign such license in its entirety on written notice to CCC in the event of a transfer of all or substantially all of User's rights in the new material which includes the Work(s) licensed under this Service.

    8.4. No amendment or waiver of any terms is binding unless set forth in writing and signed by the parties. The Rightsholder and CCC hereby object to any terms contained in any writing prepared by the User or its principals, employees, agents or affiliates and purporting to govern or otherwise relate to the licensing transaction described in the Order Confirmation, which terms are in any way inconsistent with any terms set forth in the Order Confirmation and/or in these terms and conditions or CCC's standard operating procedures, whether such writing is prepared prior to, simultaneously with or subsequent to the Order Confirmation, and whether such writing appears on a copy of the Order Confirmation or in a separate instrument.

    8.5. The licensing transaction described in the Order Confirmation document shall be governed by and construed under the law of the State of New York, USA, without regard to the principles thereof of conflicts of law. Any case, controversy, suit, action, or proceeding arising out of, in connection with, or related to such licensing transaction shall be brought, at CCC's sole discretion, in any federal or state court located in the County of New York, State of New York, USA, or in any federal or state court whose geographical jurisdiction covers the location of the Rightsholder set forth in the Order Confirmation. The parties expressly submit to the personal jurisdiction and venue of each such federal or state court.If you have any comments or questions about the Service or Copyright Clearance Center, please contact us at 978-750-8400 or send an e-mail to support@copyright.com.

v 1.1

Appendix I Manuscript 7: Characterization of Cerebrospinal Fluid DNA Methylation Age During the Acute Recovery Period Following Aneurysmal Subarachnoid Hemorrhage

Characterization of Cerebrospinal Fluid DNA Methylation Age During the Acute Recovery Period Following Aneurysmal Subarachnoid Hemorrhage

Lacey W. Heinsberg, PhD(c), RN[1*], Dongjing Liu, MS[2], John R. Shaffer, PhD,[2,3] Daniel E. Weeks, PhD,[2,4] Yvette P. Conley, PhD, FAAN[1,2]

[1]Department of Health Promotion and Development, School of Nursing, University of Pittsburgh, Pittsburgh, PA, USA

[2]Department of Human Genetics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA

[3]Department of Oral Biology, School of Dental Medicine, University of Pittsburgh, Pittsburgh, PA, USA

[4]Department of Biostatistics, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA

*** Correspondence:**
Lacey W. Heinsberg, PhD(c), RN
law145@pitt.edu

Keywords: Stroke, Epigenomics, Epigenetics, Age acceleration, Group-based trajectory analysis

Abstract

Biological aging may occur at different rates than chronological aging due to genetic, social, and environmental factors. DNA methylation (DNAm) age is thought to be a reliable measure of accelerated biological aging which has been linked to an array of poor health outcomes. Given the importance of chronological age in recovery following aneurysmal subarachnoid hemorrhage (aSAH), a type of stroke, DNAm age may also be an important biomarker of outcomes, further improving predictive models. Cerebrospinal fluid (CSF) is a unique tissue representing the local central nervous system environment post-aSAH. However, the validity of CSF DNAm age is unknown, and it is unclear which epigenetic clock is ideal to compute CSF DNAm age. Further, the stability of DNAm age post-aSAH, specifically, has not been examined and may offer insight into the biology of recovery post-aSAH. Therefore, the purpose of this study was to characterize CSF DNAm age over 14 days post-aSAH using three epigenetic clocks. Genome-wide DNAm data were available for two tissues: (1) CSF for N=273 participants with serial sampling over 14 days post-aSAH (N=850 samples) and (2) blood for a subset of n=72 participants at one time point post-aSAH. DNAm age was calculated using the Horvath, Hannum, and Levine epigenetic clocks. 'Age acceleration' was computed as the residuals of DNAm age regressed on chronological age both with and without correcting for cell-type heterogeneity (CTH). Using scatterplots, Pearson correlations, and group-based trajectory analysis, we examined the relationships between CSF DNAm age and chronological age, the concordance between DNAm age calculated from CSF versus blood, and the stability (i.e., trajectories) of CSF DNAm age acceleration over time during recovery from aSAH. We observed moderate to strong correlations between CSF DNAm age and chronological age (R = 0.66 [Levine] to R=0.86 [Horvath]), moderate to strong correlations between DNAm age in CSF versus blood

(R=0.69 [Levine] to R=0.87 [Horvath]), and stable CSF age acceleration trajectories over 14 days post-aSAH once controlling for CTH. Although correlated, CSF DNAm age differs from blood DNAm age. While CSF DNAm age was generally stable post-aSAH, the Horvath clock appears to be the least dependent on CTH.

1 Introduction

Across the spectrum of neurological injury populations, identifying therapeutic targets of intervention to improve patient outcomes has been a challenge. The aneurysmal subarachnoid hemorrhage (aSAH) population is no exception. After aSAH, it is consistently observed that younger patients do better following injury (Lantigua et al., 2015) underscoring the importance of chronological age as a predictor of outcomes. However, given within-individual variability such as genomic (e.g., genetic variability), social (e.g., socioeconomic status), and environmental (e.g., chemical exposure) factors, it is thought that 'biological aging' for many individuals happens at different rates and that chronological age is often a flawed surrogate measure of this phenomenon. For this reason, a substantial amount of work has been dedicated to identifying molecular biomarkers of aging. One of the most promising thus far is 'DNA methylation age' (DNAm age) which can be computed from 'epigenetic clocks' and is suggested to be applicable across the lifespan and in all sources of biological tissues (Horvath & Raj, 2018).

Several epigenetic clocks have been proposed over the last decade including the Horvath (Horvath, 2013, 2015), Hannum (Hannum et al., 2013), and Levine (Levine et al., 2018) clocks which use DNAm data from 353, 71, and 513 CpG sites, respectively. While the Horvath and Hannum clocks were developed to estimate chronological age, the Levine clock expanded on this to estimate a biological age metric known as 'phenotypic age', which was based not only on chronological age, but also other biological factors predictive of mortality (e.g., albumin, creatinine) (Levine et al., 2018). Another important difference between the clocks is that the Horvath clock was developed using training datasets with DNAm data generated from many biological tissues sources whereas the Hannum and Levine clocks were developed using only DNAm data generated from blood (though they have been subsequently examined and validated in other tissues). Despite these differences, DNAm ages estimated by all three epigenetic clocks are strongly correlated with chronological age. Individuals with an older DNAm age than their chronological age are said to have 'age acceleration' which has been associated with many negative health outcomes such as cancer (Kresovich et al., 2019), Parkinson's disease (Horvath & Ritz, 2015), cardiovascular disease (Roetker, Pankow, Bressler, Morrison, & Boerwinkle, 2018), and all-cause mortality (Marioni et al., 2015).

Although DNAm age has been examined in a wide range of biological tissues (e.g., blood, kidney, liver, tumor, brain (Horvath & Raj, 2018)), it has not been examined in cerebrospinal fluid (CSF). Thus, the validity and potential utility of DNAm age computed using this tissue is not understood which is an important gap in our knowledge. In many cases of neurologic injury where CSF is drained as part of clinical management, including aSAH, this tissue may offer insight into the local environment of the central nervous system. Following aSAH specifically, there are secondary injuries occurring at the cellular level that may impact biological aging. Therefore, trajectories of age acceleration over time during recovery from neurologic injury may offer insight into (1) the stability of DNAm age in acute pathological conditions such as aSAH and (2) the biology of recovery post-aSAH.

Therefore, the purpose of this longitudinal, observational study was to characterize CSF DNAm age over 14 days following aSAH. As part of this characterization, we wanted to better understand the

relationships between CSF DNAm age and chronological age, the concordance between DNAm age calculated using CSF versus blood, the stability (i.e., trajectories) of CSF age acceleration during recovery from acute neuronal injury (i.e., aSAH), and the correlations and performance of three epigenetic clock methods (Horvath (Horvath, 2013, 2015), Hannum (Hannum et al., 2013), and Levine (Levine et al., 2018)) in CSF both with and without considering the effects of cell type proportions. We hypothesized that DNAm age would be strongly correlated with chronological age but would change over time during recovery from acute neuronal injury in response to rapidly changing cell type proportions, DNAm age measured in CSF versus blood would be strongly correlated, and age acceleration trajectories based on the three epigenetic clocks would be similar.

2 Methods

2.1 Setting and Sample

This study was approved by the Institutional Review Board of the University of Pittsburgh. Upon informed consent, participants for this study were prospectively recruited from UPMC Presbyterian Neurovascular Intensive Care Unit in Pittsburgh, Pennsylvania between 2009 and 2013 as previously described (Arockiaraj et al., 2020). In brief, participants were included in this study if they were diagnosed with subarachnoid hemorrhage caused by an aneurysm rupture, were at least 18 years of age, and had no history of debilitating neurological disorder. As part of this larger study, participants were followed over 14 days post-aSAH in the hospital.

2.2 DNA Methylation Data Collection

For this study, DNA was extracted from two biological tissue sources including (1) CSF (for all study participants [N=279] with serial sampling over 14 days after aSAH) and (2) blood (for a subset of study participants [n=88] at one time point after aSAH). DNA was extracted from CSF using the Qiamp Midi kit (Qiagen, Valencia, CA, USA) and from blood using a simple salting out procedure (Miller, Dykes, & Polesky, 1988). All DNA was stored in 1X TE buffer at 4°C until DNA methylation data collection. As described, CSF samples from ventricular drains placed as standard of care were selected for targeted post-injury days of 1, 4, 7, 10, and 13 (± 1 day) (Arockiaraj et al., 2020). Genome-wide DNA methylation data were generated using the Infinium Human Methylation450 BeadChip and scanned using the Illumina iSCAN (Illumina, Incorporated, San Diego, CA, USA) at the Center for Inherited Disease Research using laboratory quality control (QC) procedures described (Arockiaraj et al., 2020). Raw genome-wide DNA methylation data were analyzed using Genome Studio Software (Illumina, Incorporated, San Diego, CA, USA). Our data cleaning and QC pipeline included removal of poorly performing samples, probes, and outliers (Arockiaraj et al., 2020) as well as functional normalization using the 'funtooNorm' function implemented in the funtooNorm package (Oros Klein et al., 2016).

2.3 DNA Methylation Age

DNAm age was calculated using three epigenetic clocks (Horvath (Horvath, 2013, 2015), Hannum (Hannum et al., 2013), and Levine (Levine et al., 2018)). These methods use linear functions and clock-specific probes and coefficients to compute DNAm age as shown in Equation 1:

$$DNAmAge = m_0 + m_1\beta_1 + m_2\beta_2 + \cdots + m_n\beta_n \quad (1)$$

where $DNAmAge$ is the predicted DNA methylation age for a given individual, m is a clock-specific coefficient corresponding to a clock-specific probe, β is the DNA methylation measurement, a beta

value as measured on a 0 to 1 scale, for a clock-specific probe within a given individual, and $ (is a clock-specific model intercept. It should be noted that the Horvath method also uses an age transformation function as described (Horvath, 2013, 2015) and shown in the Supplementary Material. These calculations were performed using a modified function from the wateRmelon package (Pidsley et al., 2013) in R (Team, 2018) (wateRmelon:agep). The wateRmelon package supplies both Horvath and Hannum coefficients for use with the 'agep' function and we modified this function to also compute Levine DNAm age as described in detail in the Supplementary Material.

DNAm age was computed using both CSF DNAm data and blood DNAm data. To allow for comparability between tissues, only clock-specific probes available in both CSF and blood were used in our analysis. Following implementation of the QC pipeline described above, for the Horvath, Hannum, and Levine epigenetic clocks, we were missing DNA methylation data for 1, 3, and 5 probes, respectively as detailed in the Supplementary Table 1. Following calculation of DNAm age, CSF data were reshaped into five cross-sectional time points including time 1 (days 0 to 2 post-aSAH), time 2 (days 3 to 5 post-aSAH), time 3 (days 6 to 8 post-aSAH), time 4 (days 9 to 11 post-aSAH), and time 5 (days 12 to 14 post-aSAH). The vast majority of the blood samples available were collected at time 1 (days 0 to 2 post-aSAH) so blood samples collected outside of this cross-sectional time point (n=16) were excluded from further analyses.

2.4 Age Acceleration

For each of the three epigenetic clocks, we computed age acceleration defined as the residuals of DNAm age regressed on chronological age within each cross-sectional time point. We computed age acceleration both with and without adjustment for cell type heterogeneity (CTH) because cell type proportions are known to vary with age (Chen et al., 2016). Because reference-based methods do not exist for CSF, all CTH data were computed using the Houseman's reference-free method which estimates proportions of cell types for each sample (Houseman, Molitor, & Marsit, 2014). In adjusting for CTH, we computed age acceleration as the residuals of DNAm age regressed on chronological age and cell type proportions. Because our cell-type variable was a proportioned phenotype which added up to one, we excluded the cell type with the lowest amount of variation within our study sample to minimize confounding the results.

2.5 Participant characteristic data

Participant data were extracted from the medical record and included standard demographic data (e.g., age, sex, and self-reported race), body mass index (BMI) and smoking history (given associations between these factors and DNA methylation levels (Li, Wong, Bui, Nguyen, Joo, Stone, Dite, Dugué, et al., 2018; Li, Wong, Bui, Nguyen, Joo, Stone, Dite, Giles, et al., 2018)), and Fisher grade, which is a clinical variable measuring the initial extent of aSAH injury (Fisher, Kistler, & Davis, 1980).

2.6 Statistical Analysis

Statistical analyses were conducted using R statistical computing environment (version 3.6.0) (Team, 2018) and SAS (version 9.4, SAS Institute Incorporated, Cary, NC, USA). Demographic and clinical characteristics of our sample were computed using standard descriptive statistics. DNAm age was compared with chronological age using scatterplots and Pearson correlations. This was done for CSF observations (available over 14 days following aSAH) and for blood (available on days 0, 1, and 2 post-aSAH). Next, for participants with both CSF and blood samples available, we compared the correlation between DNAm age calculated from CSF versus blood at time point 1 (days 0 to 2 post-

aSAH). Finally, we evaluated correlations among DNAm age calculated using the three epigenetic clocks both directly and as partial correlations controlling for age as implemented in the ppcor::pcor.test function in R. Following calculation of age acceleration as described above, we compared age acceleration computed using the Horvath (Horvath, 2013, 2015), Hannum (Hannum et al., 2013), and Levine (Levine et al., 2018) epigenetic clocks both with and without adjusting for CTH using Pearson correlations and sina plots.

Next, we examined age acceleration over time during recovery from aSAH using group-based trajectory analysis (GBTA) implemented using the Proc TRAJ macro in SAS (Jones & Nagin, 2007; Jones, Nagin, & Roeder, 2001). GBTA was performed through iterative modeling (i.e., comparison of models with varying group numbers and shapes) to infer distinct trajectory groups. Bayesian Information Criterion (BIC) was used as our primary indicator of model fit, with a larger BIC indicating a better model fit (Jones & Nagin, 2007; Jones et al., 2001). Following selection of a best-fitting model, we performed a posterior QC check of the model using several model-fit indices including ensuring (1) the average posterior probability of group assignment was at least 0.7, (2) the odds of correct classification was greater than 5, and (3) the estimated group assignment percentages were approximately equal to the observed group assignment percentages (Jones & Nagin, 2007; Jones et al., 2001). Finally, we used one way analysis of variance and chi-square tests to understand how participant characteristics differed between inferred trajectory groups and linear regression to understand bivariate associations between age acceleration metrics independent of trajectory groups and participant characteristics.

3 Results

3.1 Sample Characteristics

Our final post-QC sample sizes consisted of n=273 participants. All participants had CSF DNAm data at up to five cross-sectional time points over 14 days post-aSAH (for a total of n=850 observations). Of the overall sample, n=72 participants also had blood DNAm data available at cross-sectional time point 1 (days 0 to 2 post-aSAH). Sample characteristics are presented in Table 1. Our overall sample (n=273) had a mean (±SD) age of 52.9 (±11.1) years, was 68.5% female, and 87.2% White with Fisher grades of 2, 3, or 4 accounting for 29.7%, 49.5%, and 20.9% of the sample, respectively. Mean BMI was 28.1 (±7.2) kg/m^2 and 53.8% of participants were active smokers. We observed similar statistics in the subset of participants with both CSF and blood DNAm data available on days 0 to 2 post-aSAH (n=72). This group had a mean age of 53.0 (±11.5) years, was 69.4% female, and 84.7% White with Fisher grades of 2, 3, or 4 accounting for 27.8%, 51.4%, and 20.8% of the sample, respectively. In this smaller subset of participants with both CSF and blood DNAm data, mean BMI was 28.8 (±8.8) kg/m^2 and 51.4% of participants were active smokers. The sample characteristics observed were comparable to statistics observed in the general aSAH population (Zacharia et al., 2010).

3.2 Correlation Between DNAm Age and Chronological Age

Across all CSF samples (n=273 at up to five time points over 14 days post-aSAH), DNAm age was moderately to strongly correlated with chronological age in the Horvath (R = 0.86, p < 2.2E-16), Hannum (R = 0.82, p < 2.2E-16), and Levine (R = 0.66, p < 2.2E-16) clocks as shown in Figure 1. The relationship between DNAm age and chronological age was similar for all three clocks as shown in the overlay plot in Supplementary Figure 1. However, the scatterplots in Figure 1 highlight between participant variation and that the relationship between CSF DNAm age and chronological age differs as a function of chronological age. For CSF DNAm age computed using all three clocks,

we observe higher DNAm age than expected in younger participants and lower DNAm age than expected in older participants (all $p < 0.00001$). Within cross-sectional time points, we observe similar correlations between chronological age and DNAm age in CSF over time as shown in the scatterplots presented in Supplementary Figures 2, 3, and 4.

Within the subset of participants for which blood was available (n=72 on days 0 to 2 post-aSAH), chronological age was highly correlated with DNAm age in the Horvath ($R = 0.88$, $p < 2.2E-16$), Hannum ($R = 0.92$, $p < 2.2E-16$), and Levine ($R = 0.83$, $p < 2.2E-16$) clocks as shown in Figure 2. The relationship between DNAm age and chronological age was similar for all three clocks as shown in the overlay plot in Supplementary Figure 5. While correlations between chronological age and DNAm age were stronger in blood compared with CSF, for all three clocks, the relationship between DNAm age and chronological age differed as a function of chronological age (all $p < 0.00001$).

3.3 Correlation Between DNAm Age in CSF versus Blood

Within the subset of participants for which both CSF and blood DNAm data were available (n=72 on days 0 to 2 post-aSAH), we observed moderate to strong correlations between DNAm age measured in CSF versus blood as shown in Figure 3. Specifically, we observed correlations between DNAm age in CSF compared with blood using the Horvath ($R = 0.67$, $p < 2.2E-16$), Hannum ($R = 0.84$, $p < 2.2E-16$), and Levine ($R = 0.69$, $p = 1.2E-10$) clocks. Partial correlations comparing DNAm age in CSF versus blood given chronological age are shown in Supplementary Figure 6.

3.4 Correlation Between the Horvath, Hannum, and Levine Epigenetic Clocks

Next, we wanted to understand the correlations between age acceleration estimated using the Horvath, Hannum, and Levine epigenetic clocks. First, within cross-sectional time points available for CSF, we compared the methods using both direct correlations and partial correlations controlling for age. In comparing the Horvath and Hannum methods, we observed direct correlations of $R = 0.84$ to 0.89 and partial correlations controlling for age of $R_{age} = 0.39$ to 0.56 as shown in Supplementary Figure 7. In comparing the Horvath and Levine methods, we observed direct correlations of $R = 0.67$ – 0.76 and partial correlations controlling for age of $R_{age} = 0.14$ to 0.37 as shown in Supplementary Figure 8. Finally, in comparing the Hannum and Levine methods, we observed direct correlations of $R = 0.72$ to 0.80 and partial correlations of $R_{age} = 0.22$ to 0.58 as shown in Supplementary Figure 9. We observed similar relationships between the three epigenetic clocks in blood after aSAH as shown in Supplementary Figure 10.

Next, we wanted to understand the correlations between age acceleration with and without adjustment for CTH computed in all methods and available biospecimens (i.e., CSF and blood). As discussed above, age acceleration was computed by regressing DNAm age on chronological age both with and without adjusting for CTH resulting in a measure of age acceleration that has no association with chronological age as shown in the exemplar plots in Supplementary Figures 11 (CSF) and 12 (blood). We used the subset of participants with both CSF and blood available (n=72 at cross-sectional time point 1 on days 0 to 2 post-aSAH) to compare the methods as shown in the correlation heatmap presented in Figure 4. We observed a range of correlations as small as $R = 0.08$ (between blood age acceleration computed using the Hannum clock and CSF age acceleration computed using the Horvath clock) to as large as $R = 0.90$ (between blood age acceleration unadjusted for CTH computed using the Levine clock and blood age acceleration adjusted for CTH computed using the Levine clock).

Lastly, in the sina plot depicting the distribution and density of the age acceleration data shown in Figure 5, we observed that the CSF age acceleration computed using the Levine clock had the widest range of values while blood age acceleration adjusted for CTH computed using the Hannum clock had the narrowest range of values.

3.5 Trajectories of Age Acceleration Over Time in CSF

Finally, in an effort to understand the stability (i.e., trajectories) of CSF age acceleration over time during recovery from acute neurologic injury, we performed GBTA using the age acceleration data (both with and without adjusting for CTH) computed from the three epigenetic clocks. Inferred age acceleration trajectory groups for the Horvath and Hannum clocks are plotted in Figure 6. As discussed in more detail below, the trajectory models for the Levine clock did not pass posterior model QC, so are not presented in Figure 6. For age acceleration data computed using the Horvath clock, both unadjusted and adjusted for CTH, four distinct, flat trajectory groups (Groups 1 through 4) were inferred, suggesting that Horvath DNAm age acceleration does not change over time during recovery from neurologic injury (Figure 6, Horvath and Horvath + CTH). All unadjusted and CTH-adjusted model selection parameters including the BICs computed from iterative model testing as well as posterior model QC indices are presented in Supplementary Tables 2a, 2b, 3a, and 3b.

For age acceleration data unadjusted for CTH computed using the Hannum clock, we again inferred four distinct trajectory groups. However, while the two groups with the highest age acceleration (Groups 3 and 4) did not change over time, we observed a slow increase in age acceleration in Group 2 and an increase followed by a return to baseline in Group 1 (Figure 6, Hannum). However, when we controlled for CTH in the calculation of age acceleration, this temporal variation was washed out resulting in four flat trajectory groups with no change over time (Figure 6, Hannum + CTH). All model selection parameters including the BICs computed from iterative model testing as well as posterior model QC indices are presented in Supplementary Tables 4a, 4b, 5a, and 5b. It should be noted that the plots in Figure 6 depict inferred trajectory groups and are not directly comparable because group membership changes after adjustment for CTH as shown in Tables 2 (Horvath) and 3 (Hannum) (e.g., in Figure 6, Hannum, Group 1 has only 8 participants while in Figure 6, Hannum + CTH, Group 1 has 23 participants).

GBTA plots for the Levine clock are presented in Supplementary Figure 13. In performing GBTA for the Levine age acceleration data unadjusted for CTH, we again inferred four trajectory groups with the two highest groups (Groups 3 and 4) steady over time with no change, a third group steadily increasing over time (Group 2), and a fourth group decreasing and then returning to baseline (Group 1). Interestingly, when we controlled for CTH, this temporal variability was washed out as described with the Hannum method above, resulting in four trajectory groups with no change over time. However, neither the trajectory model unadjusted for CTH nor the trajectory model adjusted for CTH passed QC procedures due to inadequate odds of correct classification of the middle groups. In other words, while we were confident in group participant assignment in the highest and lowest DNA methylation groups (Groups 4 and 1, respectively), participant assignment couldn't be distinguished for the middle groups. All model selection parameters including BICs computed from iterative model testing as well as posterior model QC indices are presented in Supplementary Tables 6a, 6b, 7a, and 7b.

Next, we computed participant characteristics for identified trajectory groups for the trajectory models that passed posterior QC (Horvath and Hannum models) and present the results in Tables 2 and 3, respectively. For both clocks, we noticed a difference between sex with a decreasing

proportion of females as age acceleration increases, though this was only statistically significant in the Hannum clock ($p < 0.0001$). This is particularly notable in the age acceleration trajectory groups unadjusted for CTH computed using the Hannum clock. The group with the lowest age acceleration (Group 1) was 93.3% female while the group with the highest age acceleration (Group 4) was only 16.7% female. We observed no other differences in participant characteristics by trajectory group.

Lastly, we wanted to understand if DNAm age acceleration was associated with participant characteristics independent of inferred trajectory groups as shown in the bivariate results in Table 4. We observed associations between sex and Horvath CSF DNAm age acceleration ($p = 0.02$), Hannum CSF DNAm age acceleration ($p < 0.0001$), and Hannum CSF DNAm age acceleration controlling for CTH ($p = 0.0001$). We also observed associations between race and Hannum DNAm age acceleration in blood ($p = 0.04$), Levine CSF DNAm age acceleration ($p = 0.03$), and Levine CSF DNAm age acceleration controlling for CTH ($p = 0.003$). Finally, we observed an association between smoking and Levine CSF DNAm age acceleration controlling for CTH ($p = 0.003$).

4 Discussion

This study is the first to characterize CSF DNAm age over the first 14 days post-aSAH. While we observed similarities between the tissues and epigenetic clocks used in this study, some important differences should be considered. First, the relationship between both CSF and blood DNAm age and chronological age differed substantially by chronological age as shown in Figure 1 (CSF) and Figure 2 (blood). In all three clocks and both tissues examined, the chronologically youngest participants had substantially older DNAm age, whereas for the chronologically older participants the opposite was observed (all $p < 0.00001$). This finding underscored the fact that although, in general, DNAm age is highly correlated with chronological age across the lifespan, the relationship changes as a function of chronological age. This phenomenon was recently well-described in detail elsewhere (El Khoury et al., 2019) and our observations in this study further support this. Several measurements of age acceleration are reported in the literature. Most commonly, we observed (1) delta age, defined as the difference between DNAm age and chronological age and (2) age acceleration, defined as the residuals of DNAm age regressed on age (often with the addition of covariates such as CTH) and the methodology used in the current study. Initially, we performed our analyses using delta age and then realized that there was a systematic difference in delta age based on chronological age as discussed above. The age acceleration method used in the current study is a measure that is completely independent of age as shown in Supplementary Figures 11 and 12 and is therefore less likely to lead to confounding in studies using DNAm age.

In general, we observed strong correlations between DNAm age measured in CSF versus blood as shown in Figure 3. The correlations were the strongest using the Horvath ($R = 0.87$) and Hannum ($R = 0.84$) clocks, but for nearly all participants, Hannum CSF DNAm age observations were 'biologically younger' compared with blood. Given chronological age, however, we observed much lower partial correlations for the Horvath, Hannum, and Levine clocks with R_{age} values of 0.46, 0.32, and 0.44, respectively (Supplementary Figure 6). A notable finding of this study is that for both CSF and blood, we observed only moderate partial correlations between the three epigenetic clocks after adjusting for chronological age as shown in Supplementary Figures 7 through 10. There is little overlap between the probes used to compute DNAm age in each clock. Therefore, each clock captures different phenomena related to aging. In addition, while the relationship between chronological age and DNAm age was generally steady in CSF over the five cross-sectional time points examined, we did observe some possible time-dependent changes as shown in Supplementary Figures 2 through 4, though statistical significance was not reached. Moreover, as mentioned earlier,

the Levine clock was developed to predict a 'phenotypic age' rather than chronological age (Levine et al., 2018), so is therefore less directly comparable to the Horvath and Hannum methods.

A strength of this study is the availability of serial CSF samples available over 14 days allowing us to examine the stability of CSF age acceleration over time during recovery from acute neurological injury. As shown in Figure 6, our GBTA results indicate that for both the Horvath and Hannum clocks, we observe four trajectory groups with no change over time once adjusting for CTH. Although the Hannum clock unadjusted for CTH did show some change over time, an interesting observation was that this variability was no longer observed after we controlled for CTH resulting in nearly identical trajectory groups between the Horvath and Hannum methods. Although the Levine trajectory models shown in Supplementary Figure 13 did not pass posterior QC, we see results nearly identical to the Hannum clock. In the trajectory model unadjusted for CTH, we observe temporal changes over time, while in the trajectory model adjusted for CTH, we observe very similar trajectories to Horvath and Hannum clocks with four groups showing no change over time. Overall, we observed that the Horvath clock does not appear to be strongly influenced by CTH and may be the most appropriate methodology for use with CSF, even in cases where CTH data do not exist.

It is important to note that outside of CTH, we did not control for the influence of participant characteristics (e.g., sex, race, smoking, or BMI) in our calculation of age acceleration. As shown in Tables 2 and 3, sex was associated with inferred trajectory group assignment. Most notably were the between-group sex differences observed for the Hannum clock where we observe a steadily decreasing percentage of females as age acceleration increases (Table 3). This is consistent with existing literature suggesting that overall men have higher DNAm age than women (Horvath et al., 2016). This was confirmed by examining associations between participant characteristics and ungrouped age acceleration metrics (Table 4), where we again observed associations between sex and CSF DNAm age acceleration both with and without adjustment for CTH computed using the Hannum clock. An unexpected finding in this study was that the group with the lowest age acceleration (Group 1) had the highest percentage of participants with a history of smoking for both the Horvath and Hannum methods (Tables 2 and 3), which contrasts with existing literature suggesting smoking increases age acceleration (Yang et al., 2019). This was not a statistically significant finding, however, and in our ungrouped analyses (i.e., independent of trajectory groups, Table 4), we observed higher Levine CSF DNAm age acceleration in participants with a history of smoking compared with participants who had never smoked.

Although this study has many strengths, there are some limitations that should be acknowledged. First, as described above, there are several age acceleration metrics (e.g., delta age vs. age acceleration) available in the literature. Because aSAH outcomes are strongly associated with chronological age, and that delta age (the difference between DNAm age and chronological age) can vary as a function of chronological age, we chose to use the residual method. Importantly, our definition of age acceleration is an attribute that depends on the group of samples that went into the regression. So, if DNAm was measured on one participant alone, the residual-based definition of age acceleration could not be used to compute age acceleration without embedding results in a group from which the regression line could be computed. As such, the residual definition is not specific to the individual, but is also an attribute of the group. On a related note, an additional limitation of the regression method is the potential sensitivity to outlying DNAm values. We did not score adjust any outlying DNAm values which may have been influential in our regression results when computing age acceleration. The influence of outliers should be carefully considered in studies examining associations between age acceleration and health related outcomes. Clinically, delta age may be of more interest than the residual definition of age acceleration because it could be calculated for only

one participant. However, given (1) the systematic difference in delta age as a function of chronological age, (2) the wide range of chronological ages in our study (i.e., 25 to 75 years), and (3) that outcomes following aSAH are strongly associated with chronological age, we felt that the regression definition of age acceleration was most appropriate for this population.

Next, it should be noted that all blood samples were included on separate plates from CSF samples, so blood plate batch effects are intractably confounded. Despite our rigorous QC pipeline, this is something that we could not correct for and should be considered when interpreting correlations between age acceleration metrics computed in CSF versus blood. In addition, only a subset of 72 of our participants had blood DNAm data available making our comparisons involving blood DNAm age or age acceleration quite small. Likewise, for participants with blood available, the DNAm data were only collected at one cross-sectional time point (on days 0 to 2 post-aSAH) which prevented us from comparing the trajectories of blood age acceleration over time during recovery from aSAH with CSF. Finally, this study only used three epigenetic clocks to characterize CSF DNAm age. In the future, it may be important to expand comparisons of CSF and blood DNAm age and examine the utility of other epigenetic clocks.

5 Conclusion

We observed moderate to strong correlations between CSF DNAm age and chronological age that varied as a function of chronological age, strong correlations between CSF DNAm age and blood DNAm age, and a wide range of correlations between CSF and blood age acceleration using the three clocks and considering the effects of CTH. In general, once adjusting for CTH, CSF age acceleration was stable over time during recovery from aSAH. Compared with the Hannum and Levine methods, the Horvath clock appears to be less influenced by cell type in CSF.

6 Conflict of Interest

Aside from the research grant funding that supported this project, the authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

7 Author Contributions

YPC is the principal investigator of the parent study for this project and is responsible for acquisition of the data. LWH and YPC conceived and designed the study. LWH, DL, JRS, and DEW performed the statistical analysis. All authors contributed to the interpretation of the data and results. LWH drafted the initial manuscript and all authors reviewed, critically revised, and approved the final manuscript. All authors agree to be accountable for all aspects of the work.

8 Funding

Research reported in this publication was supported by the National Institute of Nursing Research of the National Institutes of Health under Award Numbers R01NR013610, F31NR017311, and T32NR009759. The content is solely the responsibility of the authors and does not represent the official views of the National Institutes of Health.

9 Acknowledgments

We thank our funding sources as well as the research participants and their families for making this work possible.

10 Supplementary Material

Supplementary Material is attached as two files (.docx and .xslx).

11 Data Availability Statement

The datasets used in this study can be accessed through dbGAP: phs001990.v1.p1.

12 Figures

Figures have been submitted separately per the author instructions and titles and captions have been provided here.

Figure 1. Chronological age versus DNAm age in CSF on days 0 to 14 post-aSAH using the Horvath, Hannum, and Levine epigenetic clocks. (A) Horvath (B) Hannum (C) Levine; Sample size, n=273 at up to 5 time points (N=850 observations over 14 days post-aSAH); dashed line, x=y; solid line, predicted model fit; DNAm, DNA methylation; CSF, cerebrospinal fluid; aSAH, aneurysmal subarachnoid hemorrhage; R, correlation computed using Pearson method

Figure 2. Chronological age versus DNAm age in blood at Time 1 (days 0 to 2) post-aSAH using the Horvath, Hannum, and Levine epigenetic clocks. (A) Horvath (B) Hannum (C) Levine; Sample size, n=72 with both CSF and blood DNA methylation data at cross-sectional time point 1 (days 0 to 2 post-aSAH); dashed line, x=y; solid line, predicted model fit; DNAm, DNA methylation; aSAH, aneurysmal subarachnoid hemorrhage; R, correlation computed using Pearson method

Figure 3. DNAm age in CSF versus blood at Time 1 (days 0 to 2) post-aSAH using the Horvath, Hannum, and Levine epigenetic clocks. (A) Horvath (B) Hannum (C) Levine; Sample size, n=72 with both CSF and blood DNA methylation data at cross-sectional time point 1 (days 0 to 2 post-aSAH); dashed line, x=y; solid line, predicted model fit; DNAm, DNA methylation; CSF, cerebrospinal fluid; aSAH, aneurysmal subarachnoid hemorrhage; R, correlation computed using Pearson method

Figure 4. Correlation heatmap of unadjusted and CTH-adjusted age acceleration at Time 1 (days 0 to 2) post-aSAH computed in CSF and blood using the Horvath, Hannum, and Levine epigenetic clocks. Sample size, n=72 with both CSF and blood DNA methylation data at cross-sectional time point 1 (days 0 to 2 post-aSAH); CSF, cerebrospinal fluid; aSAH, aneurysmal subarachnoid hemorrhage; CTH, cell-type heterogeneity; all values presented are R values indicating age acceleration correlation computed using Pearson method

Figure 5. Sina plots of unadjusted and CTH-adjusted age acceleration at Time 1 (days 0 to 2) post-aSAH computed in CSF and blood using the Horvath, Hannum, and Levine epigenetic clocks. Sample size, n=72 with both CSF and blood DNA methylation data at cross-sectional time point 1 (days 0 to 2 post-aSAH); CSF, cerebrospinal fluid; aSAH, aneurysmal subarachnoid hemorrhage; CTH, cell-type heterogeneity

Figure 6. Trajectory plots of unadjusted and CTH-adjusted CSF age acceleration from Horvath and Hannum epigenetic clocks. Note that the above plots portray inferred trajectory groups and are not directly comparable as the group membership changes between plots as shown in Table 2 (Horvath) and Table 3 (Hannum); n=273 at up to five time points (N=850 observations over 14 days post-aSAH condensed to 5 cross-sectional time points); times listed correspond to cross-sectional time points (Time 1 [Days 0 to 2]; Time 2 [Days 3 to 5]; Time 3 [Days 6 to 8]; Time 4 [Days 9 to 11]; Time 5 [Days 12 to 14]); 95% Confidence Intervals shown are for estimated trajectories and not observed trajectories; CSF, cerebrospinal fluid; CTH, cell-type heterogeneity aSAH, aneurysmal subarachnoid hemorrhage; note, the Levine trajectory models did not pass posterior quality control and are presented in the Supplementary Material

13 Tables

Table 1. Sample Characteristics

Variable	Overall sample[a] (n=273)	Sample subset with both CSF and blood (n=72)
Age, mean (SD)	52.9 (11.1)	53.0 (11.5)
Sex, female, n (%)	187 (68.5)	50 (69.4)
Race, White, n (%)	238 (87.2)	61 (84.7)
Fisher, n (%)		
2	81 (29.7)	20 (27.8)
3	135 (49.5)	37 (51.4)
4	57 (20.9)	15 (20.8)
Smoking, n (%)		
No	88 (32.2)	28 (38.8)
Yes	147 (53.8)	37 (51.4)
Social	3 (1.1)	2 (2.8)
Quit	31 (11.4)	3 (4.2)
Unknown	(1.5)	2 (2.8)
BMI, mean (SD)	28.1 (7.2)	28.8 (8.8)

[a]All participants in this study had longitudinal CSF samples available over 14 days post-aSAH; CSF, cerebrospinal fluid; SD, standard deviation; BMI, body mass index

Table 2. Summary of Horvath Cerebrospinal Fluid DNA Methylation Age Acceleration Trajectory Group Characteristics

	Horvath (CSF)						Horvath (CSF) + CTH				
Variable	Group 1	Group 2	Group 3	Group 4	*p*		Group 1	Group 2	Group 3	Group 4	p
Group membership, n(%)	8 (2.9)	115 (42.1)	112 (41.0)	38 (13.9)			23 (8.4)	135 (49.5)	95 (34.8)	20 (7.3)	
Age, mean (SD)	48.5 (14.7)	53.2 (11.7)	55.0 (10.6)	57.5 (13)	0.87[b]		48.7 (11.6)	54.0 (11.8)	52.7 (10.0)	51.7 (9.5)	0.86[b]
DNAmAge, CSF, mean (SD)	46.5 (7.1)	55.1 (6.5)	62.5 (7.5)	62.0 (5.4)	<0.0001[b]		50.7 (6.5)	56.3 (6.5)	58.9 (5.8)	62.2 (6.1)	<0.0001[b]
Age Acceleration, CSF, mean (SD)	-1.3 (2.5)	-1.3 (1.3)	1.2 (1.1)	5.0 (1.7)	<0.0001[b]		-4.3 (1.9)	-1.4 (2.3)	1.9 (1.7)	5.7 (1.9)	<0.0001[b]
DNAmAge, Blood, mean (SD)	42.1 (9.2)	56.3 (8.6)	56.7 (7.0)	59.4 (5.1)	0.068[a]		48.6 (7.0)	56.6 (7.9)	58.9 (6.8)	59.7 (3.2)	<0.0001[b]
Age Acceleration, Blood, mean (SD)	-1.1 (5.0)	-2.3 (3.8)	1.1 (2.3)	3.7 (1.7)	<0.0001[b]		-4.5 (2.8)	-1.2 (2.2)	2.3 (2.7)	4.6 (2.3)	<0.0001[b]
Sex, female, (%)	6 (75.0)	87 (75.7)	71 (63.4)	23 (60.5)	0.15[c]		19 (82.6)	97 (71.9)	56 (58.9)	15 (75.0)	0.06[c]
Race, white, n(%)	6 (75.0)	105 (91.3)	95 (84.8)	32 (84.2)	0.30[c]		19 (82.6)	121 (89.6)	82 (86.3)	16 (80.0)	0.54[c]
Fisher											
2	1 (12.5)	28 (24.3)	41 (36.6)	11 (28.9)			5 (21.7)	40 (29.6)	31 (32.6)	5 (25.0)	
3	6 (75.0)	59 (51.3)	51 (45.5)	19 (50.0)	0.36[c]		12 (52.2)	65 (48.1)	46 (48.4)	12 (60.0)	0.88[c]
4	1 (12.5)	28 (24.3)	20 (17.9)	8 (21.1)			6 (26.1)	30 (22.2)	18 (18.9)	3 (15.0)	
Smoking history											
No	2 (25.0)	38 (33.0)	35 (31.2)	13 (34.2)	0.95[c]		6 (26.1)	47 (34.8)	27 (28.4)	8 (40.0)	0.56[c]
Yes[a]	6 (75.0)	76 (66.1)	75 (67.0)	24 (63.2)			16 (69.6)	87 (64.5)	67 (70.5)	11 (55.0)	
Unknown	0	1 (0.9)	2 (1.8)	1 (2.6)			1 (4.3)	1 (0.7)	1 (1.1)	1 (5.0)	
BMI, mean (SD)	26.5 (5.2)	29.3 (8.8)	27.1 (5.8)	27.8 (5.3)	0.17[b]		27.4 (6.8)	28.8 (7.9)	27.1 (6.3)	28.4 (6.6)	0.49[b]

Table corresponds to Figure 6 (Horvath and Horvath + CTH trajectory plots); CTH, cell-type heterogeneity; SD, standard deviation; CSF, cerebrospinal fluid; BMI, body mass index; [a]Smoking history, yes (current, social, and past smokers), no (never smoked); [b]One way ANOVA; [c]Chi-square

Table 3. Summary of Hannum Cerebrospinal Fluid DNA Methylation Age Acceleration Trajectory Group Characteristics

	Hannum (CSF)					Hannum (CSF) + CTH				
Variable	**Group 1**	**Group 2**	**Group 3**	**Group 4**	***p***	**Group 1**	**Group 2**	**Group 3**	**Group 4**	***p***
Group membership, n(%)	15 (5.5)	121 (44.3)	131 (48.0)	6 (2.2)		10 (3.7)	102 (37.4)	124 (45.4)	37 (13.6)	
Age, mean (SD)	54.5 (9.9)	53.5 (11.2)	52.0 (11.3)	56.7 (7.5)	0.41	53.5 (11.7)	53.6 (11.2)	51.6 (10.7)	55.0 (11.7)	0.94[b]
DNAmAge, CSF, mean (SD)	55.0 (5.9)	56.8 (6.4)	54.5 (7.1)	62.8 (3.8)	<0.0001	55.7 (5.7)	56.9 (6.9)	56.7 (6.4)	60.1 (7.2)	<0.0001[b]
Age Acceleration, CSF, mean (SD)	3.0 (3.5)	-0.7 (2.6)	0.8 (3.3)	3.6 (1.8)	<0.0001	-1.8 (4.1)	-0.7 (2.8)	0.1 (3.2)	1.8 (3.0)	<0.0001[b]
DNAmAge, Blood, mean (SD)	54.5 (5.1)	57.5 (8.3)	55.7 (7.8)	53.0 (6.2)	0.79	53.5 (6.6)	55.4 (8.7)	57.1 (6.9)	60.1 (7.8)	0.04[b]
Age Acceleration, Blood, mean (SD)	-0.7 (2.5)	-0.3 (4.0)	0.4 (3.3)	-0.4 (2.8)	0.0002	-0.6 (2.6)	-1.7 (3.4)	1.4 (2.9)	2.5 (3.9)	<0.0001[b]
Sex, female, n(%)	14 (93.3)	100 (82.6)	72 (55.0)	1 (16.7)	<0.0001[c]	9 (90.0)	88 (86.3)	75 (60.5)	15 (40.5)	<0.0001[c]
Race, White, n(%)	11 (73.3)	104 (86.0)	118 (90.1)	5 (83.3)	0.28[c]	8 (80.0)	83 (81.4)	113 (91.1)	34 (91.9)	0.11[c]
Fisher										
2	5 (33.3)	34 (28.1)	42 (32.1)	1 (16.7)		2 (20.0)	27 (26.5)	40 (32.3)	12 (32.4)	
3	8 (53.3)	58 (47.9)	66 (50.4)	3 (50.0)	0.62[c]	6 (60.0)	52 (51.0)	59 (47.6)	18 (48.6)	0.95[c]
4	2 (13.3)	29 (24.0)	23 (17.6)	2 (33.3)		2 (20.0)	23 (22.5)	25 (20.2)	7 (18.9)	
Smoking history										
No	3 (0.2)	39 (32.2)	44 (33.6)	2 (33.3)	0.75[c]	3 (0.3)	29 (28.4)	42 (33.9)	14 (37.8)	0.68[c]
Yes[a]	12 (0.8)	79 (65.3)	85 (64.9)	4 (66.7)		6 (0.6)	73 (71.6)	79 (63.7)	23 (62.2)	
Unknown	0	3 (2.5)	2 (1.5)	0		1 (0.1)	0	3 (2.4)	0	
BMI, mean (SD)	26.3 (8.7)	28.1 (7.6)	28.5 (6.8)	26.8 (1.8)	0.48[c]	25.4 (3.0)	28.3 (8.5)	28.1 (6.6)	28.6 (6.0)	0.54[b]

Table corresponds to Figure 6 (Hannum and Hannum + CTH trajectory plots); CTH, cell-type heterogeneity; SD, standard deviation; CSF, cerebrospinal fluid; BMI, body mass index; [a] Smoking history, yes (current, social, and past smokers), no (never smoked); [b] One way ANOVA; [c] Chi-square

Table 4. Bivariate Associations Between Participant Characteristics and DNA Methylation Age Acceleration Metrics

DNA Methylation Age Acceleration Metric	Sex[b] Est (SE)	95% CI	*p*	Race[c] Est (SE)	95% CI	*p*
Horvath (CSF)	-0.96 (0.41)	-1.78 to -0.15	0.02[a]	0.31 (0.58)	-0.84 to 1.45	0.60
Horvath (CSF) + CTH	-0.63 (0.50)	-1.62 to 0.37	0.22	1.10 (0.67)	-0.22 to 2.42	0.10
Horvath (Blood)	-0.24 (0.93)	-2.07 to 1.61	0.80	-0.36 (1.23)	-2.82 to 2.10	0.77
Horvath (Blood) + CTH	-0.30 (0.73)	-1.76 to 1.15	0.96	0.34 (0.97)	-1.60 to 2.27	0.73
Hannum (CSF)	-2.71 (0.44)	-3.65 to -1.92	< 0.00001[a]	-0.92 (0.65)	-2.20 to 0.36	0.16
Hannum (CSF) + CTH	-2.18 (0.55)	-3.27 to -1.10	0.0001[a]	-0.06 (0.77)	-1.58 to 1.46	0.94
Hannum (Blood)	-0.69 (0.80)	-2.29 to 0.91	0.39	-2.23 (1.04)	-4.30 to -0.15	0.04[a]
Hannum (Blood) + CTH	-0.53 (0.58)	-1.69 to 0.63	0.37	-0.21 (0.78)	-1.76 to 1.35	0.79
Levine (CSF)	-0.30 (0.61)	-1.50 to 0.90	0.62	1.81 (0.84)	0.84 to 2.22	0.03[a]
Levine (CSF) + CTH	0.99 (0.75)	-0.49 to 2.47	0.19	2.94 (0.97)	1.02 to 4.87	0.003[a]
Levine (Blood)	-0.11 (1.05)	-2.2 to 1.99	0.92	-1.16 (1.39)	-3.93 to 1.62	0.41
Levine (Blood) + CTH	0.12 (0.95)	-1.77 to 2.01	0.90	0.17 (1.26)	-2.34 to 2.69	0.89

DNA Methylation Age Acceleration Metric	Smoking history[d] Est (SE)	95% CI	*p*	BMI Est (SE)	95% CI	*p*
Horvath (CSF)	-0.07 (0.41)	-0.89 to 0.74	0.86	-0.03 (0.03)	-0.09 to 0.02	0.22
Horvath (CSF) + CTH	0.24 (0.49)	-0.74 to 1.21	0.63	0.01 (0.03)	-0.05 to 0.06	0.86
Horvath (Blood)	-0.86 (0.88)	-0.62 to 0.90	0.33	0.02 (0.05)	-0.08 to 0.11	0.74
Horvath (Blood) + CTH	-0.83 (0.68)	-2.19 to 0.54	0.23	0.02 (0.04)	-0.05 to 0.10	0.53
Hannum (CSF)	-0.57 (0.22)	-1.49 to 0.35	0.22	0.02 (0.03)	-0.04 to 0.08	0.52
Hannum (CSF) + CTH	0.17 (0.56)	-0.94 to 1.27	0.77	0.03 (0.03)	-0.04 to 0.10	0.35
Hannum (Blood)	-1.11 (0.76)	-2.62 to 0.40	0.15	0.02 (0.04)	-0.07 to 0.10	0.69
Hannum (Blood) + CTH	0.23 (0.56)	-1.34 to 0.89	0.69	0.02 (0.03)	-0.05 to 0.08	0.61
Levine (CSF)	0.84 (0.61)	-0.36 to 2.04	0.17	0.01 (0.04)	-0.08 to 0.09	0.89
Levine (CSF) + CTH	2.09 (0.70)	-0.70 to 3.48	0.003[a]	-0.03 (0.04)	-0.11 to 0.06	0.56
Levine (Blood)	-0.06 (0.99)	-2.05 to 1.93	0.96	0.04 (0.06)	-0.07 to 0.15	0.51
Levine (Blood) + CTH	0.44 (0.50)	-1.34 to 2.23	0.62	0.03 (0.05)	-0.07 to 0.13	0.55

Results of age acceleration metrics regressed on participant characteristics; Est (SE), Unstandardized estimate and standard error; CI, Confidence Interval; [a]Significance based on an alpha of 0.05; [b]Sex, Female versus male; [c]Race, Non-White vs. White; [d]Smoking history, yes (current, social, and past smokers) versus no (never smoked)

14 References

Arockiaraj, A. I., Liu, D., Shaffer, J. R., Koleck, T. A., Crago, E. A., Weeks, D. E., & Conley, Y. P. (2020). Methylation Data Processing Protocol & Comparison of Blood and Cerebral Spinal Fluid Following Aneurysmal Subarachnoid Hemorrhage. *bioRxiv*. https://doi.org/10.1101/2020.03.24.005264

Chen, B. H., Marioni, R. E., Colicino, E., Peters, M. J., Ward-Caviness, C. K., Tsai, P. C., … Horvath, S. (2016). DNA methylation-based measures of biological age: Meta-analysis predicting time to death. *Aging*, *8*(9), 1844–1865. https://doi.org/10.18632/aging.101020

El Khoury, L. Y., Gorrie-Stone, T., Smart, M., Hughes, A., Bao, Y., Andrayas, A., … Schalkwyk, L. C. (2019). Systematic underestimation of the epigenetic clock and age acceleration in older subjects. *Genome Biology*, *20*(1), 283. https://doi.org/10.1186/s13059-019-1810-4

Fisher, C. M., Kistler, J. P., & Davis, J. M. (1980). Relation of cerebral vasospasm to subarachnoid hemorrhage visualized by computerized tomographic scanning. *Neurosurgery*, *6*(1), 1–9. https://doi.org/10.1227/00006123-198001000-00001

Hannum, G., Guinney, J., Zhao, L., Zhang, L., Hughes, G., Sadda, S., … Zhang, K. (2013). Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. *Molecular Cell*, *49*(2), 359–367. https://doi.org/10.1016/j.molcel.2012.10.016

Horvath, S. (2013). DNA methylation age of human tissues and cell types. *Genome Biology*, *14*(10), R115. https://doi.org/10.1186/gb-2013-14-10-r115

Horvath, S. (2015). Erratum to: DNA methylation age of human tissues and cell types. *Genome Biology*, *16*(1), 96. https://doi.org/10.1186/s13059-015-0649-6

Horvath, S., Gurven, M., Levine, M. E., Trumble, B. C., Kaplan, H., Allayee, H., … Assimes, T. L. (2016). An epigenetic clock analysis of race/ethnicity, sex, and coronary heart disease. *Genome Biology*, *17*(1), 171. https://doi.org/10.1186/s13059-016-1030-0

Horvath, S., & Raj, K. (2018, June 11). DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nature Reviews Genetics*. Nature Publishing Group. https://doi.org/10.1038/s41576-018-0004-3

Horvath, S., & Ritz, B. R. (2015). Increased epigenetic age and granulocyte counts in the blood of Parkinson's disease patients. *Aging*, *7*(12), 1130–42. https://doi.org/10.18632/aging.100859

Houseman, E. A., Molitor, J., & Marsit, C. J. (2014). Reference-free cell mixture adjustments in analysis of DNA methylation data. *Bioinformatics*, *30*(10), 1431–1439. https://doi.org/10.1093/bioinformatics/btu029

Jones, B. L., & Nagin, D. S. (2007). Advances in group-based trajectory modeling and an SAS procedure for estimating them. *Sociological Methods and Research*, *35*(4), 542–571. https://doi.org/10.1177/0049124106292364

Jones, B. L., Nagin, D. S., & Roeder, K. (2001). A SAS procedure based on mixture models for estimating developmental trajectories. *Sociological Methods and Research*, *29*(3), 374–393.

https://doi.org/10.1177/0049124101029003005

Kresovich, J. K., Xu, Z., O'Brien, K. M., Weinberg, C. R., Sandler, D. P., & Taylor, J. A. (2019). Methylation-based biological age and breast cancer risk. *Journal of the National Cancer Institute*, *111*(10), 1051–1058. https://doi.org/10.1093/jnci/djz020

Lantigua, H., Ortega-Gutierrez, S., Schmidt, J. M., Lee, K., Badjatia, N., Agarwal, S., … Mayer, S. A. (2015). Subarachnoid hemorrhage: who dies, and why? *Critical Care*, *19*(1), 309. https://doi.org/10.1186/s13054-015-1036-0

Levine, M. E., Lu, A. T., Quach, A., Chen, B. H., Assimes, T. L., Bandinelli, S., … Horvath, S. (2018). An epigenetic biomarker of aging for lifespan and healthspan. *Aging*, *10*(4), 573–591. https://doi.org/10.18632/aging.101414

Li, S., Wong, E. M., Bui, M., Nguyen, T. L., Joo, J. H. E., Stone, J., … Hopper, J. L. (2018). Causal effect of smoking on DNA methylation in peripheral blood: A twin and family study. *Clinical Epigenetics*, *10*(1), 18. https://doi.org/10.1186/s13148-018-0452-9

Li, S., Wong, E. M., Bui, M., Nguyen, T. L., Joo, J. H. E., Stone, J., … Hopper, J. L. (2018, November 21). Inference about causation between body mass index and DNA methylation in blood from a twin family study. *International Journal of Obesity*, pp. 1–10. Cold Spring Harbor Laboratory. https://doi.org/10.1038/s41366-018-0153-4

Marioni, R. E., Shah, S., McRae, A. F., Chen, B. H., Colicino, E., Harris, S. E., … Deary, I. J. (2015). DNA methylation age of blood predicts all-cause mortality in later life. *Genome Biology*, *16*(1), 25. https://doi.org/10.1186/s13059-015-0584-6

Miller, S. A., Dykes, D. D., & Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Research*, *16*(3), 1215. https://doi.org/10.1093/nar/16.3.1215

Oros Klein, K., Grinek, S., Bernatsky, S., Bouchard, L., Ciampi, A., Colmegna, I., … Greenwood, C. M. T. (2016). FuntooNorm: An R package for normalization of DNA methylation data when there are multiple cell or tissue types. *Bioinformatics*, *32*(4), 593–595. https://doi.org/10.1093/bioinformatics/btv615

Pidsley, R., Y Wong, C. C., Volta, M., Lunnon, K., Mill, J., & Schalkwyk, L. C. (2013). A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics*, *14*(1), 293. https://doi.org/10.1186/1471-2164-14-293

Roetker, N. S., Pankow, J. S., Bressler, J., Morrison, A. C., & Boerwinkle, E. (2018). Prospective Study of Epigenetic Age Acceleration and Incidence of Cardiovascular Disease Outcomes in the ARIC Study (Atherosclerosis Risk in Communities). *Circulation. Genomic and Precision Medicine*, *11*(3), e001937. https://doi.org/10.1161/CIRCGEN.117.001937

Team, R. C. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Yang, Y., Gao, X., Just, A. C., Colicino, E., Wang, C., Coull, B. A., … Baccarelli, A. A. (2019). Smoking-related DNA methylation is associated with DNA methylation phenotypic age

598 acceleration: The veterans affairs normative aging study. *International Journal of*
599 *Environmental Research and Public Health*, *16*(13). https://doi.org/10.3390/ijerph16132356

600 Zacharia, B. E., Hickman, Z. L., Grobelny, B. T., DeRosa, P., Kotchetkov, I., Ducruet, A. F., &
601 Connolly, E. S. (2010). Epidemiology of aneurysmal subarachnoid hemorrhage. *Neurosurgery*
602 *Clinics of North America*, *21*(2), 221–33. https://doi.org/10.1016/j.nec.2009.10.002

603

Figure 1.JPEG

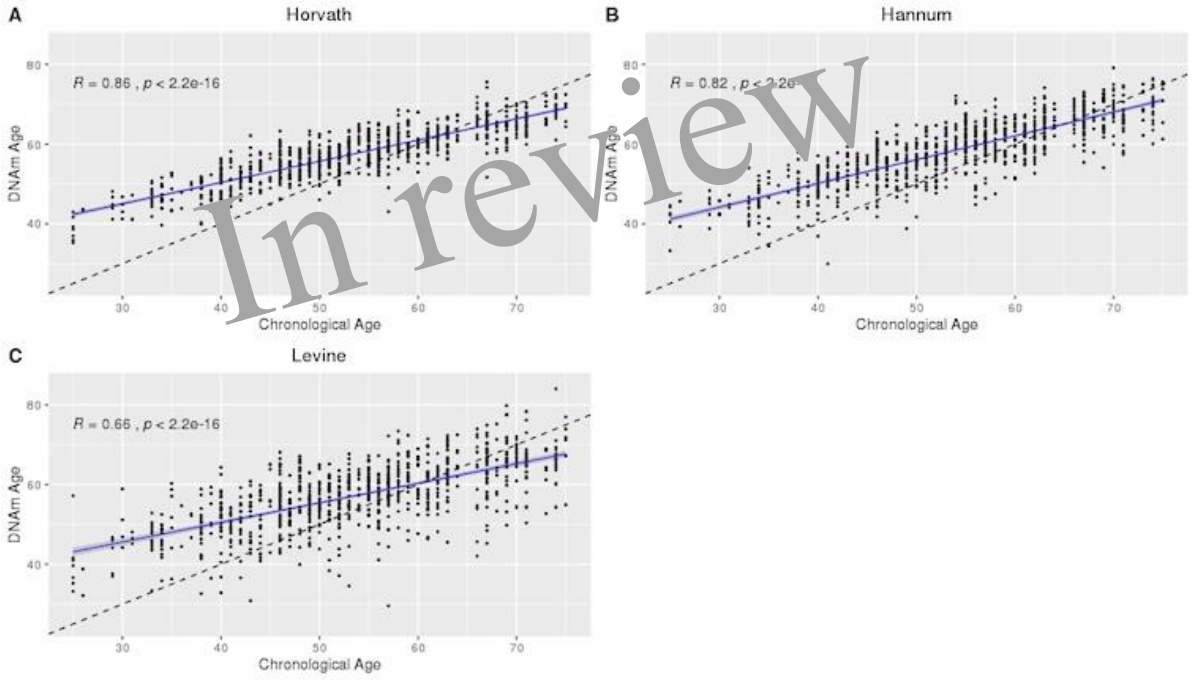


Figure 2.JPEG

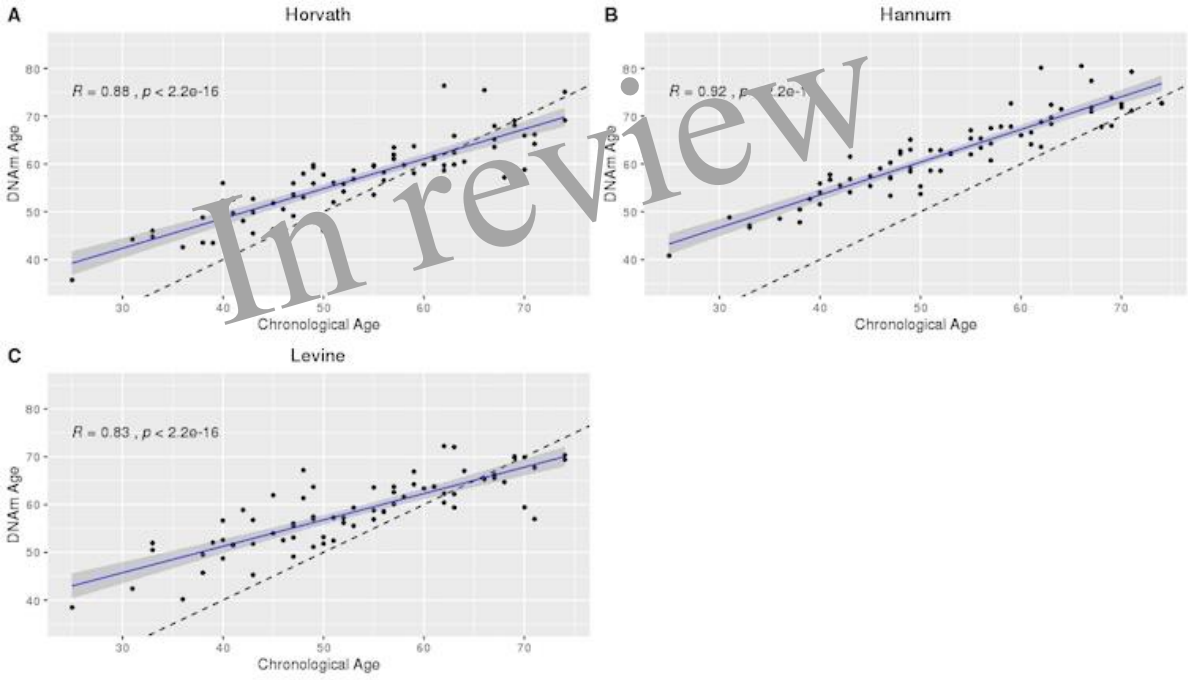


Figure 3.JPEG

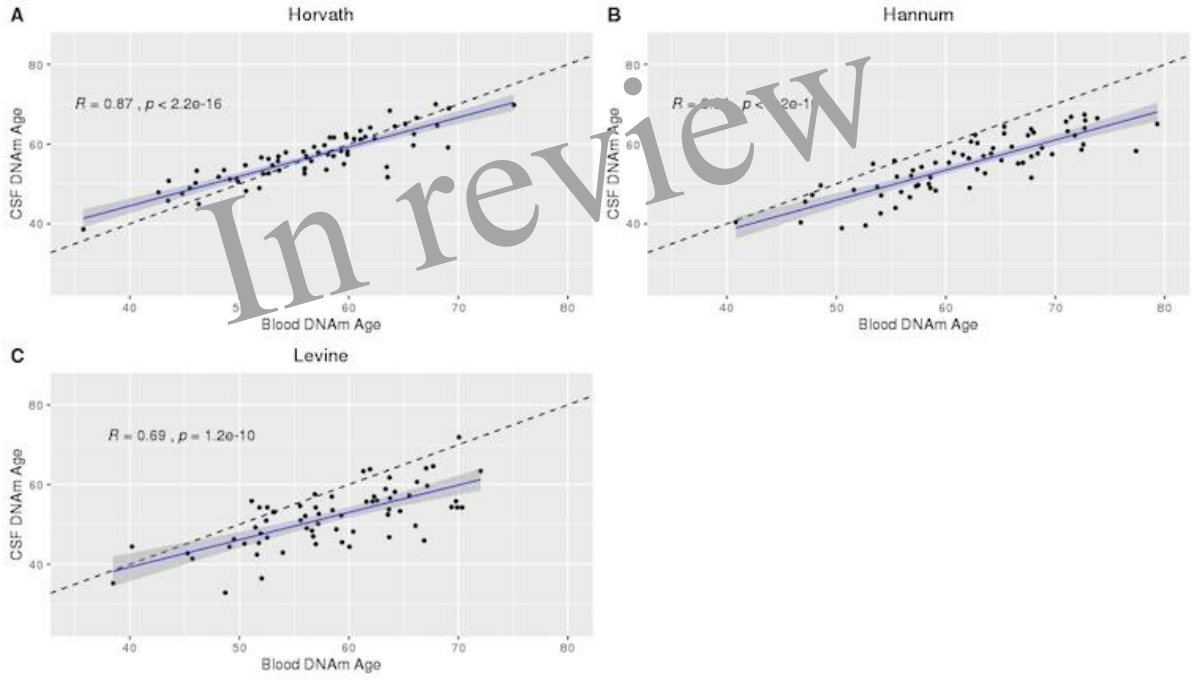


Figure 4.JPEG

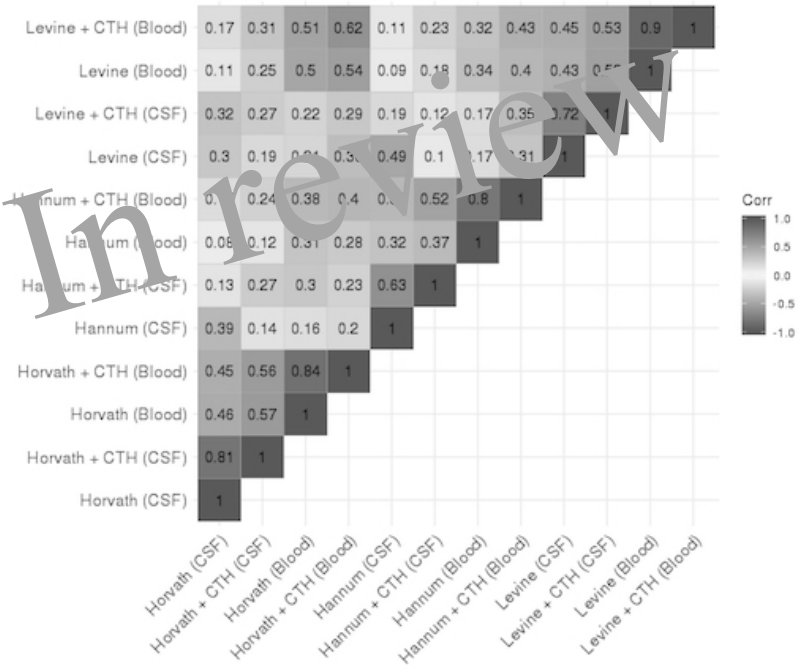


Figure 5.JPEG

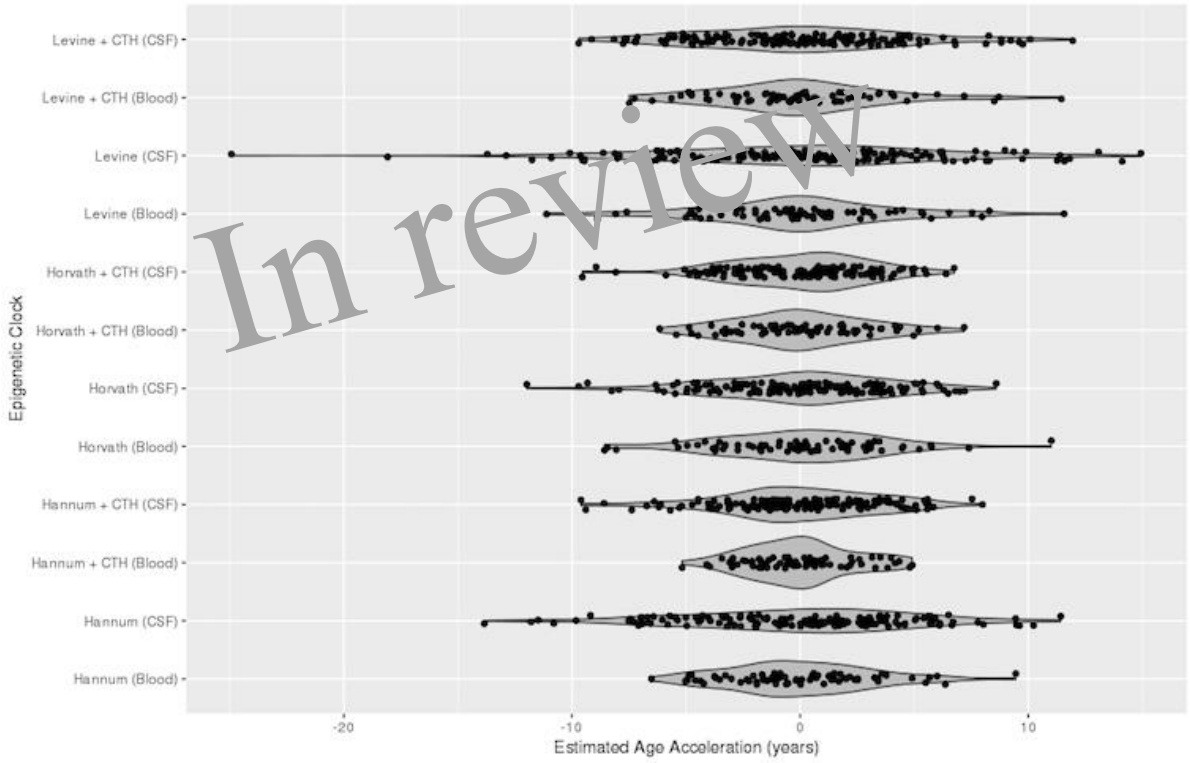
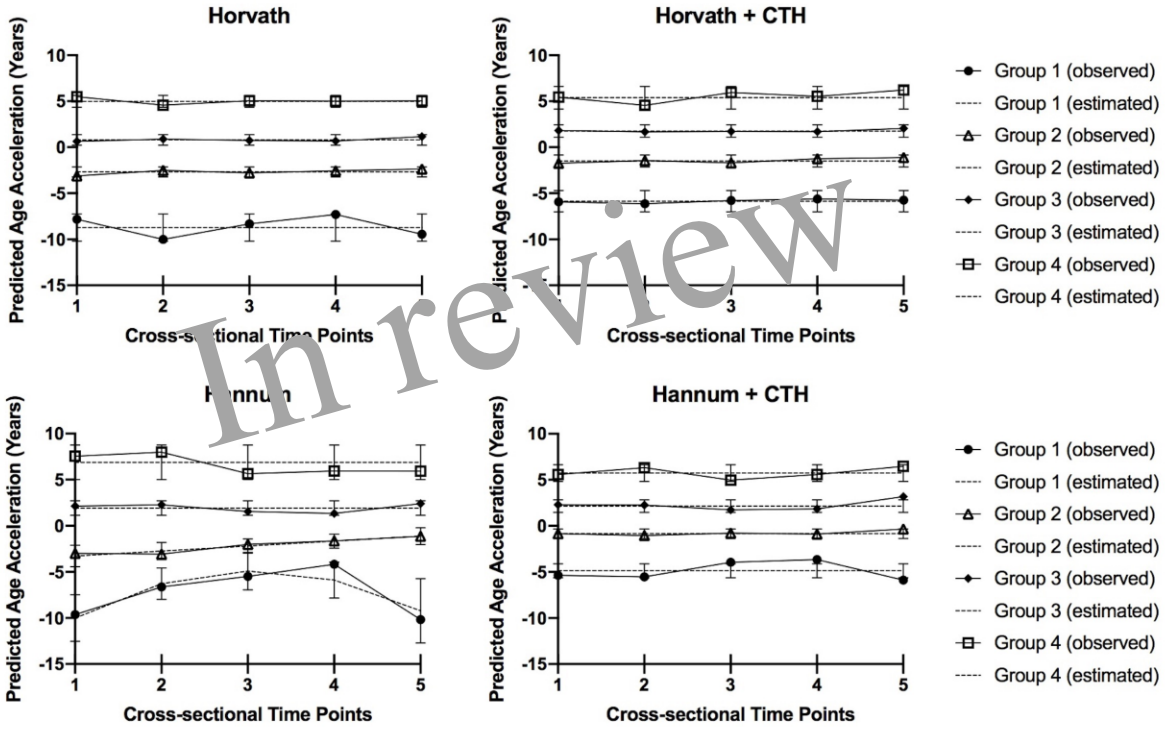


Figure 6.JPEG



Appendix J Institutional Review Board Approval

University of Pittsburgh
Institutional Review Board

Human Research Protection Office
3500 Fifth Avenue, Suite 106
Pittsburgh, PA 15213
Tel (412) 383-1480
www.hrpo.pitt.edu

APPROVAL OF SUBMISSION (Expedited)

Date:	December 9, 2019
IRB:	STUDY19100368
PI:	Lacey Heinsberg
Title:	Multi-omics of the Iron Homeostasis Pathway in Patient Outcomes Following aSAH
Funding:	Name: National Institute of Nursing Research, Funding Source ID: F31NR017311

The Institutional Review Board reviewed and approved the above referenced study. The study may begin as outlined in the University of Pittsburgh approved application and documents.

Approval Documentation

Review type:	Continuing review
Approval Date:	12/9/2019
Expiration Date:	12/8/2020

As the Principal Investigator, you are responsible for the conduct of the research and to ensure accurate documentation, protocol compliance, reporting of possibly study-related adverse events and unanticipated problems involving risk to participants or others. The HRPO Reportable Events policy, Chapter 17, is available at http://www.hrpo.pitt.edu/.

Continuing review (CR) can be submitted by clicking “Create Modification/CR” from the active study at least 5 weeks prior to the expiration date.

Clinical research being conducted in an UPMC facility cannot begin until fiscal approval is received from the UPMC Office of Sponsored Programs and Research Support (OSPARS).

If you have any questions, please contact the University of Pittsburgh IRB Coordinator, Teresa McKaveney.

Please take a moment to complete our Satisfaction Survey as we appreciate your feedback.

University of Pittsburgh
Institutional Review Board

Human Research Protection Office
3500 Fifth Avenue, Suite 106
Pittsburgh, PA 15213
Tel (412) 383-1480
www.hrpo.pitt.edu

APPROVAL OF SUBMISSION (Expedited)

Date:	February 10, 2020
IRB:	MOD19100368-001
PI:	Lacey Heinsberg
Title:	Multi-omics of the Iron Homeostasis Pathway in Patient Outcomes Following aSAH
Funding:	Name: International Society of Nurses in Genetics; Name: University of Pittsburgh; Name: University of Pittsburgh; Name: National Institute of Nursing Research, Funding Source ID: F31NR017311; Name: American Nurses Foundation; Name: Nightingale Awards of Pennsylvania; Name: Sigma Theta Tau International

The Institutional Review Board reviewed and approved the above referenced study. The study may begin as outlined in the University of Pittsburgh approved application and documents.

Approval Documentation

Review type:	Modification / Update
Approval Date:	2/10/2020
Expiration Date:	12/8/2020

Approved Documents:	• Heinsberg_ANF_App_Updated.pdf, Category: Sponsor Attachment; • Heinsberg_SONGrant2018.pdf, Category: Sponsor Attachment; • LaceyWrightHeinsberg_HoffmanGrant2017.pdf, Category: Sponsor Attachment; • LaceyWrightHeinsberg_ISONGGrant2017.pdf, Category: Sponsor Attachment; • Nightingale Awards Scholarship.pdf, Category: Sponsor Attachment; • Wright_EtaSTTI_ResearchAwardApplication.pdf, Category: Sponsor Attachment;

As the Principal Investigator, you are responsible for the conduct of the research and to ensure accurate documentation, protocol compliance, reporting of possibly study-related adverse events and unanticipated problems involving risk to participants or others. The HRPO Reportable Events policy, Chapter 17, is available at http://www.hrpo.pitt.edu/.

Continuing review (CR) can be submitted by clicking “Create Modification/CR” from the active study at least 5 weeks prior to the expiration date.

Clinical research being conducted in an UPMC facility cannot begin until fiscal approval is received from the UPMC Office of Sponsored Programs and Research Support (OSPARS).

If you have any questions, please contact the University of Pittsburgh IRB Coordinator, Teresa McKaveney.

Please take a moment to complete our Satisfaction Survey as we appreciate your feedback.

Pitt_510_EXP

Bibliography

1. Pfohman M, Criddle LM. Epidemiology of Intracranial Aneurysm and Subarachnoid Hemorrhage. *J Neurosci Nurs*. 2001;33(1):39-41. doi:10.1097/01376517-200102000-00005.

2. Rosenbaum BP, Kelly ML, Kshettry VR, Weil RJ. Neurologic disorders, in-hospital deaths, and years of potential life lost in the USA, 1988-2011. *J Clin Neurosci*. 2014;21(11):1874-1880. doi:10.1016/j.jocn.2014.05.006.

3. van Gijn J1, Kerr RS RG. Subarachnoid haemorrhage. *Lancet*. 2007;369(9558):306-318. doi:10.1136/bmj.333.7561.235.

4. Roos YBWEM, Dijkgraaf MGW, Albrecht KW, et al. Direct costs of modern treatment of aneurysmal subarachnoid hemorrhage in the first year after diagnosis. *Stroke*. 2002;33(6):1595-1599. doi:10.1161/01.STR.0000016401.49688.2F.

5. Teo M, Teo M, Turner C, et al. What Factors Determine Treatment Outcome in Aneurysmal Subarachnoid Hemorrhage in the Modern Era? A Post Hoc STASH Analysis. *World Neurosurg*. 2017;105:270-281. doi:10.1016/j.wneu.2017.05.005.

6. Lee J-Y, Keep RF, Hua Y, Pandey A, Xi G. The Role of Iron in the Brain Following Subarachnoid Hemorrhage. In: Li YV, Zhang JH, eds. *Metal Ion in Stroke*. ; 2012:273-282.

7. Dixon SJ, Lemberg KM, Lamprecht MR, et al. Ferroptosis: an iron-dependent form of nonapoptotic cell death. *Cell*. 2012;149(5):1060-1072. doi:10.1016/j.cell.2012.03.042.

8. Gomes JA, Selim M, Cotleur A, et al. Brain iron metabolism and brain injury following subarachnoid hemorrhage: iCeFISH-pilot (CSF iron in SAH). *Neurocrit Care*. 2014;21(2):285-293. doi:10.1007/s12028-014-9977-8.

9. Loftspring MC. Iron and early brain injury after subarachnoid hemorrhage. *J Cereb blood flow Metab*. 2010;30(11):1791-1792. doi:10.1038/jcbfm.2010.139.

10. Zacharia BE, Hickman ZL, Grobelny BT, et al. Epidemiology of aneurysmal subarachnoid hemorrhage. *Neurosurg Clin N Am*. 2010;21(2):221-233. doi:10.1016/j.nec.2009.10.002.

11. van Gijn J, Kerr RS, Rinkel GJE. Subarachnoid haemorrhage. *Lancet*. 2007;369(9558):306-318. doi:10.1016/S0140-6736(07)60153-6.

12. Turi E, Conley Y, Stanfill AG. A Literature Review of Psychosocial Comorbidities Related to Working Capacity After Aneurysmal Subarachnoid Hemorrhage. *J Neurosci Nurs*. 2017;49(3):179-184. doi:10.1097/JNN.0000000000000281.

13. le Roux AA, Wallace MC. Outcome and Cost of Aneurysmal Subarachnoid Hemorrhage.

Neurosurg Clin N Am. 2010;21(2):235-246. doi:10.1016/j.nec.2009.10.014.

14. Yoon S, Yoon JC, Winkler E, Liu C, Lawton MT. Nationwide Analysis of Cost Variation for Treatment of Aneurysmal Subarachnoid Hemorrhage. *Stroke*. 2019;50(1):199-203. doi:10.1161/STROKEAHA.118.023079.

15. Lee J-Y, Keep RF, Hua Y, Ernestus R-I, Xi G. Deferoxamine reduces early brain injury following subarachnoid hemorrhage. *Acta Neurochir Suppl*. 2011;112:101-106. doi:10.1007/978-3-7091-0661-7_18.

16. Yeatts SD, Palesch YY, Moy CS, Selim M. High dose deferoxamine in intracerebral hemorrhage (Hi-Def) trial: Rationale, design, and methods. *Neurocrit Care*. 2013;19(2):257-266. doi:10.1007/s12028-013-9861-y.

17. Meng H, Li F, Hu R, et al. Deferoxamine alleviates chronic hydrocephalus after intraventricular hemorrhage through iron chelation and Wnt1/Wnt3a inhibition. *Brain Res*. 2015;1602(C):44-52. doi:10.1016/j.brainres.2014.08.039.

18. Oka F, Chung DY, Suzuki M, Ayata C. Delayed Cerebral Ischemia After Subarachnoid Hemorrhage: Experimental-Clinical Disconnect and the Unmet Need. *Neurocrit Care*. January 2019. doi:10.1007/s12028-018-0650-5.

19. Dev S, Babitt JL. Overview of iron metabolism in health and disease. *Hemodial Int*. 2017;21:S6-S20. doi:10.1111/hdi.12542.

20. Ganz T. Systemic iron homeostasis. *Physiol Rev*. 2013;93(4):1721-1741. doi:10.1152/physrev.00008.2013.

21. Ponka P. Cell Biology of Heme. *Am J Med Sci*. 2016:2016.

22. Wagner KR, Sharp FR, Ardizzone TD, Lu A, Clark JF. Heme and iron metabolism: role in cerebral hemorrhage. *J Cereb blood flow Metab*. 2003;23(6):629-652. doi:10.1097/01.WCB.0000073905.87928.6D.

23. Ganz T. Hepcidin and iron regulation, 10 years later. *Blood*. 2011;117(17):4425-4433. doi:10.1182/blood-2011-01-258467.

24. Bishop GM, Robinson SR. Quantitative analysis of cell death and ferritin expression in response to cortical iron: Implications for hypoxia-ischemia and stroke. *Brain Res*. 2001;907(1-2):175-187. doi:10.1016/S0006-8993(01)02303-4.

25. Carbonell T, Rama R. Iron, oxidative stress and early neurological deterioration in ischemic stroke. *Curr Med Chem*. 2007;14(8):857-874. http://www.ncbi.nlm.nih.gov/pubmed/17430141. Accessed December 5, 2015.

26. Xi G, Keep RF, Hoff JT. Mechanisms of brain injury after intracerebral haemorrhage. *Lancet Neurol*. 2006;5(1):53-63. doi:10.1016/S1474-4422(05)70283-0.

27. Selim M. Treatment with the Iron Chelator, Deferoxamine Mesylate, Alters Serum Markers of Oxidative Stress in Stroke Patients. *Transl Stroke Res*. 2010;1(1):35-39. doi:10.1007/s12975-009-0001-0.

28. Kantor E, Bayir H, Ren D, et al. Haptoglobin genotype and functional outcome after aneurysmal subarachnoid hemorrhage. *J Neurosurg*. 2014;120(2):386-390. doi:10.3171/2013.10.JNS13219.

29. Ono S, Zhang ZD, Marton LS, et al. Heme oxygenase-1 and ferritin are increased in cerebral arteries after subarachnoid hemorrhage in monkeys. *J Cereb Blood Flow Metab*. 2000;20(7):1066-1076. doi:10.1097/00004647-200007000-00006.

30. Suzuki H, Muramatsu M, Kojima T, Taki W. Intracranial Heme Metabolism and Cerebral Vasospasm after Aneurysmal Subarachnoid Hemorrhage. *Stroke*. 2003;34(12):2796-2800. doi:10.1161/01.STR.0000103743.62248.12.

31. Felling RJ, Song H. Epigenetic mechanisms of neuroplasticity and the implications for stroke recovery. *Exp Neurol*. 2015;268:37-45. doi:10.1016/j.expneurol.2014.09.017.

32. Heinsberg LW, Arockiaraj AI, Crago EA, et al. Genetic Variability and Trajectories of DNA Methylation May Support a Role for HAMP in Patient Outcomes After Aneurysmal Subarachnoid Hemorrhage. *Neurocrit Care*. 2020;32(2):550-563. doi:10.1007/s12028-019-00787-4.

33. Baccarelli A, Wright R, Bollati V, et al. Ischemic heart disease and stroke in relation to blood DNA methylation. *Epidemiology*. 2010;21(6):819-828. doi:10.1097/EDE.0b013e3181f20457.

34. Zhang ZY, Zhang ZY, Fauser U, et al. Global hypomethylation defines a sub-population of reactive microglia/macrophages in experimental traumatic brain injury. *Neurosci Lett*. 2007;429(1):1-6. doi:10.1016/j.neulet.2007.09.061.

35. Gallek, Matthew, Alexander, Shelia, Crago, Elizabeth, Sherwood, Paula, Horowitz M, Poloyac S, Conley Y, Gallek M, Alexander S, et al. Endothelin-1 and endothelin receptor gene variants and their association with negative outcomes following aneurysmal subarachnoid hemorrhage. *Biol Res Nurs*. 2014;15(4):390-397. doi:10.1177/1099800412459674.Endothelin-1.

36. Donnelly MK, Crago E a, Conley YP, et al. 20-HETE is associated with unfavorable outcomes in subarachnoid hemorrhage patients. *J Cereb Blood Flow Metab*. 2015;35(9):1515-1522. doi:10.1038/jcbfm.2015.75.

37. Thampatty BP, Sherwood PR, Gallek MJ, et al. Role of endothelin-1 in human aneurysmal subarachnoid hemorrhage: associations with vasospasm and delayed cerebral ischemia. *Neurocrit Care*. 2011;15(1):19-27. doi:10.1007/s12028-011-9508-9.

38. Zaghlool SB, Al-Shafai M, Al Muftah WA, Kumar P, Falchi M, Suhre K. Association of dna methylation with age, gender, and smoking in an arab population. *Clin Epigenetics*.

2015;7(1):7-6. doi:10.1186/s13148-014-0040-6.

39. Aaslid R, Huber P, Nornes H. Evaluation of cerebrovascular spasm with transcranial Doppler ultrasound. *J Neurosurg*. 1984;60(1):37-41. doi:10.3171/jns.1984.60.1.0037.

40. Sloan MA, Haley EC, Kassell NF, et al. Sensitivity and specificity of transcranial Doppler ultrasonography in the diagnosis of vasospasm following subarachnoid hemorrhage. *Neurology*. 1989;39(11):1514-1518. http://www.ncbi.nlm.nih.gov/pubmed/2682350. Accessed September 29, 2016.

41. Hellawell DJ, Signorini DF. The Edinburgh Extended Glasgow Outcome Scale (EEGOS): rationale and pilot studies. *Int J Rehabil Res*. 1997;20(4):345-354. http://www.ncbi.nlm.nih.gov/pubmed/9459101. Accessed September 29, 2016.

42. Jabbarli R, Reinhard M, Roelz R, et al. The predictors and clinical impact of intraventricular hemorrhage in patients with aneurysmal subarachnoid hemorrhage. *Int J Stroke*. 2016;11(1):68-76. doi:10.1177/1747493015607518.

43. Du P, Zhang X, Huang CC, et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics*. 2010;11(1):587. doi:10.1186/1471-2105-11-587.

44. Jones BL, Nagin DS, Roeder K. A SAS procedure based on mixture models for estimating developmental trajectories. *Sociol Methods Res*. 2001;29(3):374-393. doi:10.1177/0049124101029003005.

45. Nagin DS. *Group-Based Modeling of Development*. Harvard University Press; 2005.

46. Titus AJ, Gallimore RM, Salas LA, Christensen BC. Cell-type deconvolution from DNA methylation: A review of recent applications. *Hum Mol Genet*. 2017;26(R2):R216-R224. doi:10.1093/hmg/ddx275.

47. Houseman EA, Molitor J, Marsit CJ. Reference-free cell mixture adjustments in analysis of DNA methylation data. *Bioinformatics*. 2014;30(10):1431-1439. doi:10.1093/bioinformatics/btu029.

48. Wakefield J. A Bayesian Measure of the Probability of False Discovery in Molecular Genetic Epidemiology Studies. *Am J Hum Genet*. 2007;81:208-227. doi:10.1086/519024.

49. Wakefield J. Bayes factors for Genome-wide association studies: Comparison with P-values. *Genet Epidemiol*. 2008;33(1):79-86. doi:10.1002/gepi.20359.

50. Heinsberg LW, Alexander SA, Crago EA, et al. Genetic Variability in the Iron Homeostasis Pathway and Patient Outcomes After Aneurysmal Subarachnoid Hemorrhage. *Neurocrit Care*. April 2020. doi:10.1007/s12028-020-00961-z.

51. Heinsberg LW, Turi E, Ren D, et al. Evaluation of APOE Genotype and Ability to Perform Activities of Daily Living Following Aneurysmal Subarachnoid Hemorrhage. *Biol Res*

Nurs. 2018;20(2):177-182. doi:10.1177/1099800417746504.

52. Knisely MR, Maserati M, Heinsberg LW, et al. Symptom Science: Advocating for Inclusion of Functional Genetic Polymorphisms. *Biol Res Nurs*. 2019;21(4):349-354. doi:10.1177/1099800419846407.

53. Heinsberg LW, Conley YP. Omics for Nurse Scientists Conducting Environmental Health Research. *Annu Rev Nurs Res*. 2019;38(1):35-52. doi:10.1891/0739-6686.38.35.

54. Heinsberg L, Bender C, Rohay J, Sherwood P, Crago E. Psychometric Properties of the Patient Assessment of Own Functioning Inventory Following Aneurysmal Subarachnoid Hemorrhage. *J Nurs Meas*.

55. Price EM, Robinson WP. Adjusting for batch effects in DNA methylation microarray data, a lesson learned. *Front Genet*. 2018;9(MAR). doi:10.3389/fgene.2018.00083.

56. Johnson WE, Li C, Rabinovic A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*. 2007;8(1):118-127. doi:10.1093/biostatistics/kxj037.

57. Fortin JP, Labbe A, Lemire M, et al. Functional normalization of 450k methylation array data improves replication in large cancer studies. *Genome Biol*. 2014;15(11):503. doi:10.1186/s13059-014-0503-2.

58. Aryee MJ, Jaffe AE, Corrada-Bravo H, et al. Minfi: A flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics*. 2014;30(10):1363-1369. doi:10.1093/bioinformatics/btu049.

59. Oros Klein K, Grinek S, Bernatsky S, et al. FuntooNorm: An R package for normalization of DNA methylation data when there are multiple cell or tissue types. *Bioinformatics*. 2016;32(4):593-595. doi:10.1093/bioinformatics/btv615.

60. Li Y, Liu Y, Wu P, et al. Inhibition of Ferroptosis Alleviates Early Brain Injury After Subarachnoid Hemorrhage In Vitro and In Vivo via Reduction of Lipid Peroxidation. *Cell Mol Neurobiol*. April 2020. doi:10.1007/s10571-020-00850-1.

61. Fang Y, Gao S, Wang X, et al. Programmed Cell Deaths and Potential Crosstalk With Blood–Brain Barrier Dysfunction After Hemorrhagic Stroke. *Front Cell Neurosci*. 2020;14. doi:10.3389/fncel.2020.00068.

62. Arockiaraj AI, Liu D, Shaffer JR, et al. Methylation Data Processing Protocol & Comparison of Blood and Cerebral Spinal Fluid Following Aneurysmal Subarachnoid Hemorrhage. *bioRxiv*. March 2020. doi:10.1101/2020.03.24.005264.

63. Xu Z, Niu L, Li L, Taylor JA. ENmix: A novel background correction method for Illumina HumanMethylation450 BeadChip. *Nucleic Acids Res*. 2016;44(3):e20. doi:10.1093/nar/gkv907.

64. Xu Z, Langie SAS, De Boever P, Taylor JA, Niu L. RELIC: A novel dye-bias correction method for Illumina Methylation BeadChip. *BMC Genomics*. 2017;18(1):4. doi:10.1186/s12864-016-3426-3.

65. Jennett B, Bond M. Assessment of outcome after severe brain damage. *Lancet (London, England)*. 1975;1(7905):480-484. http://www.ncbi.nlm.nih.gov/pubmed/46957. Accessed January 3, 2017.

66. Team RC. R: A language and environment for statistical computing. 2018. https://www.r-project.org/.

67. Titus AJ, Gallimore RM, Salas LA, Christensen BC. Cell-type deconvolution from DNA methylation: A review of recent applications. *Hum Mol Genet*. 2017;26(R2):R216-R224. doi:10.1093/hmg/ddx275.

68. Hou Y, Zhang S, Wang L, et al. Estrogen regulates iron homeostasis through governing hepatic hepcidin expression via an estrogen response element. *Gene*. 2012;511(2):398-403. doi:10.1016/j.gene.2012.09.060.

69. Rogers JT, Venkataramani V, Washburn C, et al. A role for amyloid precursor protein translation to restore iron homeostasis and ameliorate lead (Pb) neurotoxicity. *J Neurochem*. 2016;138(3):479-494. doi:10.1111/jnc.13671.

70. Nalivaeva NN, Turner AJ. The amyloid precursor protein: A biochemical enigma in brain development, function and disease. *FEBS Lett*. 2013;587(13):2046-2054. doi:10.1016/j.febslet.2013.05.010.

71. Gasparoni G, Bultmann S, Lutsik P, et al. DNA methylation analysis on purified neurons and glia dissects age and Alzheimer's disease-specific changes in the human cortex. *Epigenetics and Chromatin*. 2018;11(1):41. doi:10.1186/s13072-018-0211-3.

72. Kim S, Yu NK, Kaang BK. CTCF as a multifunctional protein in genome regulation and gene expression. *Exp Mol Med*. 2015;47(6):e166. doi:10.1038/emm.2015.33.

73. Vostrov AA, Quitschke WW. The zinc finger protein CTCF binds to the APBβ domain of the amyloid β-protein precursor promoter: Evidence for a role in transcriptional activation. *J Biol Chem*. 1997;272(52):33353-33359. doi:10.1074/jbc.272.52.33353.

74. Telling ND, Everett J, Collingwood JF, et al. Iron Biochemistry is Correlated with Amyloid Plaque Morphology in an Established Mouse Model of Alzheimer's Disease. *Cell Chem Biol*. 2017;24(10):1205-1215.e3. doi:10.1016/j.chembiol.2017.07.014.

75. Sobol A, Galluzzo P, Weber MJ, Alani S, Bocchetta M. Depletion of Amyloid Precursor Protein (APP) Causes G0 Arrest in Non-Small Cell Lung Cancer (NSCLC) Cells. *J Cell Physiol*. 2015;230(6):1332-1341. doi:10.1002/jcp.24875.

76. Zhang F, Tao Y, Zhang Z, et al. Metalloreductase Steap3 coordinates the regulation of iron homeostasis and inflammatory responses. *Haematol Hematol J*. 2012;97(12):1826.

http://connection.ebscohost.com/c/articles/84662557/metalloreductase-steap3-coordinates-regulation-iron-homeostasis-inflammatory-responses.

77. Han M, Xu R, Wang S, et al. Six-Transmembrane Epithelial Antigen of Prostate 3 Predicts Poor Prognosis and Promotes Glioblastoma Growth and Invasion. *Neoplasia (United States)*. 2018;20(6):543-554. doi:10.1016/j.neo.2018.04.002.

78. Howie HL, Hay AM, De Wolski K, et al. Differences in Steap3 expression are a mechanism of genetic variation of RBC storage and oxidative damage in mice. *Blood Adv*. 2019;3(15):2272-2285. doi:10.1182/bloodadvances.2019000605.

79. Shi C, Awad IA, Jafari N, et al. Genomics of human intracranial aneurysm wall. *Stroke*. 2009;40(4):1252-1261. doi:10.1161/STROKEAHA.108.532036.

80. Nanami M, Ookawara T, Otaki Y, et al. Tumor Necrosis Factor-α–Induced Iron Sequestration and Oxidative Stress in Human Endothelial Cells. *Arterioscler Thromb Vasc Biol*. 2005;25(12).

81. Scaccabarozzi A, Arosio P, Weiss G, et al. Relationship between TNF-alpha and iron metabolism in differentiating human monocytic THP-1 cells. *Br J Haematol*. 2000;110(4):978-984. http://www.ncbi.nlm.nih.gov/pubmed/11054092. Accessed October 23, 2016.

82. Shinozaki G, Braun PR, Hing BWQ, et al. Epigenetics of Delirium and Aging: Potential Role of DNA Methylation Change on Cytokine Genes in Glia and Blood Along With Aging. *Front Aging Neurosci*. 2018;10:311. doi:10.3389/fnagi.2018.00311.

83. Carraro JCC, Hermsdorff HHM, Mansego ML, et al. Higher fruit intake is related to TNF-α hypomethylation and better glucose tolerance in healthy subjects. *J Nutrigenet Nutrigenomics*. 2016;9(2-4):95-105. doi:10.1159/000448101.

84. Chen L, Dong Y, Wang X, et al. Epigenome-Wide Association Study of Dietary Fiber Intake in African American Adolescents. *Mol Nutr Food Res*. 2018;62(12):e1800155. doi:10.1002/mnfr.201800155.

85. Ma L, Jiang Y, Dong YN, Gao J, Du B, Liu DW. Anti-TNF-alpha antibody attenuates subarachnoid hemorrhage-induced apoptosis in the hypothalamus by inhibiting the activation of Erk. *Neuropsychiatr Dis Treat*. 2018;14:525-536. doi:10.2147/NDT.S154809.

86. Moore LD, Le T, Fan G. DNA methylation and its basic function. *Neuropsychopharmacology*. 2013;38(1):23-38. doi:10.1038/npp.2012.112.

87. Praticò D, Uryu K, Leight S, Trojanoswki JQ, Lee VM-Y. Increased lipid peroxidation precedes amyloid plaque formation in an animal model of alzheimer amyloidosis. *J Neurosci*. 2001;21(12):4183-4187. doi:10.1523/jneurosci.21-12-04183.2001.

88. Adibhatla RM, Hatcher JF. Altered lipid metabolism in brain injury and disorders. *Subcell Biochem*. 2008;49:241-268. doi:10.1007/978-1-4020-8831-5_9.