Nucleosides Associated With Incident Ischemic Stroke in the REGARDS and JHS Cohorts

Zsuzsanna Ament, PhD, Amit Patki, MS, Ninad Chaudhary, MBBS, Varun M. Bhave, BA,
Ana-Lucia Garcia Guarniz, MD, Yan Gao, MS, MPH, Robert E. Gerszten, MD, Adolfo Correa, MD, PhD,
Suzanne E. Judd, PhD, Mary Cushman, MD, MSc, D. Leann Long, PhD, M. Ryan Irvin, PhD, and
W. Taylor Kimberly, MD, PhD

Correspondence
Dr. Kimberly
wtkimberly@
mgh.harvard.edu

Abstract

Background and Objectives
Both genetic and environmental factors contribute to stroke risk. We sought to identify novel metabolites associated with incident stroke in the Reasons for Geographic and Racial Differences in Stroke (REGARDS) cohort and determine whether they reflected genetic or environmental variation.

Methods
This was a stroke case–cohort observational study nested in REGARDS. Cases were defined as incident stroke and metabolomic profiles were compared to a randomly selected control cohort. In baseline plasma samples, 162 metabolites were measured using liquid chromatography–tandem mass spectrometry. Cox proportional hazards models were adjusted for age, sex, race, and age by race in the base model. Fully adjusted models included traditional stroke risk factors. Mediation analyses conducted for these stroke risk factors used the metabolite as mediator. Genome-wide associations with the leading candidate metabolites were calculated using array data. Replication analyses in the Jackson Heart Study (JHS) were conducted using random effects meta-analysis.

Results
There were 2,043 participants who were followed over an average period of 7.1 years, including 1,075 stroke cases and 968 random controls. Nine metabolites were associated with stroke in the base model, 8 of which were measured and remained significant in meta-analysis with JHS. In the fully adjusted model in REGARDS, guanosine (hazard ratio [HR] 1.34, 95% CI 1.18–1.53; $p = 7.26 \times 10^{-6}$) and pseudouridine (HR 1.28, 95% CI 1.13–1.45; $p = 1.03 \times 10^{-4}$) were associated with incident ischemic stroke following Bonferroni adjustment. Guanosine also partially mediated the relationship between hypertension and stroke (17.6%) and pseudouridine did not mediate any risk factor. Genome-wide association analysis identified loci rs34631560 and rs34631560 associated with pseudouridine, but these did not explain the association of pseudouridine with stroke.

Discussion
Guanosine and pseudouridine are nucleosides associated with incident ischemic stroke independently of other risk factors. Genetic and mediation analyses suggest that environmental exposures rather than genetic variation link nucleoside levels to stroke risk.

Classification of Evidence
This study provides Class II evidence that guanosine and pseudouridine are associated with incident stroke.

From the Center for Genomic Medicine, Harvard Medical School (Z.A., W.T.K.), and Department of Neurology (Z.A., A.-L.G.G., W.T.K.), Massachusetts General Hospital, Boston; Departments of Epidemiology (A.P., N.C., R.M.I.) and Biostatistics (S.E.J., L.L.), School of Public Health, University of Alabama at Birmingham; Harvard Medical School (V.M.B.), Boston, MA; The Jackson Heart Study (Y.G., A.C.), University of Mississippi Medical Center, Jackson; Department of Medicine (R.E.G.), Beth Israel Deaconess Medical Center, Boston, MA; and Department of Medicine (M.C.), Larner College of Medicine at the University of Vermont, Burlington.

Glossary

AF = atrial fibrillation; **ARIC** = Atherosclerosis Risk in Communities; **CVD** = cardiovascular disease; **DM** = diabetes mellitus; **DMGV** = dimethylguanidino valeric acid; **HR** = hazard ratio; **JHS** = Jackson Heart Study; **LVH** = left ventricular hypertrophy; **NMR** = nuclear magnetic resonance; **REGARDS** = Reasons for Geographic and Racial Differences in Stroke; **SBP** = systolic blood pressure; **SNP** = single nucleotide polymorphism.

Stroke remains a leading cause of death and disability in the United States,[1] with significant geographic and racial disparities in its incidence. Traditional risk factors derived from the Framingham Heart Study[2] and Cardiovascular Health Study[3] have identified several core stroke risk factors, including age, sex, hypertension, systolic blood pressure, diabetes, smoking, atrial fibrillation, and cardiovascular disease. Although these risk factors are important for guiding stroke prevention strategies, there remains unaccounted-for risk, which is amplified among Black individuals.[4] The identification of additional factors that either directly contribute to or mediate the risk of stroke through established risk factors is also important to improve prevention.[5]

The composition of circulating metabolites potentially represents the convergence of physiology, gene expression, and environmental conditions. Metabolomics can therefore capture systems-level information and may help in estimating stroke risk, defining stroke diagnosis, or identifying stroke etiology.[6] Although stroke risk is complex, metabolite markers could represent underlying biological processes that account for some of the excess risk not attributable to traditional stroke risk factors. Furthermore, examining genetic variants linked to candidate metabolites could help elucidate how levels of these metabolites are regulated and provide insights into individual variation.

Our aim was to look for differences between individuals with and without incident stroke using a targeted metabolomics approach designed to measure key components of central metabolic pathways. Currently mapped metabolic networks suggest that certain endogenous or exogenous stimuli (food, stress, infections, genetic susceptibility, and other environmental factors) lead to systemic changes detectable and reflected by levels of circulating metabolites. Second, we hypothesize that, if found to associate with stroke occurrence, specific metabolites could be used as early markers of disease risk. Finally, metabolites and their associated metabolic network reflect disease phenotypes and indicate potential routes for therapeutic interventions.

The biracial Reasons for Geographic and Racial Differences in Stroke (REGARDS) cohort provides a unique opportunity for metabolomic investigations related to stroke risk, and potential differences by race, given its size and the number of incident stroke cases. The objectives of this investigation were to (1) identify novel metabolites associated with incident stroke, (2) determine whether candidate stroke metabolites were mediators of known stroke risk factors, and (3) identify racial variation in polymorphisms linked to candidate metabolites. Leading candidates were studied and meta-analyzed with an independent population-based observational study, the Jackson Heart Study (JHS).[7]

Methods

Study Populations

The REGARDS study is a prospective cohort study that enrolled 30,239 non-Hispanic Black and White participants ≥45 years of age between 2003 and 2007. We report data on incident ischemic stroke cases that were compared to a stratified cohort random subset of participants. Methods for enrollment and the design of the cohort study have been described in detail.[8,9] Briefly, participants were contacted by phone and consented individuals completed a telephone interview to obtain clinical, demographic, and lifestyle information (telephone response rate 33%; cooperation rate 49%). Exclusion criteria were medical conditions preventing long-term participation, past malignancies or active treatment for cancer, nursing home placement, inability to communicate in English, and race other than Black or White. Race classification was self-reported during the telephone interview.[10] During an in-home visit 2–3 weeks later, fasting baseline EDTA blood samples were collected by venipuncture.[11] Blood samples were stored on ice until centrifuged and subsequent plasma aliquots were stored in a central laboratory at −80°C until metabolite profiling analysis. Participants were contacted every 6 months by telephone to ascertain hospitalizations and health care encounters for stroke.

Medical records, including neuroimaging and other diagnostic reports, were retrieved and centrally reviewed by physicians to confirm the diagnosis, stroke type, and possible causes. In instances where imaging data were unavailable or if medical records were judged insufficient, a questionnaire was completed using a protocol developed for previous stroke clinical trials[12,13] and observational studies.[14,15] At least 2 physician adjudicators reviewed all the available information and classified events by stroke type and severity for every recorded potential event. In cases of disagreement, additional adjudicators reviewed the event. A stroke event was recorded if all reviewers agreed on the occurrence of stroke and stroke subtype. Further details about the study population and event verification protocols have been published elsewhere.[8]

For this analysis, we included all ischemic strokes adjudicated through April 1, 2019 (n = 1,075) without prebaseline stroke or TIA and a cohort random sample that was stratified on age, race, and sex, as previously described (n = 968).[16] Ischemic stroke was

Table 1 Characteristics

	Cohort random sample[a]	Ischemic stroke
Age, y	65 ± 9	70 ± 9
Female sex	55	50
Black race	39	41
Current smoker	14	16
SBP, mm Hg	127 ± 16	132 ± 17
Hypertension	72	85
DM	20	29
CVD	15	29
LVH	7	15
AF	8	14

Abbreviations: AF = atrial fibrillation; CVD = cardiovascular disease; DM = diabetes mellitus; HTN = hypertension; LVH = left ventricular hypertrophy; SBP = systolic blood pressure.
Values are mean ± SD or %.
[a] Random sample weighted to analytic cohort. Some participants selected for the cohort random sample also experienced stroke and became patients (n = 68).

defined as a focal neurologic deficit lasting >24 hours or nonfocal neurologic symptoms consistent with stroke on neuroimaging as defined by the WHO.[17] Hemorrhagic stroke (n = 122) cases included both intracerebral hemorrhage and subarachnoid hemorrhage, and were excluded from the study.

Covariates

We included covariates as defined in the Framingham stroke risk score function.[2,4] Age, race, and smoking status were determined by self-report. During the in-home visits, systolic blood pressure (SBP) was measured twice and the average of 2 measurements was used. Hypertension was defined as SBP ≥140 mm Hg, diastolic blood pressure ≥90 mm Hg, or self-reported use of antihypertensive medications. Diabetes mellitus (DM) was defined as either current diabetes medication use (insulin or oral glucose lowering agents) or as blood glucose concentrations of ≥126 mg/dL and ≥200 mg/dL for fasted and nonfasted states, respectively.

Cardiovascular disease (CVD) was defined as self-reported history of myocardial infarction, coronary revascularization procedure, or baseline evidence of a prior myocardial infarction on the study ECG. Left ventricular hypertrophy (LVH) was classified by ECG and atrial fibrillation (AF) was determined from medical history or presence on ECG.

Targeted Metabolomics

Polar metabolites were extracted using protein precipitation from 30 μL of EDTA plasma. Sample extraction was carried out over ice and isotopic standards, including proline ($^{13}C_5$, ^{15}N), glutamine ($^{13}C_5$, $^{15}N_2$), deuterated leucine-d10, and phenylalanine-d8 for quality control monitoring. Standards were purchased from Cambridge Isotope Laboratories. The

extracted metabolites from the supernatant were separated on Xbridge Amide columns (2.1 × 100 mm 3.5 μm; Waters) applying previously described methods using dual Infinity II 1290 high-performance liquid chromatography pumps and a 6495 QQQ tandem mass spectrometer (Agilent).[6,18-20]

Human pooled plasma samples were also extracted and injected after every 10 samples for quality control measurements. In this study, 162 metabolites were detected, and all peaks were integrated and reviewed using MassHunter QQQ Quantitative Analysis software (Agilent). Following peak integration, each metabolite was normalized to the nearest pooled plasma samples using standard approaches. Due to the nonparametric distribution of metabolite levels, all values were rank-based inverse normal transformed prior to statistical analyses.

Statistical Methods

To account for the stratified sampling of the cohort random sample, weighted proportional hazards analysis was used to evaluate the associations between exposures and the outcome, as detailed in prior studies.[21-24] Weightings were calculated based on the stratification factors age, race, and sex. At the time of analyses, the random cohort included 68 participants who developed ischemic stroke during the observation period. These participants were censored at the time of the stroke onset, as described elsewhere.[23,25] Weighted Cox proportional hazard regression models were used to calculate the hazard ratios (HRs) per unit SD of each metabolite with incident ischemic stroke. To account for the competing risk of death, cause-specific proportional hazard models were utilized, which is recommended when etiologic questions are of interest.[26] In the base model (model 1), covariates included age, sex, race, and age by race interaction, similar to other reports on stroke from REGARDS.[16,21,23] A fully adjusted model further incorporated the Framingham stroke risk factors: current smoking status, SBP, hypertension, DM, cardiovascular disease, LVH, and AF (model 2). Bonferroni correction was used to account for multiple comparisons based on 162 metabolites ($p < 3.09 \times 10^{-4}$).

The difference in coefficients mediation approach was used to examine the underlying observed relationships between clinical risk factors, metabolites, and ischemic stroke. The HRs for each clinical risk factor were calculated for incident stroke, with and without adjusting for individual metabolite mediators. Differences in HR and p values for mediation were determined. All metabolites associated with incident stroke at a nominal $p \leq 0.05$ were included in these analyses.[27,28] Bootstrapping (n = 100) was used to determine 95% CIs of the difference in HRs of stroke for each risk factor with and without metabolite adjustment. Statistical analyses were conducted using SAS version 9.4, STATA version 15, and survey 4.0 for R version 3.6.

Replication and Meta-analyses

JHS is a prospective, community-based epidemiologic study of cardiovascular disease in Black participants, which enrolled

Table 2 Metabolites Associated With Incident Ischemic Stroke

Metabolite	REGARDS			JHS		
	HR	95% CI	p Value	HR	95% CI	p Value
Guanosine	1.5	1.33–1.69	4.63E-11[a]	—	—	—
Cotinine	1.31	1.17–1.45	7.96E-07[a]	1.28	1.07–1.53	7.00E-03
Gluconic acid	1.31	1.17–1.47	5.31E-06[a]	1.18	0.96–1.44	1.11E-01
Pseudouridine	1.32	1.18–1.47	1.03E-06[a]	1.16	0.95–1.42	1.52E-01
DMGV	1.34	1.2–1.5	3.36E-07[a]	1.17	0.97–1.42	1.04E-01
Acetylneuraminic acid	1.31	1.16–1.47	6.23E-06[a]	—	—	—
Cytidine	1.25	1.12–1.4	6.03E-05[a]	—	—	—
Xanthosine	1.22	1.1–1.36	2.50E-04[a]	0.96	0.8–1.15	6.33E-01
Atorvastatin	1.22	1.1–1.35	1.60E-04[a]	1.6	0.82–3.13	1.68E-01
C4 carnitine	1.19	1.08–1.31	5.20E-04	1.19	0.98–1.45	8.09E-02
Acetylglutamate	1.21	1.08–1.35	6.07E-04	1.14	0.86–1.51	3.54E-01
Uric acid	1.19	1.07–1.32	1.48E-03	1.27	1.05–1.54	1.53E-02

Metabolite	Meta-analysis		REGARDS fully adjusted model		
	Meta HR	Meta p value	HR	95% CI	p Value
Guanosine	—	—	1.34	1.18–1.34	7.26E-06[a]
Cotinine	1.3	1.90E-08[a]	1.18	1.03–1.18	1.38E-02
Gluconic acid	1.28	2.11E-06[a]	1.14	1–1.14	4.57E-02
Pseudouridine	1.27	2.75E-05[a]	1.28	1.13–1.28	1.03E-04[a]
DMGV	1.28	5.27E-05[a]	1.18	1.04–1.18	8.27E-03
Acetylneuraminic acid	—	—	1.14	1–1.14	4.76E-02
Cytidine	—	—	1.2	1.06–1.2	3.32E-03
Xanthosine	1.09	4.63E-01	1.13	1.13–1	5.35E-02
Atorvastatin	1.23	8.12E-05[a]	1.16	1.03–1.16	1.10E-02
C4 carnitine	1.19	1.03E-04[a]	1.12	1–1.12	4.88E-02
Acetylglutamate	1.21	8.64E-05[a]	1.15	1.02–1.15	2.18E-02
Uric acid	1.21	7.64E-05[a]	1.07	0.95–1.07	2.74E-01

Abbreviations: DMGV = dimethylguanidino valeric acid; HR = hazard ratio; JHS = Jackson Heart Study; REGARDS = Reasons for Geographic and Racial Differences in Stroke.
HRs represent the HRs of incident stroke per unit of SD of the baseline level of each metabolite. REGARDS model: age + race + sex + age*race + metabolite; JHS model: age + sex + metabolite; fully adjusted model: age + race + sex + age*race + smoking + atrial fibrillation + hypertension + diabetes + cardiovascular disease + left ventricular hypertrophy.
[a] Values exceed the Bonferroni-corrected threshold.

5,306 participants in 2000–2004. The study design, recruitment, and data collection for the JHS have been described elsewhere.[29] There were 1,944 individuals with metabolomics data available, including 119 incident ischemic stroke cases and 1,825 participants without stroke. Metabolomics measurements in JHS were performed using targeted liquid chromatography with tandem mass spectrometry and have been previously reported.[30,31] There were 126 metabolites in common between JHS and REGARDS. Missing data were excluded from analyses. Statistically significant results from our discovery in REGARDS (6 metabolites) were evaluated for replication (with $p < 0.008$) using Cox proportional hazard analysis of each metabolite and incident ischemic stroke. Given the limited number of events in JHS, models were adjusted for age and sex; because all participants were Black, race was not included as a covariate. The initial candidate metabolites did not meet the prespecified threshold for replication but were directionally consistent with the

Table 3 Associations Between Risk Factors and Metabolites

	Hypertension			DM2			CVD		
	OR	95% CI	p Value	OR	95% CI	p Value	OR	95% CI	p Value
Guanosine	1.51	1.23–1.84	6.40E-05[a]	1.29	1.07–1.55	8.42E-03	1.31	1.02–1.69	3.40E-02
Cotinine	1.10	0.91–1.32	3.11E-01	0.94	0.79–1.13	5.39E-01	1.20	0.98–1.47	8.54E-02
Gluconic acid	1.23	1.02–1.48	2.64E-02	2.63	2.03–3.40	2.42E-13[a]	1.04	0.81–1.33	7.51E-01
Pseudo-uridine	1.24	1.03–1.48	2.17E-02	0.97	0.81–1.15	6.94E-01	1.18	0.94–1.49	1.60E-01
DMGV	1.51	1.26–1.81	1.01E-05[a]	1.51	1.25–1.81	1.18E-05[a]	1.23	0.97–1.56	8.58E-02
Acetyl-neuraminic acid	1.65	1.37–1.99	2.06E-07[a]	2.03	1.60–2.57	6.18E-09[a]	1.24	0.98–1.56	6.84E-02
Cytidine	1.46	1.22–1.75	5.01E-05[a]	1.08	0.90–1.30	3.81E-01	1.29	1.05–1.59	1.64E-02
Uric acid	1.54	1.29–1.86	3.49E-06[a]	1.64	1.36–1.99	3.12E-07[a]	1.11	0.91–1.35	3.13E-01
Atorvastatin	1.24	0.91–1.69	1.70E-01	1.49	1.15–1.94	2.90E-03	1.86	1.41–2.46	1.16E-05[a]
C4-carnitine	1.29	1.08–1.53	4.40E-03	1.14	0.94–1.38	1.73E-01	1.27	1.04–1.55	1.91E-02
Acetyl-glutamate	1.11	0.93–1.33	2.51E-01	1.54	1.27–1.86	9.62E-06[a]	1.22	0.98–1.52	7.78E-02

	AF			Smoking			LVH		
	OR	95% CI	p Value	OR	95% CI	p Value	OR	95% CI	p Value
Guanosine	0.99	0.73–1.35	9.60E-01	1.11	0.89–1.38	3.72E-01	1.04	0.80–1.35	7.68E-01
Cotinine	1.23	0.95–1.61	1.23E-01	16.18	7.60–34.46	7.39E-13[a]	1.09	0.86–1.37	4.95E-01
Gluconic acid	1.09	0.81–1.48	5.61E-01	1.32	1.05–1.65	1.65E-02	1.23	0.97–1.57	9.37E-02
Pseudo-uridine	1.15	0.89–1.48	2.75E-01	0.96	0.77–1.20	7.47E-01	1.03	0.80–1.34	8.06E-01
DMGV	1.22	0.92–1.61	1.69E-01	1.08	0.86–1.36	5.27E-01	1.19	0.92–1.54	1.88E-01
Acetyl-neuraminic acid	1.13	0.85–1.50	4.10E-01	1.35	1.10–1.67	4.71E-03	1.26	0.98–1.62	6.83E-02
Cytidine	0.96	0.74–1.26	7.79E-01	0.97	0.79–1.19	7.88E-01	1.11	0.86–1.45	4.20E-01
Uric acid	0.99	0.76–1.28	9.26E-01	1.01	0.84–1.22	9.04E-01	1.32	0.99–1.77	5.85E-02
Atorvastatin	1.40	0.95–2.06	9.30E-02	1.21	0.86–1.69	2.71E-01	1.11	0.76–1.62	6.04E-01
C4-carnitine	1.25	0.98–1.60	7.41E-02	0.89	0.71–1.12	3.16E-01	0.99	0.74–1.32	9.47E-01
Acetyl-glutamate	0.97	0.76–1.23	8.04E-01	1.28	1.04–1.59	2.29E-02	1.05	0.79–1.39	7.52E-01

Abbreviations: AF = atrial fibrillation; CVD = cardiovascular disease; DM2 = diabetes mellitus type 2; DMGV = dimethylguanidino valeric acid; LVH = left ventricular hypertrophy; OR = odds ratio.
ORs represent the cross-sectional odds of each risk factor per unit of SD in the baseline level of each metabolite. Models were adjusted for age, sex, and race.
[a] Values exceed the Bonferroni-corrected threshold.

findings in REGARDS. Given the limited number of events in JHS, we subsequently performed meta-analysis of the 126 metabolites identified as common between the 2 studies under a random effects model using the open-source METASOFT software tool. For meta-analysis, significance was set at $p < 3.97 \times 10^{-4}$, corresponding to the Bonferroni threshold for 126 tests.[32]

Genome-wide Association Studies

Genotyping was conducted using the Illumina Infinium Multi-Ethnic AMR/AFR Bead Chip (MEGA) array[33] in REGARDS. Preimputation quality control included removal of internal duplicates, HapMap controls, and sex mismatches. Principal component analysis was performed, using EIGEN-SOFT software (version 7.2.1), to estimate population stratification. Imputation using cleaned single nucleotide polymorphisms (SNPs) was conducted with the Transomics for Precision Medicine cosmopolitan reference panel. A total of 15, 248, and 746 variants with imputation quality (R^2) > 0.3 and minor allele frequency >0.01 were retained. SNPs in Hardy-Weinberg disequilibrium ($p < 10^{-6}$) were removed from further analysis. Genome-wide association was performed to identify SNPs associated with the top candidate metabolites and to examine associations separately by race. Models were adjusted for age, sex, and the first 3 genetic principal components. These models were conducted separately by race and meta-analyzed, with analysis carried out in PLINK. The regional association plots were created using SAS version 9.4. The multiple

Table 4 Mediation Role of Metabolites Between Each Risk Factor and Incident Ischemic Stroke

Stroke risk factors, mediating metabolites	HR of risk factor without metabolite	HR of risk factor with metabolite	HR of metabolite	Mediation %	*p* Value
Smoking					
Cotinine	1.84 (1.37–2.49)	1.37 (0.97–1.94)	1.23 (1.09–1.39)	48.3	1.60E-04
Hypertension					
Guanosine	1.91 (1.49–2.46)	1.71 (1.32–2.21)	1.44 (1.27–1.63)	17.6	7.57E-04
DMGV	1.91 (1.49–2.46)	1.76 (1.36–2.27)	1.28 (1.14–1.43)	13.2	1.15E-04
Acetylneuraminic acid	1.91 (1.49–2.46)	1.76 (1.36–2.29)	1.24 (1.10–1.40)	12.6	6.67E-03
Cytidine	1.91 (1.49–2.46)	1.81 (1.40–2.34)	1.21 (1.08–1.35)	8.2	1.52E-03
Uric acid	1.91 (1.49–2.46)	1.83 (1.42–2.37)	1.14 (1.03–1.27)	6.5	3.64E-02
DM					
Gluconic acid	1.73 (1.36–2.20)	1.51 (1.18–1.93)	1.23 (1.09–1.39)	25.5	1.37E-04
DMGV	1.73 (1.36–2.20)	1.61 (1.27–2.05)	1.30 (1.16–1.45)	13.0	1.18E-03
Acetylneuraminic acid	1.73 (1.36–2.20)	1.54 (1.20–1.97)	1.23 (1.09–1.38)	21.2	5.31E-03
Uric acid	1.73 (1.36–2.20)	1.64 (1.29–2.09)	1.13 (1.01–1.26)	9.4	2.28E-02
Acetylglutamate	1.73 (1.36–2.20)	1.62 (1.27–2.07)	1.17 (1.05–1.31)	11.8	8.44E-03
CVD					
Atorvastatin	1.98 (1.53–2.56)	1.91 (1.48–2.47)	1.18 (1.06–1.31)	5.3	1.22E-03

Abbreviations: CVD = cardiovascular disease; DM = diabetes mellitus; DMGV = dimethylguanidino valeric acid; HR = hazard ratio.
Model: adjusted for age, race, sex, age by race interaction, risk factor.

testing corrected significance threshold for genome-wide significance was set at $p < 5 \times 10^{-8}$.

Standard Protocol Approvals, Registrations, and Patient Consents

The REGARDS study was approved by the institutional review boards of all participating institutions and written informed consent was obtained from all participants. The metabolomics analysis was also approved by the Mass General Brigham institutional review board.

Data Availability

Qualified investigators may request access to obtain de-identified data in accordance with institutional data sharing agreements.

Results

Study Population

Over an average follow-up of 7.1 ± 4.5 years, there were 1,075 ischemic stroke cases identified in REGARDS. Participants who developed an ischemic stroke during the follow-up period were approximately 4 years older, more likely to be Black, and had higher rates of comorbidities as defined by the Framingham stroke risk score function[2,4] and including hypertension, DM, CVD, LVH, and AF (Table 1).

Associations of Metabolites With Ischemic Stroke

Nine metabolites surpassed the Bonferroni-adjusted threshold in Cox proportional hazards models adjusting for age, sex, race, and age by race interaction (Table 2). The 9 metabolites included 4 nucleosides (guanosine, pseudouridine, cytidine, and xanthosine), 3 amino acid derivatives (dimethylguanidino valeric acid [DMGV], acetylneuraminic acid, and gluconic acid), and 2 exogenous compounds (atorvastatin and the nicotine metabolite cotinine).

Replication and Meta-analyses With the JHS Cohort

There were 126 metabolites measured in common between the REGARDS and JHS cohorts. Of the 9 leading metabolites identified in REGARDS, 3 were not measured in JHS (guanosine, acetylneuraminic acid, and cytidine), while an additional 3 (uric acid, C4-carnitine, and acetylglutamate) crossed the significance threshold following meta-analysis (Table 2). Except for xanthosine, HRs were similar for each metabolite between the 2 cohorts. Uridine, the precursor of pseudouridine,[34] had the strongest association with incident ischemic stroke in JHS (HR 0.69 per unit of SD in uridine level, 95% CI 0.57–0.83; $p = 7.00 \times 10^{-5}$) and had a similar direction of association in the REGARDS study (HR 0.87, 95% CI 0.79–0.97; $p = 1.23 \times 10^{-2}$), but did not remain significant following meta-analysis (meta HR 0.78; meta $p = 3.61 \times 10^{-2}$)

Table 5 Single Nucleotide Polymorphisms Associated With Plasma Pseudouridine

SNP	Gene	African American ancestry		European ancestry		All	
		β	p Value	β	p Value	β	p Value
rs9498623	GRIK2	0.29	4.84E-08[a]	−0.19	3.31E-01	0.23	4.07E-06
rs71654810	GRIK2	0.29	4.64E-08[a]	−0.24	1.76E-01	0.22	7.63E-06
rs17062194	GRIK2	0.30	2.99E-08[a]	−0.19	3.31E-01	0.24	1.94E-06
rs66484859	GRIK2	0.30	2.64E-08[a]	−0.19	3.31E-01	0.24	1.72E-06
rs7741281	GRIK2	0.30	2.71E-08[a]	−0.19	3.31E-01		
rs12900886	MFGE8; ABHD2	0.11	1.78E-01	0.26	2.97E-06	0.24	2.15E-08[a]
rs12324224	MFGE8; ABHD2	0.11	1.86E-01	0.26	2.97E-06	0.24	2.48E-08[a]
rs71149248	MFGE8; ABHD2	0.11	1.78E-01	0.25	6.50E-06	0.23	3.39E-08[a]
rs4360903	MFGE8; ABHD2	0.15	1.35E-01	0.27	1.12E-06	0.25	2.36E-08[a]
rs11854739	MFGE8; ABHD2	0.15	1.35E-01	0.27	2.12E-06	0.25	3.00E-08[a]
rs35864015	MFGE8; ABHD2	0.15	1.22E-01	0.26	3.88E-06	0.25	4.52E-08[a]
rs34631560	MFGE8; ABHD2	0.11	2.99E-01	0.29	4.63E-07	0.26	2.11E-08[a]

Abbreviation: SNP = single nucleotide polymorphism.
[a] Exceeded the genome-wide association study p value threshold (5.0×10^{-8}).

after correction for multiple testing. However, cotinine, gluconic acid, pseudouridine, DMGV, and atorvastatin as well as 3 additional metabolites (uric acid, C4-carnitine, and acetylglutamate) were significant after meta-analysis (Table 2).

Metabolites and Traditional Risk Factors for Stroke

We next examined which of the leading metabolites were associated with stroke independent of traditional stroke risk factors. We carried forward the metabolites identified in meta-analysis (cotinine, gluconic acid, pseudouridine, DMGV, atorvastatin, uric acid, C4-carnitine, and acetylglutamate) and the 3 metabolites that were unique in REGARDS (guanosine, acetylneuraminic acid, and cytidine). Following adjustment for covariates defined by the Framingham stroke risk function, 2 metabolites remained associated with incident ischemic stroke: guanosine and pseudouridine (Table 2).

To understand the potential role of the top candidate metabolites, we next examined associations with individual stroke risk factors. In age-, sex-, and race-adjusted multivariable models, hypertension was associated with guanosine, DMGV, acetylneuraminic acid, cytidine, and uric acid (Table 3). Metabolites associated with DM included gluconic acid, DMGV, acetylneuraminic acid, uric acid, and acetylglutamate. Cotinine was associated with a history of smoking and atorvastatin was associated with CVD. Pseudouridine and C4-carnitine were not associated with any risk factors (Table 3).

We next determined whether the metabolites mediated the associations of risk factors with incident stroke (Table 4). As a proof of principle, we first examined cotinine, a metabolite of nicotine after smoking,[35] reasoning that this metabolite would partially mediate the association between current smoking and incident ischemic stroke. Cotinine was highly associated with current smoking status (odds ratio 16.2 of being a smoker per unit of SD of cotinine level, 95% CI 7.6–34.5; $p = 7.39 \times 10^{-13}$) and mediated 48% of the association between smoking and incident ischemic stroke (HR without mediator 1.84, 95% CI 1.37–2.49; HR with mediator 1.37, 95% CI 0.97–1.94; mediation $p = 1.60 \times 10^{-4}$). Guanosine was associated with hypertension (Table 4) and mediated the incidence of ischemic stroke through hypertension by 17.6% (HR without mediator 1.91, 95% CI 1.49–2.46; HR with mediator 1.71, 95% CI 1.32–2.21; mediation $p = 7.57 \times 10^{-4}$). Pseudouridine was not a significant mediator of any of the stroke risk factors.

Genome-wide Association Studies

Because pseudouridine and guanosine were identified independent of traditional stroke risk factors, we next evaluated whether genetic variation could account for some of the difference in levels of these metabolites using genome-wide association studies. Black and White participants were first analyzed separately and then jointly in these analyses (Table 5). Among individuals of Black ancestry, we found a genome-wide significant association signal for pseudouridine on chromosome 6 (Figure, A). The lead SNP rs66484859 (GRCh37/hg19 genome assembly, chr6:101578136) minor allele was associated with higher pseudouridine level (β = 0.30; $p = 2.71 \times 10^{-8}$) (Figure, B), and this was located within the GRIK2 (glutamate ionotropic receptor kainate type subunit 2) locus.

Figure Manhattan Plots and Regional Association Plots for Pseudouridine Genome-wide Association Studies
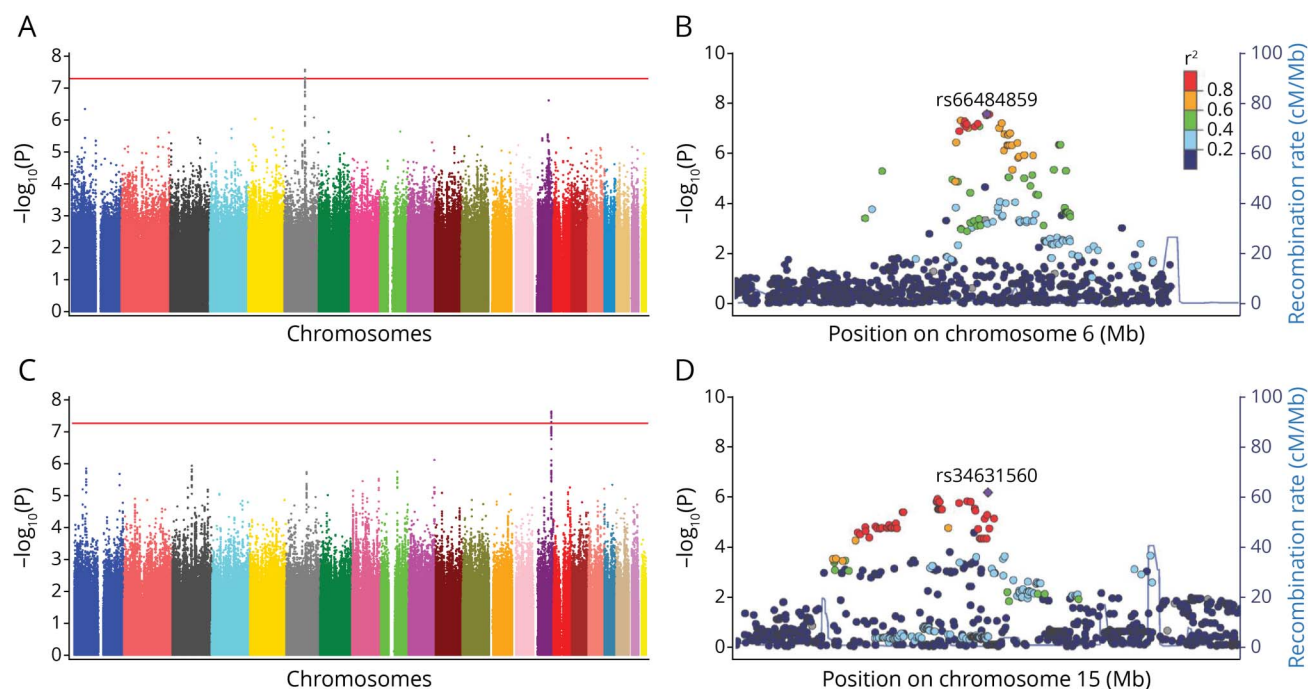
(A) Manhattan plot of participants with African American ancestry for markers associated with pseudouridine. (B) Manhattan plot of participants with European American ancestry for markers associated with pseudouridine. (C) Manhattan plot of participants with African and European American ancestry for markers associated with pseudouridine. (D) Regional association plot for risk loci on chromosome 6 (African American ancestry population only; rs66484859, GRIK2 region). (E) Regional association plot for risk loci on chromosome 16 (rs34631560, MFGE8 region).

In individuals of White ancestry, the association of pseudouridine with the chromosome 6 *GRIK2* locus was not present (Table 5), nor was any other locus; however, when both Black and White ancestry groups were combined, a second region of association was identified on chromosome 15, with lead SNP rs34631560 (GRCh37/hg19 genome assembly, chr15:88973692) adjacent to the *MFGE8* and *ABHD2* (CAMRA) loci (Figure, C and D; Table 5). This association was larger in White participants (Table 5). Neither of the leading SNPs for pseudouridine was associated with ischemic stroke risk and we did not identify any genetic loci in association with guanosine level. Of the other metabolites identified in the base model, DMGV level was associated with the *AGXT2* (alanine-glyoxylate aminotransferase 2) gene locus, which has previously been reported.[36,37]

Classification of Evidence

This study provides Class II evidence that guanosine and pseudouridine are associated with incident stroke.

Discussion

In this study, we performed targeted metabolomic profiling in the biracial REGARDS cohort to identify plasma metabolites associated with risk of incident ischemic stroke, followed by meta-analysis with the JHS cohort. Following adjustment for traditional stroke risk factors, there were 2 nucleoside metabolites, guanosine and pseudouridine, that remained significant. We also examined whether the leading metabolites mediated associations of other risk factors with stroke and found that guanosine mediated 17.6% of the association of hypertension with stroke. Finally, we observed a race-specific genetic locus associated with pseudouridine level, but this was not associated with stroke risk. Taken together, our findings identify 2 novel nucleosides linked to stroke risk that do not appear to be related to underlying genetic variation or to known stroke risk factors. Rather, these nucleosides most likely reflect unrecognized environmental or biological risk pathways.

Our analysis demonstrated that guanosine was a partial mediator of the relationship between hypertension and incident ischemic stroke risk and it also remained associated with stroke risk independent of risk factors. These results could indicate that, in addition to hypertension, guanosine may also serve as a biochemical marker of a process not captured by traditional stroke risk factors. For example, some evidence suggests that guanosine directly contributes to endothelial injury, limiting proliferation and inhibiting cell viability in a dose-dependent fashion.[38] In addition, the level of guanosine and other nucleosides has been linked to dietary intake and subsequent effects on colonic microbiota.[39–42] It is therefore possible that the increased risk associated with elevated guanosine may be related to dietary patterns, which is an important area for future research.

We also found that pseudouridine, a modified isomer of uridine, was associated with incident ischemic stroke. Furthermore, no mediation effects of pseudouridine were detected in the relationship between risk factors and stroke risk. Pseudouridine is the most abundant posttransitional modification of RNA, in which a uridine residue is irreversibly isomerized into pseudouridine. Pseudouridylation can convert nonsense codons into sense codons,[43] alter mRNA splicing and stability,[44] and can also facilitate noncanonical base pairing, each of which is regulated in response to environmental stress signals.[34,45] As a biomarker, pseudouridine is not converted back to uridine and enters the circulation following release after RNA hydrolysis. Therefore, pseudouridine levels may serve as a reflection of cellular stress and turnover. We also observed that uridine was inversely related to stroke risk in both cohorts, although the significance level did not surpass adjustment for multiple hypothesis testing. Nevertheless, associations in the opposite direction for uridine and pseudouridine highlight the possible precursor–product role of these pyrimidines in stroke risk.

To further investigate the role of guanosine and pseudouridine as risk markers for stroke, we examined whether genetic variation could account for the observed risk. We did not identify any SNPs associated with guanosine. SNPs within or adjacent to 2 separate loci were associated with pseudouridine, including rs66484859 in the *GRIK2* gene in individuals of African ancestry and rs34631560 and rs34631560 adjacent to the *MFGE8* and *ABHD2* genes, predominantly in individuals of European ancestry. None of these SNPs or the loci in general have been reported in association with ischemic stroke,[46] including in recent studies that collectively studied more than 16,000 stroke cases.[47,48] Taken together with the available known biology of guanosine and pseudouridine, we hypothesize that elevated levels of these nucleosides have an acquired rather than genetic basis for the relationship to stroke risk. Given the substantial systematic and structural differences in the lived experience between Black and White individuals in the United States, there are a number of potential social and environmental factors that could contribute to metabolite composition such as income, occupation, education, diet, or housing conditions. Future work is needed to systematically evaluate racial disparities in these factors that are linked to each of the identified risk markers. It is also important to note that our findings would not imply that either nucleoside is a direct causal agent for incident stroke, but rather (and perhaps more likely) each metabolite may be an indirect marker.

Several prior studies have evaluated metabolites associated with stroke, although important differences exist in comparison to our study. A recent meta-analysis of 7 cohorts was restricted to European ancestry populations and included 147 metabolites measured by nuclear magnetic resonance (NMR) spectroscopy.[49] The combined cohorts included 38,797 controls and 1,277 ischemic stroke events, a similar number as this study. Based on the lower sensitivity of NMR, higher abundance metabolites were studied. Histidine, pyruvate, phenylalanine, and 7 types of lipoprotein subparticles were associated with incident stroke. Atherosclerosis Risk in Communities (ARIC), a prospective epidemiologic study of participants of European and African ancestry, measured 245 metabolites that included 346 incident ischemic stroke cases.[50] The ARIC study identified tetradecanedioate and hexadecanedioate as stroke risk factors, but no other associations remained significant in fully adjusted models.[50] It is possible that we did not verify these associations due to differences in metabolome coverage, differences in the cohort design, and possibly difference in the time intervals from sampling to stroke onset. However, it is also possible the prior findings, from a smaller number of cases, were chance findings.

There are several strengths of our study that are worth noting, including the large number of incident stroke cases and the inclusion of Black and White participants. We also leveraged a targeted tandem mass spectrometry platform that was enriched for biologically informative metabolites and has high specificity and sensitivity due to the tandem quadrupole design. There are also some limitations. First, there is some uncertainty about the role of long-term metabolite dynamics, diluting any observable relationship with future events in longitudinal studies. Among metabolomics studies, there are also analytical differences depending on the measurement platform, sampling handling and storage, and extraction procedures, each of which could introduce analytical variation. However, these factors would tend to bias results towards the null hypothesis, and therefore may strengthen our findings identified in replication and meta-analysis across 2 independent studies.

Our study identified that pseudouridine and guanosine were associated with incident ischemic stroke independently of traditional stroke risk factors. Mediation and genetic analyses suggest that, together with the known biology of these nucleosides, they may represent markers for acquired and possibly lifestyle-related exposures. Future research may further elucidate the role of these metabolites in linking to stroke risk.

Disclosure

Publication History

Appendix Authors

Name	Location	Contribution
Zsuzsanna Ament, PhD	Center for Genomic Medicine, Massachusetts General Hospital, Harvard Medical School; Department of Neurology, Massachusetts General Hospital, Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; analysis or interpretation of data
Amit Patki, MS	Department of Epidemiology, School of Public Health, University of Alabama at Birmingham	Drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data

Appendix (continued)

Name	Location	Contribution
Ninad Chaudhary, MBBS	Department of Epidemiology, School of Public Health, University of Alabama at Birmingham	Analysis or interpretation of data
Varun M. Bhave, BA	Harvard Medical School, Boston, MA	Drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data
Ana-Lucia Garcia Guarniz, MD	Department of Neurology, Massachusetts General Hospital, Boston	Drafting/revision of the manuscript for content, including medical writing for content; analysis or interpretation of data
Yan Gao, MS, MPH	The Jackson Heart Study, University of Mississippi Medical Center, Jackson	Analysis or interpretation of data
Robert E. Gerszten, MD	Department of Medicine, Beth Israel Deaconess Medical Center, Boston, MA	Drafting/revision of the manuscript for content, including medical writing for content
Adolfo Correa, MD, PhD	The Jackson Heart Study, University of Mississippi Medical Center, Jackson	Drafting/revision of the manuscript for content, including medical writing for content
Suzanne E. Judd, PhD	Department of Biostatistics, School of Public Health, University of Alabama at Birmingham	Study concept or design
Mary Cushman, MD, MSc	Department of Medicine, Larner College of Medicine at the University of Vermont, Burlington	Study concept or design
D. Leann Long, PhD	Department of Biostatistics, School of Public Health, University of Alabama at Birmingham	Analysis or interpretation of data
M. Ryan Irvin, PhD	Department of Epidemiology, School of Public Health, University of Alabama at Birmingham	Drafting/revision of the manuscript for content, including medical writing for content; study concept or design; analysis or interpretation of data
W. Taylor Kimberly, MD, PhD	Center for Genomic Medicine, Massachusetts General Hospital, Harvard Medical School; Department of Neurology, Massachusetts General Hospital, Boston	Drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data

References

1. Benjamin EJ, Muntner P, Alonso A, et al. Heart disease and stroke statistics: 2019 update: a report from the American Heart Association. *Circulation.* 2019;139(10):e56-e528.
2. Wolf PA, D'Agostino RB, Belanger AJ, Kannel WB. Probability of stroke: a risk profile from the Framingham Study. *Stroke.* 1991;22(3):312-318.
3. Manolio TA, Kronmal RA, Burke GL, O'Leary DH, Price TR. Short-term predictors of incident stroke in older adults: the Cardiovascular Health Study. *Stroke.* 1996;27(9):1479-1486.
4. Howard G, Cushman M, Kissela BM, et al. Traditional risk factors as the underlying cause of racial disparities in stroke. *Stroke.* 2011;42:3369-3375.
5. Kamin Mukaz D, Zakai NA, Cruz-Flores S, McCullough LD, Cushman M. Identifying genetic and biological determinants of race-ethnic disparities in stroke in the United States. *Stroke.* 2020;51:3417-3424.

6. Nelson SE, Ament Z, Wolcott Z, Gerszten RE, Kimberly WT. Succinate links atrial dysfunction and cardioembolic stroke. *Neurology.* 2019;92(8):e802-e810.

7. Booth JN, Abdalla M, Tanner RM, et al. Cardiovascular health and incident hypertension in blacks: JHS (the Jackson Heart Study). *Hypertension.* 2017;70(2):285-292.

8. Howard VJ, Cushman M, Pulley L, et al. The reasons for geographic and racial differences in stroke study: objectives and design. *Neuroepidemiology.* 2005;25(3):135-143.

9. Howard VJ, Kleindorfer DO, Judd SE, et al. Disparities in stroke incidence contributing to disparities in stroke mortality. *Ann Neurol.* 2011;69(4):619-627.

10. Howard G, Kissela BM, Kleindorfer DO, et al. Differences in the role of black race and stroke risk factors for first vs. recurrent stroke. *Neurology.* 2016;86(7):637-642.

11. Gillett SR, Boyle RH, Zakai NA, McClure LA, Jenny NS, Cushman M. Validating laboratory results in a national observational cohort study without field centers: the Reasons for Geographic and Racial Differences in Stroke cohort. *Clin Biochem.* 2014; 47(16-17):243-246.

12. Spence JD, Howard VJ, Chambless LE, et al. Vitamin Intervention for Stroke Prevention (VISP) trial: rationale and design. *Neuroepidemiology.* 2001;20(1):16-25.

13. Lefkowitz DS, Brust JC, Goldman L, et al. A pilot study of the end point verification system in the Asymptomatic Carotid Atherosclerosis Study. *J Stroke Cerebrovasc Dis.* 1992;2(2):92-99.

14. Wagenknecht LE, Mayer EJ, Rewers M, et al. The Insulin Resistance Atherosclerosis Study (IRAS) objectives, design, and recruitment results. *Ann Epidemiol.* 1995;5(6):464-472.

15. Chambless LE, Toole JF, Nieto FJ, Rosamond W, Paton C. Association between symptoms reported in a population questionnaire and future ischemic stroke: the ARIC study. *Neuroepidemiology.* 2004;23:33-37.

16. Cushman M, Judd SE, Howard VJ, et al. N-terminal pro-B-type natriuretic peptide and stroke risk: the Reasons for Geographic and Racial Differences in Stroke cohort. *Stroke.* 2014;45(6):1646-1650.

17. Howard G, Cushman M, Kissela BM, et al. Traditional risk factors as the underlying cause of racial disparities in stroke: lessons from the half-full (empty?) glass. *Stroke.* 2011;42:3369-3375.

18. Stapleton CJ, Acharjee A, Irvine HJ, Wolcott ZC, Patel AB, Kimberly WT. High-throughput metabolite profiling: identification of plasma taurine as a potential biomarker of functional outcome after aneurysmal subarachnoid hemorrhage. *J Neurosurg.* 2019:1-8.

19. Ament Z, Bevers MB, Wolcott Z, Kimberly WT, Acharjee A. Uric acid and gluconic acid as predictors of hyperglycemia and cytotoxic injury after stroke. *Transl Stroke Res.* 2021;12(2):293-302.

20. Kimberly WT, O'Sullivan JF, Nath AK, et al. Metabolite profiling identifies anandamide as a biomarker of nonalcoholic steatohepatitis. *JCI Insight.* 2017;2:e92989.

21. Chaudhary NS, Bridges SL, Saag KG, et al. Severity of hypertension mediates the association of hyperuricemia with stroke in the REGARDS case cohort study. *Hypertension.* 2020;75(1):246-256.

22. Olson NC, Cushman M, Judd SE, et al. Associations of coagulation factors IX and XI levels with incident coronary heart disease and ischemic stroke: the REGARDS study. *J Thromb Haemost.* 2017;15(6):1086-1094.

23. Onland-Moret NC, van der ADL, van der Schouw YT, et al. Analysis of case-cohort data: a comparison of different methods. *J Clin Epidemiol.* 2007;60:350-355.

24. Zakai NA, Judd SE, Alexander K, et al. ABO blood type and stroke risk: the Reasons for Geographic and Racial Differences in Stroke study. *J Thromb Haemost.* 2014; 12(4):564-570.

25. Barlow WE, Ichikawa L, Rosner D, Izumi S. Analysis of case-cohort designs. *J Clin Epidemiol.* 1999;52(12):1165-1172.

26. Austin PC, Lee DS, Fine JP. Introduction to the analysis of survival data in the presence of competing risks. *Circulation.* 2016;133(6):601-609.

27. Nguyen QC, Osypuk TL, Schmidt NM, Glymour MM, Tchetgen Tchetgen EJ. Practical guidance for conducting mediation analysis with multiple mediators using inverse odds ratio weighting. *Am J Epidemiol.* 2015;181(5):349-356.

28. Tchetgen Tchetgen EJ. Inverse odds ratio-weighted estimation for causal mediation analysis. *Stat Med.* 2013;32:4567-4580.

29. Tahir UA, Katz DH, Zhao T, et al. Metabolomic profiles and heart failure risk in black adults: insights from the Jackson Heart Study. *Circ Heart Fail.* 2021;14(1):e007275.

30. Kakkar AK, Cimminiello C, Goldhaber SZ, Parakh R, Wang C, Bergmann JF. Low-molecular-weight heparin and mortality in acutely ill medical patients. *N Engl J Med.* 2011;365(26):2463-2472.

31. Cheng S, Rhee EP, Larson MG, et al. Metabolite profiling identifies pathways associated with metabolic risk in humans. *Circulation.* 2012;125(18):2222-2231.

32. Han B, Eskin E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *Am J Hum Genet.* 2011;88(5):586-598.

33. Zhao X, Geng X, Srinivasasainagendra V, et al. A PheWAS study of a large observational epidemiological cohort of African Americans from the REGARDS study. *BMC Med Genomics.* 2019;12(suppl 1):26.

34. Schwartz S, Bernstein DA, Mumbach MR, et al. Transcriptome-wide mapping reveals widespread dynamic-regulated pseudouridylation of ncRNA and mRNA. *Cell.* 2014; 159(1):148-162.

35. Kyeremateng GA, Damiano MD, Dvorchik BH, Vesell ES. Smoking-induced changes in nicotine disposition: application of a new HPLC assay for nicotine and its metabolites. *Clin Pharmacol Ther.* 1982;32(6):769-780.

36. Ottosson F, Ericson U, Almgren P, et al. Dimethylguanidino valerate: a lifestyle-related metabolite associated with future coronary artery disease and cardiovascular mortality. *J Am Heart Assoc.* 2019;8:e012846.

37. O'Sullivan JF, Morningstar JE, Yang Q, et al. Dimethylguanidino valeric acid is a marker of liver fat and predicts diabetes. *J Clin Invest.* 2017;127:4394-4402.

38. Han Z, Wyche JH. Guanosine induces necrosis of cultured aortic endothelial cells. *Am J Pathol.* 1994;145(2):423-427.

39. Vedder D, Walrabenstein W, Heslinga M, et al. Dietary interventions for gout and effect on cardiovascular risk factors: a systematic review. *Nutrients.* 2019;11:1-19.

40. Sauer N, Mosenthin R, Bauer E. The role of dietary nucleotides in single-stomached animals. *Nutr Res Rev.* 2011;24(1):46-59.

41. Furuhashi M. New insights into purine metabolism in metabolic diseases: role of xanthine oxidoreductase activity. *Am J Physiol Endocrinol Metab.* 2020;319(5): E827-E834.

42. Lee JB, Radhi M, Cipolla E, et al. A novel nucleoside rescue metabolic pathway may be responsible for therapeutic effect of orally administered cordycepin. *Sci Rep.* 2019; 9(1):15760-15812.

43. Karijolich J, Yu YT. Converting nonsense codons into sense codons by targeted pseudouridylation. *Nature.* 2011;474(7351):395-398.

44. Wu G, Xiao M, Yang C, Yu YT. U2 snRNA is inducibly pseudouridylated at novel sites by Pus7p and snR81 RNP. *EMBO J.* 2011;30(1):79-89.

45. Carlile TM, Rojas-Duran MF, Zinshteyn B, Shin H, Bartoli KM, Gilbert WV. Pseudouridine profiling reveals regulated mRNA pseudouridylation in yeast and human cells. *Nature.* 2014;515(7525):143-146.

46. Malik R, Chauhan G, Traylor M, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet.* 2018;50(4):524-537.

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

48. NINDS Stroke Genetics Network (SiGN), International Stroke Genetics Consortium (ISGC). Loci associated with ischaemic stroke and its subtypes (SiGN): a genome-wide association study. *Lancet Neurol.* 2016;15:174-184.

49. Vojinovic D, Kalaoja M, Trompet S, et al. Association of circulating metabolites in plasma or serum and risk of stroke. *Neurology.* 2021;96:e1110-e1123.

50. Sun D, Tiedt S, Yu B, et al. A prospective study of serum metabolites and risk of ischemic stroke. *Neurology.* 2019;92(16):e1890-e1898.