INTRODUCTION

Alzheimer’s disease (AD) is the most common form of dementia and is characterized by significant impairment in memory, behavioral changes and gradual loss of functional autonomy. The prevalence of AD dementia is as high as 20–30% in persons aged 75–84 years, and up to 50% in individuals aged ≥85 years. When 2800 subjects who were free of dementia were followed for 29 years in the Framingham Heart Study, the lifetime risk for dementia was reported to be 1 in 5 for women and 1 in 10 for men. There is currently no known cure or preventive intervention. Cognitive function, including memory and executive function, is influenced by both genetic and environmental factors. The human metabolome is a reflection of the interaction between genes and the environment, and studies examining the relationship between metabolomic profiles and cognitive function may lead to the development of biomarkers used to detect cognitive decline or AD before clinical diagnostic criteria for impairment are met. In this context, Mapstone et al. have recently reported that a set of 10 lipids identified in a metabolomics screen in peripheral blood could be used with 90% accuracy to predict conversion from normal cognitive status to amnestic mild cognitive impairment (MCI) or AD dementia over a 2–3-year period in adults aged ≥70 years. Several other investigators have also found significant alterations in metabolic profiles in comparisons of patients with MCI and AD dementia to cognitively normal subjects. The goal of this study is to determine whether metabolites measured in serum in middle-aged African American adults are associated with cognitive function and cognitive change in the Atherosclerosis Risk in Communities (ARIC) study. African Americans are affected disproportionately with AD dementia, therefore, this investigation may also provide insight into the biological basis of this health disparity.

MATERIALS AND METHODS

The ARIC Study

The ARIC Study is a prospective longitudinal investigation of the development of atherosclerosis and its clinical sequelae, in which 15792 individuals aged 45–64 years were enrolled at the baseline examination. A detailed description of the ARIC study has been reported previously. At the inception of the study in 1987–1989, the participants were selected by probability sampling from four communities in the United States: Forsyth County, North Carolina; Jackson, Mississippi (African Americans only); University of Texas Health Science Center at Houston, Houston, TX, USA; and 5 Division of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA; 6 Department of Medicine, Division of Geriatrics, University of Texas Health Science Center at Houston, Houston, TX, USA; 7 Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA; 8 Department of Biostatistics, Gillings School of Public Health, University of North Carolina, Chapel Hill, NC, USA and 9 Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA. Correspondence: Dr E Boerwinkle, Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, 1200 Pressler Street, Houston, TX 77030, USA.

Received 5 August 2016; revised 5 April 2017; accepted 20 April 2017

ORIGINAL ARTICLE

Metabolomics and cognition in African American adults in midlife: the atherosclerosis risk in communities study

J Bressler, B Yu, TH Mosley, DS Knopman, RF Gottesman, A Alonso, AR Sharrett, LM Wruck and E Boerwinkle

Clinical studies have shown alterations in metabolic profiles when patients with mild cognitive impairment and Alzheimer’s disease dementia were compared to cognitively normal subjects. Associations between 204 serum metabolites measured at baseline (1987–1989) and cognitive change were investigated in 1035 middle-aged community-dwelling African American participants in the biracial Atherosclerosis Risk in Communities (ARIC) Study. Cognition was evaluated using the Delayed Word Recall Test (DWRT; verbal memory), the Digit Symbol Substitution Test (DSST; processing speed) and the Word Fluency Test (WFT; verbal fluency) at visits 2 (1990–1992) and 4 (1996–1998). In addition, Cox regression was used to analyze the metabolites as predictors of incident hospitalized dementia between baseline and 2011. There were 141 cases among 1534 participants over a median 17.1-year follow-up period. After adjustment for established risk factors, one standard deviation increase in N-acetyl-1-methylhistidine was significantly associated with greater 6-year change in DWRT scores (β = 0.66 words; P = 3.65 × 10⁻³). Two metabolites (one unnamed and a long-chain omega-6 polyunsaturated fatty acid found in vegetable oils (docosapentaenoate (DPA, 22:5 n-6)) were significantly associated with less decline on the DSST (DPA: β = 1.25 digit-symbol pairs, P = 9.47 × 10⁻³). Two unnamed compounds and three sex steroid hormones were associated with an increased risk of dementia (all P < 3.9 × 10⁻³). The association of 4-androstene-3beta, 17beta-diol disulfate 1 with dementia was replicated in European Americans. These results demonstrate that screening the metabolome in midlife can detect biologically plausible biomarkers that may improve risk stratification for cognitive impairment at older ages.

Translational Psychiatry (2017) 7, e1173; doi:10.1038/tp.2017.118; published online 18 July 2017
suburbs of Minneapolis, Minnesota; and Washington County, Maryland. Four examinations were carried out at 3-year intervals (examination 1, 1987–1988; examination 2, 1990–1992; examination 3, 1993–1995; examination 4, 1996–1998). Subjects were contacted annually to update their medical histories between examinations. Cognitive testing was performed at examinations 2 and 4. Metabolomics profiles were available for 1977 African American study participants. Cognitive function at the baseline examination was examined in 1577 men and women after excluding individuals who had an unknown history of definite or probable stroke or a history of physician-diagnosed stroke prior to visit 2 (n = 51), did not attend the second clinical examination (n = 297), were missing test scores for all three cognitive tests (n = 50) or were missing information on educational attainment (n = 2). Among the participants with cognitive test scores obtained at both examinations, individuals were not included in the analysis of 6-year cognitive change if they had an unknown history of definite or probable stroke or a history of physician-diagnosed stroke prior to visit 2 (n = 51) or between visits 2 and 4 (n = 74), or did not attend the second (n = 297) or fourth (n = 361) clinical examination. Additional exclusions were made for missing cognitive data for all three neuropsychological tests at either visit 2 (n = 45) or visit 4 (n = 112), or for missing information concerning the highest level of education attained (n = 2); the final study sample consisted of 1035 participants. Individuals hospitalized for dementia through the end of 2011 were identified using ICD-9 codes listed in the hospital discharge record (AD (331.0); vascular dementia (290.4); or other forms of dementia (290.0, 290.1, 290.2, 290.3, 290.9, 294.1, 294.2, 294.3, 294.9, 331.2, 331.3, 331.4, 331.5, 331.6, 331.9) a copy of all hospital records of the study participants. The study participants were excluded from the analysis of incident hospitalized dementia if they had prevalent stroke at visit 1 (n = 67), were missing information about the number of years of education completed (n = 5) or if follow-up time was missing (n = 371), leaving a total of 1534 individuals. Written informed consent was provided by all study participants, and the study design and methods were approved by the institutional review boards at the four collaborating institutions: University of Mississippi Medical Center Institutional Review Board (Jackson Field Center); Wake Forest University Health Sciences Institutional Review Board (Forsyth County Field Center); University of Minnesota Institutional Review Board (Minnesota Field Center); and the Johns Hopkins School of Public Health Institutional Review Board (Washington County Field Center).

Cognitive tests

Cognitive function was assessed by three neuropsychological tests at the second and fourth clinical examinations that have been described previously: (1) the Delayed Word Recall Test (DWRT) is a test of verbal memory requiring recall of a word list after a short delay (score range 0–10); (2) the Digit Symbol Substitution Test (DSST) is a subtest of the Wechsler Adult Intelligence Scale-Revised involving timed translation of numbers to symbols in 90 s using a key, and measures psychomotor performance (score range 0–93); and (3) the Word Fluency Test (WFT) is a test of executive function. The score is the number of correct words produced beginning with F, A and S. For all of the neuropsychological tests, lower scores indicate a lower measure of cognition. Six-year change in cognitive function was analyzed as the difference between the test score obtained at visit 4 and the test score obtained at visit 2 for each test.

Clinical and laboratory measurements

The clinical and laboratory measurements used for this study were assessed during the first clinical examination for the analyses of incident hospitalized dementia, and during the second clinical examination for the analyses of cognitive change. The detection of proteinuria was estimated glomerular filtration rate (eGFR), which were evaluated at the first examination. Education was included as a covariate in regression models as an ordinal variable based on the highest level attained (≤11 years; 12–16 years; >16 years). Serum creatinine was measured using a Jaffé method and calibrated to nationally representative estimates as described previously. GFR was estimated based on serum creatinine using the equation:

\[ \text{eGFR} = \frac{186 \times \text{serum creatinine}^1.154 \times \text{age}^{-0.203}}{0.742} \]

Plasma total cholesterol, triglycerides and low-density lipoprotein (LDL) cholesterol were measured as previously described. Hypertension was defined by diastolic blood pressure ≥ 90 mm Hg, systolic blood pressure ≥ 140 mm Hg or use of antihypertensive medication. The prevalence of diabetes was defined using a fasting glucose level ≥ 7.0 mmol l⁻¹, a nonfasting glucose level ≥ 11.1 mmol l⁻¹ and/or self-reported physician diagnosis or treatment for diabetes. Body mass index (BMI) was calculated as weight in kilograms/(height in meters)². Information on cigarette smoking and alcohol consumption was obtained using an interviewer-administered questionnaire, and smoking and drinking status were classified as current, former or never. Usual intake of alcohol in grams/week was calculated for current drinkers, the usual weekly intake was set to zero for former or never drinkers. Genotyping of apolipoprotein E (APOE) polymorphisms at codons 112 and 158 (ref. 31) was performed using the TaqMan system (Thermo Fisher Scientific, Waltham, MA, USA) to generate the six APOE genotypes. The genotype call rate, or the percentage of samples to which a genotype was assigned, was determined prior to exclusion of individuals from the analysis and was 92.5%.

Metabolomics

Metabolomic profiles were measured in a subsample of 1977 randomly selected African American study participants from the Jackson, Mississippi field center who had given consent for use of genetic information as previously described. Metabolites were detected and quantified by Metabolon (Durham, NC, USA) in serum isolated from individuals who had fasted ≥ 8 h before the first clinical examination using an untargeted, gas chromatography/mass spectrometry, and a liquid chromatography–mass spectrometry-based method. Instrument variability was determined by calculating the median of the median of all technical replicate values that were added to each sample prior to injection into the mass spectrometers. Overall process variability was determined by calculating the median relative s.d. for all endogenous metabolites (that is, non-instrument standards) present in 100% of the technical replicate samples created from a homogeneous pool of human plasma. As the measurements span multiple days, a data-normalization step was performed to correct variations resulting from instrument tuning differences between days. Metabolites were compared to library entries of purified standards that included retention times, molecular weights, preferred adducts, in-source fragments and associated fragmentation spectra of the intact parent ions, or to recurrent unknown compounds. After laboratory quality-control procedures described in detail elsewhere, this approach yielded a total of 602 metabolites, including 361 named compounds and 241 unnamed compounds that did not have a chemical standard. The unknown compounds are designated by X followed by numbers (e.g., X-12345) by Metabolon. After conducting a repeatability study to determine the stability of metabolites in two serum samples collected 4–6 weeks apart from 60 individuals, 204 metabolites were selected that met the criteria of a reliability coefficient ≥ 0.6 and < 80% of values that were either missed or below the limit of detection. The reliability coefficient is an intraclass-correlation coefficient calculated as the ratio of between-individual variance and the sum of between-individual variance and within-individual variance over time. Using the intraclass-correlation coefficient, a group of metabolites was defined where the between-person variance accounts for most of the variability in metabolite concentrations, whereas within-person variance remains relatively low, thus optimizing their usefulness for risk assessment.

Statistical analysis

The 204 metabolites described above were divided into two groups for statistical analysis (Supplementary Table S1). The first group was composed of 187 metabolites (108 named and 79 unnamed compounds) that were analyzed as continuous variables and had values above the limit of detection ≥ 50% of samples; values below the limit of detection were assigned the lowest detected value for that metabolite in all samples. The lowest detected values for the metabolites found to be significantly associated with incident hospitalized dementia or interindividual variation in performance on neuropsychometric tests in this study are shown in Supplementary Table S2. Each of the 187 metabolites was centered by its mean and scaled by its s.d. prior to the analysis. A second group of 17 metabolites (10 named and 7 unnamed) with 50–80% of values that were below the limit of detection were analyzed as ordinal variables with the levels specified as follows: 1 = below the limit of detection; 2 = detected value (M); 3 = detected value (M) + 0.5 × lower detection limit; 4 = detected value (M) + 1 × lower detection limit; 5 = detected value (M) + 1.5 × lower detection limit; 6 = detected value (M) + 2 × lower detection limit. Proportions, means and s.d.s were calculated for clinical and demographic characteristics for individuals categorized by incident dementia case status, and for all individuals for the analyses of cognitive function. Groups were compared using X²-tests for
RESULTS

The clinical and demographic characteristics of the study sample are summarized in Table 1 for individuals included in the analyses of incident hospitalized dementia who were free of dementia and self-reported stroke or transient ischemic attack at baseline. Among the 1534 participants, 9.2% (n = 141) were hospitalized with dementia during a median follow-up period of 17.1 years. Participants who developed dementia were significantly more likely to be older at baseline, have had 11 years or less of formal education, have hypertension or diabetes, have a lower eGFR and to bear at least one copy of the APOE ε4 allele that has been reproducibly associated with increased risk of AD41–43 and variation in cognitive function in non-demented community-dwelling adults35,44–47 (P < 0.005). Individuals without prevalent stroke at visit 1 or incident clinical stroke between visits 1 and 4 were included in the analyses of baseline cognitive function and 6-year cognitive change, and the characteristics of this study sample are shown in Table 2. As a randomly selected subsample of African American ARIC study participants with metabolomics profiles were included in this analysis, the clinical and demographic characteristics of those with and without available data were compared (Supplementary Table S3). Individuals with metabolomics data were significantly younger and less likely to be male or have diabetes or hypertension, and also had a higher eGFR than ARIC participants who were not chosen for measurements of the metabolome.

Two unnamed compounds (X-11423, X-11491) were significantly associated with incident hospitalized dementia in the minimally adjusted Cox proportional hazards model but were no longer associated with susceptibility to dementia after further adjustment for a panel of established risk factors, whereas another metabolite implicated in sex steroid metabolism (pregnen-diol disulfate and 4-androsten-3 beta, 17 beta-diol disulfate) and an unnamed compound (X-12851) were significantly associated with elevated dementia risk only after adjustment for all covariates (Table 3). Three metabolites were significantly associated with increased hospitalized dementia using both the minimally adjusted and full Cox proportional hazards regression models including two named compounds involved in sex steroid metabolism (pregnen-diol disulfate and 4-androsten-3 beta, 17 beta-diol disulfate) and one metabolite with unknown structural identity (X-11440). Secondary analyses stratified by gender revealed that the two androgen sulfates were either significantly or marginally significantly associated with incident dementia in men (Table 4), although there was no association observed in women. The associations appeared to be similar in men and women for pregnen-diol disulfate.

In the analyses of 6-year cognitive change, the N-acetylated amino acid N-acetyl-T-methylhistidine was associated with greater decline in scores on the DWRT, and two metabolites

---

Table 1. Clinical and demographic characteristics stratified by incident hospitalized dementia status

| Characteristic | All; N = 1534; N (%) | N\(^a\) Incident dementia; N = 141; N (%) | N\(^a\) Non-case; N = 1393; N (%) | P\(^b\) |
|---------------|---------------------|------------------------------------------|-------------------------------|-------|
| Male          | 559 (36.4)          | 53 (37.6)                                | 506 (36.3)                    | 0.766 |
| Female        | 975 (63.6)          | 88 (62.4)                                | 887 (63.7)                    | 0.002 |
| Education     |                     |                                         |                               |       |
| ≤ 11 Years    | 659 (43.0)          | 80 (56.7)                                | 579 (41.6)                    |       |
| > 11 Years and ≤ 16 years | 418 (27.2) | 30 (21.3)                               | 388 (28.7)                     |       |
| > 16 Years    | 457 (29.8)          | 31 (22.0)                                | 426 (30.6)                    |       |
| Current smokers | 465 (30.4) | 35 (24.8)                               | 430 (30.9)                    | 0.133 |
| Hypertension  | 862 (56.5)          | 94 (67.1)                                | 768 (55.4)                    | 0.008 |
| Diabetes      | 269 (17.5)          | 39 (27.7)                                | 230 (16.5)                    | 0.001 |
| APOE (at least 1 ε4 allele) | 591 (43.2) | 81 (61.8)                               | 727 (58.8)                    | < 0.001 |
| APOE (ε4/ε4)  | 63 (4.4)            | 15 (11.4)                                | 48 (3.9)                      | < 0.001 |

| Mean (s.d.)   | Mean (s.d.)         | Mean (s.d.)     | Mean (s.d.)     |
|---------------|---------------------|-----------------|-----------------|
| Age (years)   | 53.4 (5.8)          | 56.4 (5.6)      | 53.1 (5.7)      | < 0.001 |
| BMI (kg/m\(^2\)) | 29.8 (6.1) | 30.1 (6.5)     | 29.8 (6.1)      | 0.532  |
| LDL cholesterol (mmol l\(^{-1}\)) | 3.6 (1.1) | 3.6 (1.2)      | 3.6 (1.1)       | 0.677  |
| Ethanol (g per week) | 32.8 (104.2) | 34.9 (123.3) | 32.6 (102.1) | 0.809  |
| eGFR\(_{CKD-EPI}\) (ml min\(^{-1}\) per 1.73 m\(^2\)) | 104.0 (18.4) | 103.3 (18.9) | 104.4 (18.4) | 0.011  |

Abbreviations: APOE, apolipoprotein E; BMI, body mass index; CKD-EPI, Chronic Kidney Disease Epidemiology Collaboration; eGFR, estimated glomerular filtration rate; LDL, low-density lipoprotein cholesterol; N, number. \(\text{N}a\)Number of participants with data for clinical and demographic characteristics. \(\text{N}b\)Individuals with and without incident hospitalized dementia were compared using \(\chi^2\)-tests for categorical variables and \(t\)-tests for continuous variables.

categorical variables and \(t\)-tests for continuous variables. Cox proportional hazards models were used to estimate the hazard ratio (HR) for developing hospitalized dementia; the median follow-up period was 17.1 years (25th percentile = 10.8 years; 75th percentile = 21.2 years). All HRs were calculated and reported per one s.d. increase for continuous variables, and per one category change for the categorical variables. Linear regression models were applied separately for each of the three cognitive tests (DWRT, DSST and WFT) for each metabolite to estimate its association with baseline cognitive function and 6-year cognitive change. Two multivariable models were used to evaluate the relationship between individual metabolites and both incident dementia and cognitive function. Model 1 was adjusted for age, gender, education defined as an ordinal variable (≤11 years; 12–16 years; > 16 years) and eGFR, whereas Model 2 was adjusted for the covariates in Model 1 with the addition of potential confounding variables including diabetes and hypertension case status, BMI, LDL cholesterol, current smoking, alcohol intake and APOE genotype (carriage of at least one ε4 allele) after exclusion of individuals with missing data.\(^{27,28}\) The Dubey/Armitage-Parmar algorithm,\(^{29}\) a modified stepwise Bonferroni procedure, was used to correct for multiple comparisons as previously described,\(^{18}\) and a two-tailed \(P\)-value of 3.9 \times 10^{-4} was considered statistically significant for each individual test. A power calculation was performed using the fixed sample size for each analysis and a pre-specified effect size reported from previous studies. For incident hospitalized dementia, more than 90% power will be reached if the sample size is 1534 (141 cases), the HR is 1.7,\(^{38,39}\) the correlation between metabolites is 0.3 and the alpha level is 3.9 \times 10^{-4}. For the tests of association of baseline neurocognitive test scores and 6-year score change, assuming a small effect size of 0.02,\(^{35}\) five predictors (Model 1) and an alpha level of 3.9 \times 10^{-4}, there will be more than 90% and 80% power with sample sizes of 1577 and 1035, respectively. All statistical analyses were performed using R.\(^{20}\)
dementia, the omega-6 fatty acid DPA and less decline in processing speed, and an N-acetylated amino acid and greater decline in verbal memory were identified in a large-scale screen of serum metabolites in middle-aged African American participants in the ARIC study. Although the phenotype examined in this study was incident hospitalized dementia that encompassed AD and vascular dementia as well as other forms of dementia, several investigators have previously observed significant alterations in metabolic profiles in cerebrospinal fluid,8,38–51 or blood6,59–64 in comparisons of patients with MCI or AD to cognitively normal subjects. Because sample acquisition is less invasive, identification of blood-based biomarkers would be advantageous for clinical diagnosis, AD-screening programs and monitoring of response to therapy. Mapstone et al. recently reported that a set of 10 lipids identified in a metabolomics screen in peripheral blood could be used with 90% accuracy to predict conversion from normal cognitive status to amnestic MCI or AD dementia over a 2–3-year period in 525 community-dwelling adults ≥70 years.6 Although these findings were not replicated in three independent cohorts including the ARIC study,52,53 higher concentrations of one of the 10 phospholipids was significantly associated with a decreased prevalence of dementia.54 Alterations in lipid metabolic pathways detected in the blood of patients with AD dementia have also been described by other investigators.9–14 For example, significantly decreased plasma concentrations of three phosphatidylincholines were observed when individuals with AD dementia were compared to controls from the King’s College London Dementia Case Register and the AddNeuroMed study, and six ether-containing phosphatidylincholines and sphingomyelins were associated with abnormal levels of ApoA4 in the cerebrospinal fluid in participants in the Alzheimer’s Disease Neuroimaging Initiative-1 cohort.14 A group of 154 metabolites associated with AD dementia was identified in the Mayo Clinic Study of Aging including etiocholanolone sulfate (5-androstan-3-alpha-ol-17-one sulfate) and testosterone sulfate, two metabolites of testosterone.7 Most of the metabolites described in the studies above were not included in the panel analyzed in the ARIC study, with the exception of uridine,7 creatine,7 dimethylglycine,7 tryptophan,49 2-hydroxysobuturate7 and valine.55 Providing evidence in support of an initial discovery in Alzheimer’s Disease Neuroimaging Initiative-1, higher levels of the branched-chain amino acid valine were significantly positively associated with general cognitive ability and conferred a decreased risk of incident AD dementia over a median 9.7-year follow-up period in 2505 dementia-free participants in the population-based Rotterdam Study.12 However, among these metabolites, only dimethylglycine was nominally associated with incident hospitalized dementia using the fully adjusted Cox regression model (P = 0.0252).

In the ARIC study, profiling of serum metabolites by an untargeted mass spectrometry-based method revealed associations between increased risk of incident hospitalized dementia and three metabolites involved in sex steroid metabolism including 4-androsten-3-beta, 17 beta-diol disulfate 1 (sulfate of 4-androstenenediol; HR = 1.25, P = 1.44 × 10−5), 5 alpha-androstan-3 beta, 17 beta-diol disulfate (HR = 1.26, P = 1.64 × 10−5) and pregnen-diol disulfate (HR = 1.35, P = 5.49 × 10−4) and two unnamed metabolites after adjustment for a panel of established risk factors for cognitive function.55 The regression models were also adjusted for eGFR, as the kidney freely filters molecules weighing less than 10 000 Da from the blood that are then reabsorbed, catabolized and/or secreted so that interindividual variation in renal function can influence metabolite concentrations.55 The association with 4-androsten-3-beta, 17 beta-diol disulfate 1 was replicated in European American ARIC study participants, providing stronger evidence that levels of this hormone influence susceptibility to dementia. As both androgens and estrogens have been shown to exert neuroprotective effects

| Table 2. Clinical and demographic characteristics—baseline cognitive function and 6-year cognitive change |
| --- |
| Characteristic | N | N (%) | Cognitive change | N (%) |
| --- | --- | --- | --- | --- |
| Male | 1577 | 538 (34.1) | 1035 | 341 (33.0) |
| Female | 1039 (65.9) | 694 (67.0) | 1035 | 341 (33.0) |
| Education | 1577 | 1035 | 1035 | 341 (33.0) |
| ≤11 Years | 615 (39.0) | 364 (35.2) | 364 (35.2) | 283 (27.3) |
| >11 Years and ≤16 years | 432 (27.4) | 283 (27.3) | 283 (27.3) | 283 (27.3) |
| >16 Years | 530 (33.6) | 388 (37.5) | 388 (37.5) | 283 (27.3) |
| Current smokers | 1575 | 385 (24.4) | 1034 | 214 (20.7) |
| Hypertension | 1570 | 810 (51.6) | 1028 | 502 (48.8) |
| Diabetes | 1561 | 344 (22.0) | 1023 | 209 (20.4) |
| APOE (at least 1 ε4 allele) | 1404 | 492 (22.2) | 912 | 392 (43.0) |
| APOE (≤4/≤4) | 1404 | 63 (4.5) | 912 | 47 (5.2) |

| Mean ± s.d. | Mean ± s.d. |
| --- | --- |
| Cognitive tests | --- |
| DWRT (words) | 1575 | 6.2 (1.6) | 1034 | 6.3 (1.6) |
| DSST (digit-symbol pairs) | 1565 | 31.4 (13.6) | 1030 | 32.9 (13.3) |
| WFT (words) | 1570 | 27.9 (13.3) | 1031 | 29.3 (13.2) |
| 6-Year change (visit 2–visit 4) | --- |
| DWRT (words) | --- | --- | 1034 | −0.25 (1.72) |
| DSST (digit-symbol pairs) | --- | --- | 1018 | −2.06 (8.82) |
| WFT (words) | --- | --- | 1026 | −0.89 (8.41) |

Abbreviations: APOE, apolipoprotein E; BMI, body mass index; CKD-EPI, Chronic Kidney Disease Epidemiology Collaboration; DSST, Digit Symbol Substitution Test; DWRT, Delayed Word Recall Test; eGFR, estimated glomerular filtration rate; LDL cholesterol, low-density lipoprotein cholesterol; N; number; WFT, Word Fluency Test. Number of participants with data for clinical and demographic characteristics.
Table 3. Significant associations of metabolites with incident hospitalized dementia

| Metabolitea | Pathway | Platform | Model 1 (N = 1534) |
|-------------|---------|----------|---------------------|
|             |         |          | HR                  |
|             |         |          | 95% CI              |
|             |         |          | P                  |
|             | Mean (s.d.) | Median | Q1 | Q3 |
| 4-androsten-3 beta, 17 beta-diol sulfate 1 | Sterol/steroid | LC/MS neg | 1.27 | 1.14, 1.41 | 1.50 × 10⁻⁵ | 0.025(1.05) | −0.244 | −0.410 | 0.064 |
| 4 α-androstan-3 beta, 17 beta-diol disulfate | Sterol/steroid | LC/MS neg | 1.25 | 1.11, 1.42 | 2.94 × 10⁻⁴ | 0.026(1.04) | −0.278 | −0.599 | 0.280 |
| 5 α-androstan-3 beta, 17 beta-diol disulfate | Sterol/steroid | LC/MS neg | 1.19 | 1.08, 1.31 | 3.83 × 10⁻⁴ | 0.032(1.05) | −0.230 | −0.516 | 0.225 |
| X-11440 | None | LC/MS neg | 1.33 | 1.15, 1.53 | 1.10 × 10⁻⁴ | 0.029(1.02) | −0.214 | −0.638 | 0.360 |
| X-11423 | None | LC/MS neg | 1.32 | 1.14, 1.54 | 2.40 × 10⁻⁴ | 0.009(1.02) | −0.106 | −0.328 | 0.176 |
| X-11491 | None | LC/MS neg | 1.19 | 1.07, 1.32 | 9.03 × 10⁻⁴ | 0.006(1.02) | −0.280 | −0.490 | 0.092 |
| 5 α-androstan-3 beta, 17 beta-diol disulfate | Sterol/steroid | LC/MS neg | 1.23 | 1.08, 1.39 | 1.67 × 10⁻³ | −0.018(0.966) | −0.333 | −0.502 | −0.029 |

Abbreviations: APOE, apolipoprotein E; BMI, body mass index; CI, confidence interval; CKD-EPI, Chronic Kidney Disease Epidemiology Collaboration; eGFR, estimated glomerular filtration rate; HR, hazard ratio; LC/MS neg, liquid chromatography/mass spectrometry-negative ion mode; LDL cholesterol, low-density lipoprotein cholesterol; Q1, quartile 1; Q3, quartile 3. Model 1, adjusted for age, gender, education, eGFR₂₅₂₀₀, Model 2, Model 1 + diabetes, hypertension, BMI, LDL cholesterol, current smoking, alcohol intake, APOE genotype (at least 1 ε4 allele/no ε4 allele). *All metabolites in Group 1, analyzed as continuous variables. **Threshold for statistical significance < 3.9 × 10⁻⁵; bold P-values are statistically significant.

Table 4. Significant associations of metabolites with incident hospitalized dementia stratified by gender

| Metabolitea | Pathway | Platform | Women (N = 1348) | Men (N = 508) |
|-------------|---------|----------|-----------------|--------------|
|             |         |          | HR | 95% CI | P      | Mean (s.d.) | Median | Q1 | Q3 | HR | 95% CI | P      | Mean (s.d.) | Median | Q1 | Q3 |
| 4-androsten-3 beta, 17 beta-diol sulfate 1 | Sterol/steroid | LC/MS neg | 1.25 | 1.11, 1.40 | 1.44 × 10⁻⁴ | 0.034(1.08) | −0.241 | −0.409 | 0.074 |
| 4 α-androstan-3 beta, 17 beta-diol disulfate | Sterol/steroid | LC/MS neg | 1.26 | 1.12, 1.42 | 1.64 × 10⁻⁴ | 0.017(1.02) | −0.274 | −0.488 | 0.105 |
| 5 α-androstan-3 beta, 17 beta-diol disulfate | Sterol/steroid | LC/MS neg | 1.35 | 1.17, 1.56 | 5.59 × 10⁻⁵ | 0.031(1.04) | −0.278 | −0.597 | 0.285 |
| X-11440 | None | LC/MS neg | 1.37 | 1.18, 1.60 | 4.52 × 10⁻⁵ | 0.038(1.04) | −0.208 | −0.638 | 0.376 |
| X-11423 | None | LC/MS neg | 1.39 | 1.12, 1.72 | 2.66 × 10⁻³ | 0.002(0.815) | −0.104 | −0.325 | 0.188 |
| X-11491 | None | LC/MS neg | 1.14 | 1.03, 1.27 | 0.0101 | 0.041(1.04) | −0.223 | −0.508 | 0.238 |
| X-12851 | None | LC/MS neg | 1.26 | 1.12, 1.43 | 1.92 × 10⁻⁴ | −0.010(0.976) | −0.331 | −0.502 | −0.025 |

Abbreviations: APOE, apolipoprotein E; BMI, body mass index; CI, confidence interval; CKD-EPI, Chronic Kidney Disease Epidemiology Collaboration; eGFR, estimated glomerular filtration rate; HR, hazard ratio; LC/MS neg, liquid chromatography/mass spectrometry-negative ion mode; LDL cholesterol, low-density lipoprotein cholesterol; Q1, quartile 1; Q3, quartile 3. Model 1, adjusted for age, education, eGFR₂₅₂₀₀, diabetes, hypertension, BMI, LDL cholesterol, current smoking, alcohol intake, APOE genotype (at least 1 ε4 allele/no ε4 allele). *All metabolites in Group 1, analyzed as continuous variables. **Threshold for statistical significance < 3.9 × 10⁻⁵; bold P-values are statistically significant. *Adjusted for age, education, eGFR₂₅₂₀₀, diabetes, hypertension, BMI, LDL cholesterol, current smoking, alcohol intake, APOE genotype (at least 1 ε4 allele/no ε4 allele).
### Table 5. Significant associations of metabolites with baseline cognitive function and 6-year cognitive change

| Test; baseline | Metabolite | Pathway | Platform | Model 1 (N = 1565) | Model 2 (N = 1355) | 6-Year change | Model 1 (N = 1028) | Model 2 (N = 883) | Model 1 (N = 1012) | Model 2 (N = 870) |
|---------------|------------|---------|----------|---------------------|---------------------|---------------|---------------------|---------------------|---------------------|---------------------|
| DSST          | Mannose    | Fructose, mannose | GC/MS | \( \beta \) | s.e. | \( P^b \) | Mean (s.d.) | Median | Q1 | Q3 | Mean (s.d.) | Median | Q1 | Q3 |
|               | N-acetyl-1-methylhistidine | Histidine metabolism | LC/MS pos | -1.174 | 0.286 | \( 4.37 \times 10^{-5} \) | -0.052 (0.904) | -0.211 | -0.525 | 0.123 |
|               | X-12846    | None | LC/MS neg | 6.591 | 1.681 | \( 9.24 \times 10^{-5} \) | -0.046 (0.166) | -0.065 | -0.229 | 0.035 |
|               | Andro steroid monosulfate 2 | Sterol/steroid | LC/MS neg | -1.131 | 0.295 | \( 1.33 \times 10^{-4} \) | -0.067 (0.877) | -0.328 | -0.644 | 0.156 |
|               | Glucose    | Glycolysis | GC/MS | -1.045 | 0.283 | \( 2.26 \times 10^{-4} \) | -0.042 (0.917) | -0.246 | -0.452 | -0.012 |
| DSST          | Mannose    | Fructose, mannose | GC/MS | -0.437 | 0.347 | 0.208 | -0.051 (0.917) | -0.210 | -0.522 | 0.116 |
|               | N-acetyl-1-methylhistidine | Histidine metabolism | LC/MS pos | -2.414 | 0.734 | \( 1.04 \times 10^{-3} \) | -0.056 (0.417) | -0.083 | -0.335 | 0.096 |
|               | X-12846    | None | LC/MS neg | -4.547 | 1.822 | 0.0127 | -0.043 (0.168) | -0.061 | -0.194 | 0.036 |
|               | Andro steroid monosulfate 2 | Sterol/steroid | LC/MS neg | -0.818 | 0.316 | \( 9.73 \times 10^{-3} \) | -0.051 (0.893) | -0.314 | -0.644 | 0.174 |
|               | Glucose    | Glycolysis | GC/MS | -0.306 | 0.355 | 0.390 | -0.040 (0.929) | -0.246 | -0.452 | -0.012 |

6-Year change

| DWRT          | N-acetyl-1-methylhistidine | Histidine metabolism | LC/MS pos | -0.504 | 0.170 | \( 3.18 \times 10^{-3} \) | -0.086 (0.333) | -0.096 | -0.362 | 0.085 |
|               | X-12844    | None | LC/MS neg | 0.977 | 0.290 | \( 7.84 \times 10^{-4} \) | -0.045 (0.966) | -0.113 | -0.726 | 0.527 |
| DSST          | Docosapentaenoate (n-6 DPA) | Essential fatty acid | LC/MS neg | 0.977 | 0.290 | \( 7.84 \times 10^{-4} \) | -0.045 (0.966) | -0.113 | -0.726 | 0.527 |
|               | X-12844    | None | LC/MS neg | 0.977 | 0.290 | \( 7.84 \times 10^{-4} \) | -0.045 (0.966) | -0.113 | -0.726 | 0.527 |

Abbreviations: \( \beta \), beta coefficient; APOE, apolipoprotein E; BMI, body mass index; CKD-EPI, Chronic Kidney Disease Epidemiology Collaboration; DPA, docosapentaenoate; DSST, Digit Symbol Substitution Test; DWRT, Delayed Word Recall Test; eGFR, estimated glomerular filtration rate; GC/MS, gas chromatography/mass spectrometry; LC/MS neg, liquid chromatography/mass spectrometry-negative ion mode; LC/MS pos, liquid chromatography/mass spectrometry-positive ion mode; Q1, quartile 1; Q3, quartile 3. As many of the same metabolites have recently been measured in 1553 European American ARIC participants, replication was sought for compounds that were identified using the fully adjusted models for both phenotypes (Supplementary Table S4). Among the available metabolites, 4-androsten-3-beta, 17-beta-dihol disulfate 1 was also found to be significantly associated with incident hospitalized dementia in individuals of European ancestry after application of the same exclusion criteria used for African Americans (clinical and demographic characteristics are shown in Supplementary Tables S5 and S6). Model 1, Adjusted for age, gender, education, eGFRCKD-EPI; Model 2, Model 1 + diabetes, hypertension, BMI, LDL cholesterol, current smoking, alcohol intake, APOE genotype (at least 1 ε4 allele/no ε4 allele). \(^a\)All metabolites in Group 1, analyzed as continuous variables. \(^b\)Threshold for statistical significance \( < 3.9 \times 10^{-4} \); bold \( P \)-values are statistically significant.
including prevention of amyloid beta deposition and tau hyperphosphorylation.\textsuperscript{56-60} Age-related depletion of sex steroid hormones might be expected to contribute to an increased susceptibility to neurodegenerative diseases. Accordingly, age-related decline in circulating estrogens and testosterone has previously been reported to be a risk factor for AD dementia or MCI in women and men, respectively, in some\textsuperscript{61-67} but not all studies.\textsuperscript{68} The relationship between steroid hormones and AD has also been evaluated directly in samples of neuronal tissue.\textsuperscript{69,70} Although these previous reports provide support for the suggestion that interindividual variation in the levels of the three sex steroid metabolites may be biologically linked to cognitive impairment, important caveats are that dementia as defined in this study was based on hospital ICD-9 codes and was not restricted to AD, and that even many cases of clinically diagnosed dementias including AD show a mixture of AD and vascular neuropathology at autopsy.\textsuperscript{71-73}

Four-androstenediol is a precursor of both testosterone and estradiol, and is used as a prohormone by men to elevate levels of serum and urinary testosterone.\textsuperscript{74,75} 5-alpha-androstan-3-beta, 17-beta-diol (3-betaAdiol) is a metabolite of the androgen dihydro-testosterone. Pregnen-diol disulfate has not been well characterized. An increase in the levels of sulfated intermediates in steroid metabolism in individuals with incident hospitalized dementia could reflect alterations in a number of biological processes. The major site of clearance and inactivation of steroids is the liver where conjugation with sulfuric acid is one mechanism whereby the compounds become water-soluble and can be excreted in the urine.\textsuperscript{76} Changes in the rate of steroid clearance in individuals with dementia may be reflected in the increased levels detectable in the serum. Another possibility is that, as hydrolysis of steroid sulfates by steroid sulfatase in the liver, kidney and other tissues can contribute to the pool of unconjugated androgens,\textsuperscript{77} as yet undescribed variation in this enzymatic reaction may also be a factor.

To our knowledge, this study is among the first to examine the association between the human metabolome and change in cognitive function among middle-aged adults who are not cognitively impaired using scores on standardized neuropsychological tests. Higher levels of an N-acetylated amino acid previously associated with lower eGFR in African American ARIC study participants were significantly associated with greater decline in scores on the DWT, a test of verbal declarative memory.\textsuperscript{59} Although lower eGFR has been shown to be associated with impairment in delayed word recall in patients with chronic kidney disease,\textsuperscript{78,79} N-acetyl-1-methylhistidine was significantly associated with memory performance after adjustment for eGFR, suggesting an effect on cognition that is independent of its role in renal function. The DSST is a test widely used to measure information-processing speed while an individual translates numbers to letters on a paper and pencil test. The rate of information-processing speed has been shown to decrease in older individuals,\textsuperscript{80,81} and low scores on the DSST indicative of poor performance in this domain have been associated with MCI and early-stage dementia.\textsuperscript{82,83} Higher n-6 DPA levels were associated with less decline in DSST scores over a 6-year period in ARIC study participants ($\beta = 1.25$, $P = 9.47 \times 10^{-5}$). DPA is a long-chain n-6 PUFA found in vegetable oil that can also be metabolized from the essential fatty acid linoleic acid (18:2 n-6) or from arachidonic acid (20:4 n-6) by a series of chain elongation and desaturation reactions.\textsuperscript{84} The higher serum DPA levels may indicate an increased rate of conversion from either linoleic acid or arachidonic acid to n-6 DPA in those with less pronounced DSST score change. Genetic variants in enzymes involved in the n-6 metabolic pathway have been associated with linoleic acid and arachidonic acid levels, and could potentially underlie the observed variation in cognitive status.\textsuperscript{85,86}

Dietary intake of omega-6 fatty acids has previously been shown to be associated with a lower risk of AD, MCI, cognitive decline and all-cause mortality.\textsuperscript{87-90} Although increased risk has also been reported.\textsuperscript{91-93} For example, intake of omega-6 polyunsaturated fat was inversely associated with incident clinically diagnosed AD in a biracial sample of 815 participants from the Chicago Health and Aging Project.\textsuperscript{87} And the odds ratio of MCI decreased as intake of omega-6 fatty acids increased in 1233 individuals in the Mayo Clinic Study of Aging.\textsuperscript{88} In one of the only studies to specifically address the association between n-6 DPA and cognitive status, erythrocyte DPA levels were higher in adults older than 65 years with MCI than in healthy controls.\textsuperscript{94} Of particular relevance to the results reported here, a higher ratio between total erythrocyte n-6-PUFAs and n3-PUFAs was significantly negatively correlated with baseline DSST scores in the Lothian Birth Cohort 1936 study.\textsuperscript{89} However, as all n-6 PUFAs were considered as a group, it is possible that individual components such as DPA exerted a range of effects in different directions that could not be distinguished in the assay used. Although there may be many reasons for the inconsistent results reported across studies, many relied on food frequency questionnaires or interviews\textsuperscript{87-89,92,93} that are subject to recall bias rather than on laboratory measurements of metabolites as reported here. Most importantly, both the neuropsychological test batteries used to evaluate cognition and the individual omega-6 fatty acids chosen for analysis either alone or as components of a composite measure often varied between cohorts, and a large-scale hypothesis-free screen of the metabolome was not undertaken except in the ARIC study. Further research optimally relying on the use of the same cognitive tests to assess change over a similar time period may help to clarify whether increased levels of serum n-6 DPA are consistently associated with less decline in processing speed in early middle age.

Taken together, these results demonstrate that screening the metabolome in midlife can be used to discover and prioritize biologically plausible biomarkers that may improve risk stratification for cognitive impairment at older ages. The strengths of the study include the prospective design and a large deeply phenotyped cohort that allowed the detection of an association between novel biomarkers and incident hospitalized dementia decades before it came to medical attention, as well as early cognitive change in non-demented community-dwelling adults. There are also limitations. Only one cognitive test was used to measure each of three cognitive domains, and cognitive change was analyzed over a short period in middle age when relatively few participants will have undergone substantial decline. Finally, there was only a single measurement of the serum metabolites, although rigorous quality-control procedures were undertaken to assure medium term reliability of the data, and the identity of some of the metabolites significantly associated with both phenotypes is currently unknown. While the association between a sex steroid hormone significantly associated with incident hospitalized dementia in African Americans was replicated in European American study participants, evaluation of the same metabolites in an independent replication sample of African Americans and across different ethnicities to address the generalizability of the findings is warranted.

**CONFLICT OF INTEREST**

Dr Knopman served as Deputy Editor for *Neurology* until October 2015 and received compensation from the American Academy of Neurology. He serves on a Data Safety Monitoring Board for Lundbeck Pharmaceuticals and for the DIAN study, and is an investigator in clinical trials sponsored by TauRX Pharmaceuticals, Lilly Pharmaceuticals and the Alzheimer’s Disease Cooperative Study. The remaining authors declare no conflict of interest.
ACKNOWLEDGMENTS

The Atherosclerosis Risk in Communities Study is carried out as a collaborative study supported by the National Heart, Lung, and Blood Institute contracts (HHSN262820100005C, HHSN262820100006C, HHSN262820100007C, HHSN262820100008C, HHSN262820100009C, HHSN262820100010C, HHSN262820100011C and HHSN262820100012C. We thank the staff and participants of the ARIC study for their important contributions. The metabolomics research was sponsored by the National Human Genome Research Institute (grant 3U01HG004402-02).
