



Published in final edited form as:

Alzheimers Dement. 2016 July ; 12(7): 815–822. doi:10.1016/j.jalz.2015.12.008.

Blood metabolite markers of preclinical Alzheimer's disease in two longitudinally followed cohorts of older individuals

Ramon Casanova, PhD¹, Sudhir Varma, PhD², Brittany Simpson, BA^{3,4}, Kim Min, MSci⁵, Yang An, MS³, Santiago Saldana, MS¹, Carlos Riveros, PhD⁶, Pablo Moscato, PhD⁶, Michael Griswold, PhD⁷, Denise Sonntag, PhD⁸, Judith Wahrheit, PhD⁸, Kristaps Klavins, PhD⁸, Palmi V Jonsson⁹, Gudny Eiriksdottir⁹, Thor Aspelund⁹, Lenore J Launer¹⁰, Vilmundur Gudnason⁹, Cristina Legido Quigley, PhD⁵, and Madhav Thambisetty, MD, PhD³

¹Wake Forest School of Medicine, Winston, Salem, NC, USA ²HiThru Analytics LLC, Laurel, MD, USA ³Clinical and Translational Neuroscience Unit, National Institute on Aging, National Institutes of Health, Baltimore, MD, USA ⁴School of Medicine, University of Mississippi Medical Center, Jackson, MS, US ⁵King's College, London, United Kingdom ⁶University of Newcastle, Callaghan, Australia ⁷Center of Biostatistics and Bioinformatics, University of Mississippi Medical Center, Jackson, MS, US ⁸BIOCRATES Life Sciences AG, Innsbruck, Austria ⁹Icelandic Heart Association, Kopavogur, Iceland ¹⁰Laboratory of Epidemiology, Demography and Biometry, National Institute on Aging, Bethesda, Maryland

Abstract

Recently, quantitative metabolomics identified a panel of 10-plasma lipids that were highly predictive of **conversion to Alzheimer's disease (AD)** in cognitively normal older individuals (N=28, area-under-the-curve; AUC=0.92, sensitivity/specificity of 90%/90%). We failed to replicate these findings in a substantially larger study from two independent cohorts - the Baltimore Longitudinal Study of Aging (BLSA, N=93, AUC=0.642, sensitivity/specificity of 51.6%/65.7%) and the Age, Gene/Environment Susceptibility-Reykjavik Study (AGES-RS, N=100, AUC=0.395, sensitivity/specificity of 47.0%/36.0%). **In analyses applying machine learning methods to all 187 metabolite concentrations assayed, we find a modest signal in the BLSA with distinct metabolites associated with the preclinical and symptomatic stages of AD, whereas the same methods gave poor classification accuracies in the AGES-RS samples.** We believe that ours is the largest blood biomarker study of **preclinical AD** to date. These findings underscore the importance of large-scale independent validation of index findings from biomarker studies with relatively small sample sizes.

Conflict of Interest: The authors confirm that they do not have any conflicts of interest to disclose.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

1. INTRODUCTION

Non-invasive and accurate peripheral biomarkers of preclinical Alzheimer's disease (AD) are a critical unmet need. The recent incorporation of cerebrospinal fluid (CSF) and imaging biomarkers into the diagnostic guidelines for AD, mild cognitive impairment (MCI) and preclinical AD represent a paradigm shift in the field [1–3]. The application of metabolomics technology for the discovery of AD biomarkers is receiving increasing attention [4]. As small metabolites represent the end result of cellular regulatory complexity, they are thought to be reliable proximal reporters of disease processes [5]. Several studies have reported on alterations in metabolite concentrations in blood between AD and healthy controls [6–12]. Most of these previous studies have relied upon relatively small sample sizes and different methodological approaches making comparative assessment of results difficult. Moreover, very few studies have focused attention on the discovery of predictive biomarkers of AD i.e. those indicative of greater risk of subsequent conversion to AD in older individuals who are cognitively normal. Such biomarkers would represent a substantial breakthrough as they would allow for the effective screening of large numbers of at-risk elderly and facilitate the testing of disease-modifying treatments in patients in very early stages of the AD disease process. If such biomarkers could accurately identify cognitively normal elderly at risk of subsequent AD, they would be of immediate clinical utility and merit use in routine clinical practice.

A recent study reported the discovery of a 10-metabolite panel in plasma that could discriminate cognitively normal older individuals who developed incident AD within 3 years (n=10; validation sample; age; 79.3±5.49 years) from healthy controls who remained cognitively normal (n=20; validation sample; age; 81.35±3.25 years) [13]. This panel was described as having impressive accuracy (0.92 area-under-the-receiver-operating-characteristic-curve; sensitivity/specificity; 90%/90%), suggesting considerable clinical utility of these analytes as antecedent biomarkers of memory impairment in cognitively normal individuals who will eventually develop AD.

Here, we test these index findings in a substantially larger sample from two well-characterized and longitudinally followed cohorts of older individuals from North America and Europe-the Baltimore Longitudinal Study of Aging (BLSA) [14] and the Age, Gene/Environment Susceptibility Study-Reykjavik (AGES-RS) [15] using the same targeted metabolomics platform utilized in the index study [13] (AbsoluteIDQ p180 assay, BIOCRATES, Life Science AG, Innsbruck, Austria). We examined serum metabolomic profiles at two time points for two groups – ‘non-converters’ who remained cognitively normal between 2 time points, approximately 5 years apart, and ‘converters’ who were cognitively normal at baseline and converted to AD within the same interval as the non-converters. In addition to testing whether we could replicate the previously reported findings by Mapstone and colleagues using their 10-metabolite panel [13], we also used a data-driven approach using machine-learning methods to analyze the entire targeted metabolomic data we acquired to examine whether other metabolite signatures could be identified as predictors of incident AD.

2. METHODS

2.1 BLSA and AGES-RS cohorts

Several previous publications have described details of the BLSA [14, 16, 17] and AGES-RS cohorts [15, 18]. A summary of participant characteristics and study procedures, including diagnostic approaches to AD/dementia in the two studies are included in supplementary material.

2.2 Serum samples

Serum samples were collected after **overnight fasting** in both BLSA and AGES-RS participants fasting (between 6 and 7 AM in BLSA and between 8AM-11 AM in AGES-RS). Details of pre-analytical procedures including sample storage are provided in supplementary material. BLSA participants who were cognitively normal at baseline and developed AD during follow up (i.e. ‘converters’: BLSA; n=93; baseline age 77.9 ± 6.5 years) were age- and sex-matched to participants who remained cognitively normal throughout follow up (i.e. ‘non-converters’: BLSA; n=99, baseline age 76.6 ± 6.7 years) in a case-control design (Table-1A). Two serial serum samples were analyzed from each participant in both the ‘converter’ and ‘non-converter’ groups (total number of serum samples assayed=384). Serial serum samples from the BLSA were obtained as follows:

- i. Baseline: 5 years (4.8 ± 1.2) prior to onset of cognitive impairment in individuals meeting consensus clinical diagnosis of incident AD (i.e. converters) and age-matched samples from healthy controls (i.e. non-converters).
- ii. Samples concurrent (0.69 ± 0.83 years) to onset of cognitive impairment in individuals meeting consensus clinical diagnosis of incident AD (i.e. converters) and age-matched samples from healthy controls (i.e. non-converters).

Serum samples from AGES-RS participants were obtained over the course of two separate waves of the study. Wave-1 was completed between 2002 and 2006 and wave-2, between 2007 and 2011. 100 participants who were cognitively normal during wave-1 and subsequently diagnosed with AD during wave-2 (i.e. ‘converters’, baseline age 78.18 ± 4.4) were age- and sex-matched 1:1 to 100 participants who remained cognitively normal over the course of both waves-1 and 2 (i.e. ‘non-converters’, baseline age 78.23 ± 4.4). Baseline and follow-up samples were obtained at wave-1 and wave-2 respectively in both converter and non-converter groups. The average interval between baseline and follow-up samples was 5.22 ± 0.25 years (total number of serum samples assayed=400). We estimated the approximate time-to-conversion to AD in the converter group as the midpoint of the time interval between their baseline and follow-up samples (2.62 ± 0.14 years) (Table-1B).

2.3 Sample size, data exclusion and randomization

Using the previously published estimate of within-class standard deviation of 1.5 by Mapstone et al [13] for metabolite concentrations assayed on the AbsoluteIDQ p180

platform, we calculated that a sample size of 100 controls and 100 AD cases would achieve >85% power to detect a log-fold change of 1.5 between control and converter groups.

In the BLSA samples, data from one subject diagnosed with mild cognitive impairment (MCI) was excluded. Data from an additional seven subjects who only had serum samples at one time point were also excluded. In the AGES-RS samples data from all subjects were included.

Converter and control samples were randomly divided into 6 groups and each group processed in separate batches during acquisition of serum metabolite concentrations. Serum metabolite concentrations were assayed in a blinded manner to diagnoses. For analysis of serum metabolite data, we used rigorous cross validation (including metabolite selection) to estimate classifier accuracy in an unbiased manner.

2.4 Data acquisition using BIOCRATES AbsoluteIDQ p180 metabolomics platform

BIOCRATES commercially available kit plates were used for the quantification of amino acids, acylcarnitines, sphingomyelins, phosphatidylcholines, hexoses, and biogenic amines. This validated assay uses two different mass spectrometric methods with isotope labelled and other internal standards for quantification. The acylcarnitines, lipids, and hexose are analyzed by flow injection analysis mass spectrometry (FIA-MS/MS). The amino acids and biogenic amines are subjected to phenylisothiocyanate (PITC)-derivatization and analyzed by HPLC-MS/MS using an AB SCIEX 4000 QTrap® mass spectrometer (AB SCIEX, Darmstadt, Germany) with electrospray ionization. A more detailed description of the assay can be found elsewhere [19]. The analytical platform was identical to that reported previously by Mapstone and colleagues [13].

2.5 Data reproducibility and variation within groups

The AbsoluteIDQ p180 platform yielded data that was highly reproducible with a median coefficient of variance of 5.3% (6.4% in AGES-RS) with 86% (90% in AGES-RS) of metabolites measured with CV<15% (calculated using quality control spike-in plasma samples). Equality of variance between groups was tested using the Levene test for homogeneity of variance. In the BLSA samples, 163 metabolites passed the test with a false detection ratio (FDR) > 0.05. In the AGES-RS data, all 187 metabolites passed the test with FDR>0.05. Normality of data within each group was not tested since the machine learning algorithms we used do not assume normality.

2.6 Evaluation of classifiers performance

Absolute serum concentrations of 187 targeted metabolites (including the 10 reported previously¹) were analyzed for discrimination between converters and non-converters at both time points (i.e. 'baseline' and 'follow-up'). Average values of classification accuracy (area-under-the-receiver-operating-characteristic-curve, AUC), and sensitivity/specificity (using disease probability thresholds of 0.5) were derived using several machine learning methods (elastic net regularized logistic regression (EN-RLR), random forest classifier (RF), support vector machines (SVM) and L₁ regularized logistic regression (L₁-RLR) (supplementary material and supplementary table-1) to assess how well these serum metabolite

concentrations discriminated between converter and non-converter samples. To avoid upward bias in the estimation of these metrics, we divided our data into training and testing sets. The 192 samples in BLSA and 200 samples in AGES-RS were thus split 80%/20% (BLSA; 152/40 and AGES-RS; 160/40 samples) for training and testing respectively. To account for variability due to random partitioning of training/testing data, the process was repeated 100 times and the performance metrics averaged over the 100 splits. The proportion of converters: non-converters was maintained approximately at 1:1 after partitioning.

3.0 RESULTS

The demographic characteristics of participants from the BLSA and AGES-RS-RS studies included in this report are described in Table-1A and B.

First, we used an identical logistic regression model applied by Mapstone et al. [13] to test the accuracy of their 10-metabolite panel as a predictor of preclinical AD in our cohorts. In the BLSA samples, this panel gave an area-under the-curve (AUC) of 0.64 and sensitivity/specificity of 51.6%/65.7% for discriminating baseline converter samples (i.e. pre-conversion) from non-converters. In the AGES-RS samples, the 10-metabolite panel gave an AUC of 0.394 and sensitivity/specificity of 47%/36% for discriminating baseline converter samples (i.e. pre-conversion) from non-converters. In comparison, Mapstone and colleagues reported an AUC of 0.92 and sensitivity/specificity of 90%/90% for discrimination between baseline converter and non-converter groups.

Next, we tested the 10-metabolite panel as a biomarker of current disease using the same logistic regression model to discriminate samples concurrent to symptom onset in converters (i.e. post-conversion) versus non-converters. In the BLSA samples, this analysis yielded an AUC of 0.58 and sensitivity/specificity of 53.8%/62.6%. In the AGES-RS samples, the 10-metabolite panel gave an AUC of 0.481 and sensitivity/specificity of 52%/48% for discriminating AD samples (i.e. post-conversion) from non-converters. In comparison, the index study reported an AUC of 0.77 for discrimination between post-conversion and non-converter samples (sensitivity/specificity were not reported for this analysis) [13].

Next, we analyzed the entire metabolite dataset of 187 metabolites without *a priori* hypotheses about the nature and/or identity of candidate AD biomarkers. In the BLSA samples, a Random Forest (RF) classifier discriminated pre-conversion samples from non-converters with a classification accuracy of 64.2% and sensitivity/specificity of 55.5%/73.0%. L₁-Regularized Logistic Regression (L₁-RLR) gave the best discrimination between samples concurrent to symptom onset (i.e. post-conversion) versus non-converters with classification accuracy of 67.2% and sensitivity/specificity of 67.7%/66.7%. Two additional algorithms i.e. Support Vector Machine (SVM) and Elastic Net-Regularized Logistic Regression (EN-RLR) yielded similar performance characteristics (supplementary table-1). Different metabolites, in particular phospholipids with fatty acid chains (summed to C30 to C44 carbon-carbon bonds) appeared to be important in discrimination between pre- and post-conversion samples relative to non-converters (supplementary table-2), suggesting that alterations in distinct metabolic pathways may underlie the pre-symptomatic and symptomatic phases of AD.

In the AGES-RS samples, our machine learning analyses demonstrated poor discrimination between both pre-conversion and non-converters as well as post-conversion samples relative to non-converters. All the classifiers tested showed performance metrics similar to random classification of samples between groups. The L₁-RLR classifier yielded classification accuracy of 46.5%, sensitivity/specificity of 45%/48% for pre-conversion samples and slightly better but substantially similar results for post-conversion samples. The other three classifiers gave similar results (not presented).

4.0 DISCUSSION

An accurate and non-invasive blood biomarker associated with preclinical AD is likely to revolutionize the care of patients with this devastating disease and accelerate the development of novel disease-modifying treatments by targeting them in individuals at greatest risk. The identification of such blood biomarkers is therefore likely to be of immediate benefit to patients and their caregivers in clinical as well as research settings. A recent report that a panel of ten plasma metabolites in blood could accurately predict conversion to AD in cognitively normal older individuals, raised hope that such biomarkers are within reach and received wide attention both within and outside the scientific community [13, 20]. However, other high profile biomarker findings in the AD field have not been replicated [21, 22], highlighting the need for large-scale confirmatory studies performed in well-characterized cohorts [23].

In this report, we first attempted to confirm the index findings by Mapstone et al. using the same metabolomics platform and analytical methodology used in their report. We undertook these analyses in two independent cohorts from longitudinal studies of normal aging i.e the BLSA and AGES-RS studies. We assayed metabolite concentrations from nearly 800 serial serum samples and found that we were unable to replicate the high performance metrics of the 10-metabolite panel in the sample of 28 converters reported in the index study. To the best of our knowledge, our current report is the largest blood biomarker study of preclinical AD to date.

It is important to first consider the major methodological differences between the current report and that of Mapstone et al. These are summarized in table-2. The most obvious consideration is the choice of matrix used to measure metabolite concentrations. Whereas the original study assayed plasma, we used serum samples in our study. However, recent studies have provided strong evidence for a high concordance between concentrations of metabolites assayed in plasma versus serum on the BIOCRATES Absolute^{IDQ}™ platform. In a large study sample from participants in the Cooperative Health Research in the Region of Augsburg (KORA) (n=377; 180 female, 197 male, age range from 51 to 84 years), Yu and colleagues reported on the concordance between plasma versus serum concentrations of 122 metabolites assayed on the BIOCRATES Absolute^{IDQ}™ platform [24]. Using their published data (available at: <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0021230#s5>), we examined the correlation between plasma and serum for each of the 10 metabolites reported by Mapstone et al. These findings are summarized in table-3 and show that most of these metabolites are highly correlated between the two matrices, with seven of the 10 analytes showing a correlation coefficient >0.85. Taken together, these

previous reports clearly establish that the majority of metabolite concentrations measured using the BIOCRATES Absolute*IDQ*TM method, and especially the ten metabolites reported by Mapstone et al. are highly correlated between plasma and serum. Therefore differences in the matrix used between our report and the index publication are unlikely to account for our inability to replicate the previously reported results.

The second major factor that could potentially account for differences between our results and those previously reported is the time-to-conversion to AD. Cognitively normal participants in the original report developed AD or mild cognitive impairment (MCI) over an average follow-up interval of 2.1 years, whereas those in the BLSA had an average time to symptom onset of 4.8 years. We therefore analyzed samples from an independent cohort of older individuals from the AGES-RS study with an estimated average time-to-conversion to AD (2.62 ± 0.14 years) that was more similar to that reported in Mapstone et al. to confirm that our failure to replicate the index findings were not driven by a slightly longer time to symptom onset in the BLSA cohort. It is also worth noting here that our inability to replicate the previous findings extend to both the prediction of subsequent AD in cognitively normal individuals (i.e. pre-conversion versus non-converters) as well as the discrimination between AD (i.e. post-conversion) and non-converter samples. If the inconsistencies in results were due entirely to differences in time-to-symptom onset, we would expect replication of the previous findings at least as markers of concurrent disease.

We also considered whether duration of sample storage may account for the inconsistency in results between our study and the index study. In their report, Mapstone and colleagues did not provide details on either the duration of sample storage or when participant recruitment to their study began. We are therefore unable to determine whether differences in duration of sample storage may contribute to the inconsistent results. In both the BLSA and AGES-RS studies, serum samples were collected after overnight fasting in standard serum separator tubes, centrifuged and stored in cryovials at -80°C until further use. The mean duration of serum sample storage ranged from 15.67 ± 8.12 years in the BLSA and 8.04 ± 2.77 years in the AGES-RS study. Serum aliquots used in our current report were not subject to additional freeze-thaw prior to the metabolomics assays reported herein. Breier et al. have examined the role of several pre-analytical variables such as sample storage duration, temperature and freeze-thaw cycles in targeted metabolomics analyses in blood using the BIOCRATES Absolute*IDQ*TM assay [25]. They conclude that the majority of metabolites measured are stable for up to 24 hours on cool packs and at room temperature even in non-centrifuged tubes. Additionally, serum metabolite concentrations were mostly unaffected by tube type and one or two freeze-thaw cycles. However, as some amino acids and biogenic amines may be unstable on cool packs, the authors recommend that blood samples be processed and frozen immediately after collection, as was done in both the BLSA and AGES-RS studies. We believe therefore, that serum metabolite concentrations measured in the BLSA and AGES-RS samples reported herein, followed optimal sampling conditions and storage times. Hence this pre-analytical variable i.e. sample storage duration is also unlikely to account for our failure to replicate the index findings of Mapstone et al.

Having assessed obvious methodological differences between our current study and that of Mapstone et al. which may account for the inconsistent results, we next considered the

analytical procedures adopted to examine the metabolomics data derived in the respective studies. Our first analysis was based on an *a priori* selection of the 10 metabolites reported in the original study. The index report provides few details about how these 10-metabolites were eventually selected from the approximately 180 that are assayed on the BIOCRATES Absolute*IDQ*TM platform [13]. Thus it is unclear how these data were processed, the exact procedures followed in constructing the final classifier and what, if any measures were adopted to avoid overfitting bias in estimating the performance metrics of the 10 metabolites. After our primary analysis with the *a priori* selected 10 metabolites that failed to replicate their original results, we examined the entire metabolomics data acquired in both BLSA and AGES-RS samples using machine learning approaches. By using rigorous partitioning of samples into training/test sets in two independent samples from the BLSA and AGES-RS studies, we guarded against over-fitting of data which is a common cause of inflated classification accuracies while using high-dimensional datasets such as metabolomics and proteomics on relatively small sample sizes [26]. In both the BLSA and AGES-RS samples, the analyses did not reveal a high level of discrimination between groups based on serum metabolite concentrations. In the BLSA, consensus metabolites were modestly associated with pre-symptomatic and symptomatic stages of AD and appeared to represent distinct metabolic pathways.

In conclusion, we were unable to replicate the high prediction performance for detecting preclinical AD reported previously for a metabolite panel assayed on the same platform using a substantially larger sample size drawn from two independent well-characterized longitudinal cohorts. Our study further underscores the importance of well-designed validation of exploratory biomarker studies. While the emerging technology of metabolomics holds great promise in epidemiological studies to accurately measure environmental exposure and quantify risk factors, our current report highlights the importance of performing large-scale replication of findings emerging from small index studies [27].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to the Baltimore Longitudinal Study of Aging and the Age, Gene/Environment Susceptibility Reykjavik Study participants for their dedication to these studies. This work was supported by the Intramural Research Program, National Institute on Aging, National Institutes of Health. AGES-RS-RS was supported by the National Institutes of Health contract N01-AG-12100, the National Institute on Aging Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament).

References

1. Albert MS, DeKosky ST, Dickson D, Dubois B, Feldman HH, Fox NC, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association*. 2011; 7:270–9.
2. McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR Jr, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-

- Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association*. 2011; 7:263–9.
3. Sperling RA, Aisen PS, Beckett LA, Bennett DA, Craft S, Fagan AM, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association*. 2011; 7:280–92.
 4. Trushina E, Mielke MM. Recent advances in the application of metabolomics to Alzheimer's Disease. *Biochim Biophys Acta*. 2014; 1842:1232–9. [PubMed: 23816564]
 5. Lewis GD, Gerszten RE. Toward metabolomic signatures of cardiovascular disease. *Circ Cardiovasc Genet*. 2010; 3:119–21. [PubMed: 20407099]
 6. Greenberg N, Grassano A, Thambisetty M, Lovestone S, Legido-Quigley C. A proposed metabolic strategy for monitoring disease progression in Alzheimer's disease. *Electrophoresis*. 2009; 30:1235–9. [PubMed: 19288586]
 7. Han X, Rozen S, Boyle SH, Hellegers C, Cheng H, Burke JR, et al. Metabolomics in early Alzheimer's disease: identification of altered plasma sphingolipidome using shotgun lipidomics. *PloS one*. 2011; 6:e21643. [PubMed: 21779331]
 8. Oresic M, Hyotylainen T, Herukka SK, Sysi-Aho M, Mattila I, Seppanen-Laakso T, et al. Metabolome in progression to Alzheimer's disease. *Transl Psychiatry*. 2011; 1:e57. [PubMed: 22832349]
 9. Sato Y, Nakamura T, Aoshima K, Oda Y. Quantitative and wide-ranging profiling of phospholipids in human plasma by two-dimensional liquid chromatography/mass spectrometry. *Anal Chem*. 2010; 82:9858–64. [PubMed: 21062019]
 10. Sato Y, Suzuki I, Nakamura T, Bernier F, Aoshima K, Oda Y. Identification of a new plasma biomarker of Alzheimer's disease using metabolomics technology. *J Lipid Res*. 2012; 53:567–76. [PubMed: 22203775]
 11. Trushina E, Dutta T, Persson XM, Mielke MM, Petersen RC. Identification of altered metabolic pathways in plasma and CSF in mild cognitive impairment and Alzheimer's disease using metabolomics. *PloS one*. 2013; 8:e63644. [PubMed: 23700429]
 12. Whiley L, Sen A, Heaton J, Proitsi P, Garcia-Gomez D, Leung R, et al. Evidence of altered phosphatidylcholine metabolism in Alzheimer's disease. *Neurobiol Aging*. 2014; 35:271–8. [PubMed: 24041970]
 13. Mapstone M, Cheema AK, Fiandaca MS, Zhong X, Mhyre TR, MacArthur LH, et al. Plasma phospholipids identify antecedent memory impairment in older adults. *Nature medicine*. 2014; 20:415–8.
 14. Ferrucci L. The Baltimore Longitudinal Study of Aging (BLSA): a 50-year-long journey and plans for the future. *J Gerontol A Biol Sci Med Sci*. 2008; 63:1416–9. [PubMed: 19126858]
 15. Harris TB, Launer LJ, Eiriksdottir G, Kjartansson O, Jonsson PV, Sigurdsson G, et al. Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *American journal of epidemiology*. 2007; 165:1076–87. [PubMed: 17351290]
 16. Kawas C, Gray S, Brookmeyer R, Fozard J, Zonderman A. Age-specific incidence rates of Alzheimer's disease: the Baltimore Longitudinal Study of Aging. *Neurology*. 2000; 54:2072–7. [PubMed: 10851365]
 17. Shock, NW. Gerontology Research C. Normal human aging : the Baltimore longitudinal study of aging. [Baltimore, Md.]; Washington, D.C.: U.S. Dept. of Health and Human Services, Public Health Service, National Institutes of Health, National Institute on Aging, Gerontology Research Center; 1984. For sale by the Supt. of Docs., U.S. G.P.O
 18. Qiu C, Cotch MF, Sigurdsson S, Jonsson PV, Jonsdottir MK, Sveinbjrnsdottir S, et al. Cerebral microbleeds, retinopathy, and dementia: the AGES-Reykjavik Study. *Neurology*. 2010; 75:2221–8. [PubMed: 21172845]
 19. Koal T, Klavins K, Seppi D, Kemmler G, Humpel C. Sphingomyelin SM(d18:1/18:0) is significantly enhanced in cerebrospinal fluid samples dichotomized by pathological amyloid-beta42, tau, and phospho-tau-181 levels. *Journal of Alzheimer's disease : JAD*. 2015; 44:1193–201. [PubMed: 25408209]

20. Diagnosing dementia This is not spinal tap. ECONOMIST-LONDON-ECONOMIST NEWSPAPER LIMITED-. 2014:79.
21. Bjorkqvist M, Ohlsson M, Minthon L, Hansson O. Evaluation of a previously suggested plasma biomarker panel to identify Alzheimer's disease. PloS one. 2012; 7:e29868. [PubMed: 22279551]
22. Ray S, Britschgi M, Herbert C, Takeda-Uchimura Y, Boxer A, Blennow K, et al. Classification and prediction of clinical Alzheimer's diagnosis based on plasma signaling proteins. Nature medicine. 2007; 13:1359–62.
23. Perlis RH. Translating biomarkers to clinical practice. Molecular psychiatry. 2011; 16:1076–87. [PubMed: 21709685]
24. Yu Z, Kastenmuller G, He Y, Belcredi P, Moller G, Prehn C, et al. Differences between human plasma and serum metabolite profiles. PloS one. 2011; 6:e21230. [PubMed: 21760889]
25. Breier M, Wahl S, Prehn C, Fugmann M, Ferrari U, Weise M, et al. Targeted metabolomics identifies reliable and stable metabolites in human serum and plasma samples. PloS one. 2014; 9:e89728. [PubMed: 24586991]
26. Thambisetty M, Lovestone S. Blood-based biomarkers of Alzheimer's disease: challenging but feasible. Biomark Med. 2010; 4:65–79. [PubMed: 20387303]
27. Tzoulaki I, Ebbels TM, Valdes A, Elliott P, Ioannidis JP. Design and analysis of metabolomics studies in epidemiologic research: a primer on -omic technologies. American journal of epidemiology. 2014; 180:129–39. [PubMed: 24966222]

Research in context

Systematic review

We first reviewed (using PubMed) all publications reporting the use of metabolomics in plasma/serum of AD patients and healthy controls. Few studies have used metabolomics to discover blood biomarkers of preclinical AD. A recent report identified a 10-metabolite panel in blood that could predict risk of conversion to AD in cognitively normal individuals with greater than 90% accuracy. Using the same targeted metabolomics method, we tried to confirm these results in a substantially larger study using more than 700 serial serum samples from two well-characterized, longitudinally followed cohorts of older individuals.

Interpretation

We were unable to replicate the recent findings implicating a panel of 10 metabolites as highly accurate blood biomarkers of preclinical AD. Our results underscore the need for large-scale validation of small exploratory biomarker studies.

Future directions

Future studies that expand the number and classes of analytes beyond those reported here will be required to discover novel blood metabolite biomarkers of AD. Such studies will require independent validation in well-characterized samples to establish their clinical utility.

Table 1A

Demographic characteristics of BLSA participants whose serum samples were analyzed in the current report. Data presented as mean (standard deviation).

	Whole Sample	Control	AD	DIFFERENCE (P-VALUE)
N	192	99	93	
Sex (F/M)	94/98	44/55	50/43	0.20
Age at first serum sample analyzed (years)	77.2 (6.6)	76.6 (6.7)	77.9 (6.5)	0.13
Age at second serum sample analyzed (years)	81.5 (6.2)	81.0 (6.2)	82.0 (6.3)	0.18
Interval between first and second samples analyzed (years)	4.3 (1.2)	4.4 (1.2)	4.1 (1.1)	0.11
Education (years)	16.4 (2.8)	16.2 (3.0)	16.5 (2.5)	0.75
Interval between pre-conversion sample and onset of cognitive impairment (years)			4.8 (1.2)	
Interval between post-conversion sample and onset of cognitive impairment (years)			0.69 (0.83)	

The differences between Control group and AD groups were tested using Mann–Whitney U test for continuous variables and Chi-Square test for categorical variables.

Table 1B

Demographic characteristics of AGES-RS participants whose serum samples were analyzed in the current report. Data presented as mean (standard deviation).

	Whole Sample	Control	AD	DIFFERENCE (P-VALUE)
N	200	100	100	
Sex (F/M)	109/91	55/45	54/46	
Age at first serum sample analyzed (years)	78.20 (4.38)	78.23 (4.39)	78.18 (4.40)	−0.05 (0.934)
Age at second serum sample analyzed (years)	83.43 (4.42)	83.43 (4.41)	83.43 (4.45)	0.00 (0.978)
Interval between first and second samples analyzed (years)	5.22 (0.25)	5.20 (0.20)	5.24 (0.29)	0.04 (0.379)
Education *	2.03 (0.91)	2.17 (0.95)	1.90 (0.85)	−0.27 (0.048)
Interval between pre-conversion sample and onset of cognitive impairment (years)			2.62 (0.14)	
Interval between post-conversion sample and onset of cognitive impairment (years)			2.62 (0.14)	

The differences between Control group and AD groups were tested using Mann–Whitney U test for continuous variables and Chi-Square test for categorical variables.

* Education was graded semi-quantitatively according to the highest level of completed education as follows: 1-primary school or less; 2-secondary school; 3-college; 4-University degree.

Table 2

Comparison of participant characteristics, study design, experimental methodology and results between the index study by Mapstone et al [13] and the current report.

Blood samples analyzed	Mapstone et al.	Current report	
		<u>BLSA</u>	<u>AGES-RS</u>
Healthy controls	73	99	100
Converters with both pre- and post-conversion samples	28 [§]	93	100
Age at sampling			
Healthy controls; baseline	81.49 (3.48)	76.6 (6.7)	78.23 (4.39)
Healthy controls; follow-up	Not reported	81.0 (6.2)	83.43 (4.41)
Pre-conversion	80.21 (4.02)	77.9 (6.5)	78.18 (4.40)
Post-conversion	82.23 (3.95)	82.0 (6.3)	83.43 (4.45)
Follow-up interval			
Pre-conversion sample to symptom onset (years)	2.1	4.8 (1.2)	2.62 (0.14)
Sex (F/M)			
Healthy controls	62/39	94/98	109/91
Converters	46/27	44/55	55/45
Converters	16/12	50/43	54/46
Education (years)			
Healthy controls	15.52 (2.36)	16.2 (3.0)	2.17 (0.95)*
Converters	15.04 (2.74)	16.5 (2.5)	1.9 (0.85)*
Sample assayed	Plasma	Serum	Serum
Metabolomics assay used	AbsoluteIDQ p180 BIOCRATES	AbsoluteIDQ p180 BIOCRATES	
Performance metrics of 10-metabolite panel described in Mapstone et al.			
Pre-conversion versus controls AUC/ Sensitivity/Specificity	0.92/90%/90%	<u>BLSA</u> 0.642/51.6%/65.7%	<u>AGES-RS</u> 0.395/47.0%/36.0%
Post-conversion versus controls AUC/ Sensitivity/Specificity	0.77 sensitivity/specificity not reported	0.575/53.8%/62.6%	0.482/52.0%/48.0%
Unbiased assessment of all targeted metabolites assayed			
Methodology	Not reported	All 187 metabolite levels assessed as potential AD biomarkers using multiple machine learning classifiers without <i>a priori</i> assumptions**	
Pre-conversion versus controls Accuracy/Sensitivity/Specificity	Not reported	<u>BLSA</u> 0.642/55.5%/73.0%	<u>AGES-RS</u> 0.415/45.0%/48.0%
Post-conversion versus controls Accuracy/Sensitivity/Specificity	Not reported	0.672/67.7%/66.7%	0.515/53.0%/50.0%

[§] Details available at <http://www.nature.com/nm/journal/v20/n4/extref/nm.3466-S1.pdf>

** Education was graded semi-quantitatively according to the highest level of completed education as follows: 1-primary school or less; 2-secondary school; 3-college; 4-University degree.

** Only results from the L1-RLR classifier are reported for AGES-RS. Other classifiers tested yielded similar results.

Table 3

Correlation coefficients ('R between plasma and serum') for concentrations of the 10-metabolites reported by Mapstone et al [13]. Data extracted from Yu, et al [24] (n=377; 180 female, 197 male, age range from 51 to 84 years), and available online at <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0021230#s5>

Metabolite names	Mean \pm SD (μ M) in plasma	Mean \pm SD (μ M) in serum	Relative mean difference (%)	R between plasma and serum
PC aa C36:6	1.15 \pm 0.49	1.27 \pm 0.56	9.30	0.93
PC aa C38:6	82.47 \pm 26.46	92.59 \pm 28.9	11.43	0.90
PC aa C40:6	27.39 \pm 9.51	30.27 \pm 10.13	10.00	0.90
lysoPC a C18:2	26.13 \pm 8.45	30.01 \pm 9.82	13.68	0.89
PC aa C38:0	3.09 \pm 0.84	3.52 \pm 0.98	12.44	0.88
C3	0.04 \pm 0.01	0.04 \pm 0.02	7.00	0.86
PC ae C40:6	4.89 \pm 1.28	5.51 \pm 1.52	11.26	0.86
PC aa C40:1	0.41 \pm 0.09	0.45 \pm 0.1	7.70	0.79
PC aa C40:2	0.33 \pm 0.09	0.37 \pm 0.1	-26.2	0.33
C16:1-OH	0.01 \pm 0.01	0.01 \pm 0.01	0	0.01