Journal of Alzheimer's Disease xx (2022) x–xx
DOI 10.3233/JAD-215640
IOS Press

Augmenting Imaging Biomarker Performance with Blood-Based Gene Expression Levels for Predicting Alzheimer's Disease Progression

Vitaly I. Dobromyslin, Dalila B. Megherbi*, for the Alzheimer's Disease Neuroimaging Initiative[1]
Center for Computer Machine/Human Intelligence Networking and Distributed Systems, University of Massachusetts, Lowell, MA, USA

Handling Associate Editor: Ruth Barrientos

Accepted 25 February 2022
Pre-press 15 March 2022

Abstract.
Background: Structural brain imaging metrics and gene expression biomarkers have previously been used for Alzheimer's disease (AD) diagnosis and prognosis, but none of these studies explored integration of imaging and gene expression biomarkers for predicting mild cognitive impairment (MCI)-to-AD conversion 1-2 years into the future.
Objective: We investigated advantages of combining gene expression and structural brain imaging features for predicting MCI-to-AD conversion. Selection of the differentially expressed genes (DEGs) for classifying cognitively normal (CN) controls and AD patients was benchmarked against previously reported results.
Methods: The current work proposes integrating brain imaging and blood gene expression data from two public datasets (ADNI and ANM) to predict MCI-to-AD conversion. A novel pipeline for combining gene expression data from multiple platforms is proposed and evaluated in the two independents patient cohorts.
Results: Combining DEGs and imaging biomarkers for predicting MCI-to-AD conversion yielded 0.832-0.876 receiver operating characteristic (ROC) area under the curve (AUC), which exceeded the 0.808-0.840 AUC from using the imaging features alone. With using only three DEGs, the CN versus AD predictive model achieved 0.718, 0.858, and 0.873 cross-validation AUC for the ADNI, ANM1, and ANM2 datasets.
Conclusion: For the first time we show that combining gene expression and imaging biomarkers yields better predictive performance than using imaging metrics alone. A novel pipeline for combining gene expression data from multiple platforms is proposed and evaluated to produce consistent results in the two independents patient cohorts. Using an improved feature selection, we show that predictive models with fewer gene expression probes can achieve competitive performance.

Keywords: Alzheimer's disease, gene expression, MCI-to-AD conversion, predictive model

*Correspondence to: Professor Dalila B. Megherbi, Center for Computer Machine/Human Intelligence Networking and Distributed Systems, University of Massachusetts, Lowell, MA, USA. E-mail: Dalila_Megherbi@uml.edu.

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

INTRODUCTION

According to a 2018 press release by the US Centers for Disease Control and Prevention (CDC), the population growth in the US will be accompanied by more than doubling of Alzheimer's disease (AD) prevalence by 2060 [1]. Longer lifespan in developed countries may also lead to a greater portion of the population experiencing age-related cognitive deficits such as stable mild cognitive impairment (sMCI), and progressive MCI (pMCI) leading to AD. Research by Ganguli et al. suggests that individuals with some form of MCI have up to 20 times greater risk of converting to AD than their cognitively normal (CN) counterparts [2]. The much higher AD risk for the MCI group makes AD screening for this group clinically more relevant than screening healthy aging adults.

Extensive research over the past 40 years into AD pathology has elucidated several key aspects of neuronal changes associated with AD, yet the exact mechanism of disease progression continues to be unknown. The AD pathology in the brain can be broadly divided into three types: 1) damage related to accumulation of toxic material; 2) lesions characterized by tissue loss; and 3) reactive processes such as inflammation and plasticity [3]. Currently, AD is diagnosed with neuropsychological assessment, followed by brain imaging to confirm characteristic AD pathology in the brain. Alzheimer's Disease Neuroimaging Initiative (ADNI) and AddNeuroMed (ANM) datasets provide access to multimodal data related to AD diagnosis and progression, collected from multi-center patient population in the US and Europe, respectively. While structural MRI data may likely be the most commonly accessed part of these repositories, the ADNI and ANM datasets also contain clinical and demographic information, Apolipoprotein E (*APOE*) status, and gene expression levels collected from blood samples. Lebedev et al. used 24 cortical thickness and volume markers, along with *APOE*, age, and education, for predicting CN versus AD (84.2-90.7% Sensitivity, 82.9-88.3% Specificity) and MCI-to-AD conversion (78.0-79.0% Sensitivity, 82.9% Specificity) in the ADNI and ANM cohorts [4]. Besides requiring a significant amount of testing, incurring substantial costs, and potentially requiring an invasive injection of a radioactive substance, this approach is limited by the availability of a magnetic resonance imaging (MRI) or positron emission tomography (PET) scanner, and scales poorly to resource-limited populations.

On the other hand, blood biomarkers, such as gene expression, provide a readily available and cost-effective source of diagnostic information. The major issue with using blood gene expression for AD diagnosis is that multiple organs and tissues contribute to the observed transcriptional profile. Some other challenges with using transcriptional profile biomarkers for predicting disease progression arise from a varying sample dilution affecting differentially expressed gene (DEG) identification [5], lab-to-lab variability [6], and inter-platform differences between various manufacturers [7]. In the case of neurological conditions such as AD, it remains to be seen whether DEGs in the blood can provide useful diagnostic information by themselves, and in addition to the known structural MRI biomarkers. To simultaneously evaluate both sensitivity and specificity of the DEGs and imaging biomarkers we will use the receiver operating characteristic (ROC) area under the curve (AUC) metric, which represents the AUC of the model sensitivity plotted versus the false positive rate, defined as (1-Specificity). This metric ranges from zero to one, with the latter representing ideal model performance.

Lee et al. recently built a predictive model for distinguishing individuals with AD from healthy controls using ADNI and ANM (ANM1 and ANM2) datasets [8]. They identified 334 DEGs that were used with L1 normalized logistic regression (L1-LR) to train on ADNI dataset and test on ANM1 dataset, achieving ROC AUC of 0.70. The pathway analysis revealed that AD-related genes were enriched with inflammation, mitochondria, and Wnt signaling pathways. Voyle et al. [9] compared the performance of three Random Forest predictive models for distinguishing individuals with AD from non-demented controls using ANM gene expression data: 1) demographic data model included sample collection site, age, years of full-time education; 2) demographic and gene expression model; and 3) pathway analysis model. The demographic model achieved the best AUC ROC of 0.771, followed by pathway and gene expression models achieving 0.729 and 0.724 AUC ROC. The authors concluded that the pathway model did not have any performance advantage over the gene expression model.

While most studies have focused on classifying AD versus CN condition, few have examined MCI either as a separate category, or considered predicting conversion from MCI to AD. Using the ADNI dataset, Miller et al. found the CLIC1 gene to be the only DEG between CN, MCI, and AD conditions [10]. While their algorithm achieved high cross-validation AUC

of 0.906 for predicting AD, the performance was not verified on an internal or external test data set. Lunnon et al. reported 0.78 sensitivity for correctly predicting conversion from MCI to AD within 2 years in the ANM1 dataset, albeit with poor specificity of 0.25 [11]. Their random forest classifier used 48 genes, corresponding to 50 probes. Using a relatively small MCI dataset ($N = 66$, 34 pMCI) collected from 8 centers in Norway and Sweden, Roed et al. reported 0.73 sensitivity and 0.81 specificity for predicting MCI-to-AD conversion within 2 years [12]. The performance metrics above were calculated using leave-one-out cross-validation, which in some cases is known to provide a more optimistic estimate of performance than a more conservative evaluation with an independent test set.

Challenges for wider adoption of gene expression markers in clinical practice arise in some instances from the inability to reproduce classifier performance in independent population cohorts and very little consistency between the sets of gene expression markers identified in various studies [13]. One of the strengths of the current study stems from using two independent datasets, the ADNI and ANM population cohorts originating from US and Europe, respectively. ADNI and ANM datasets have been acquired on different platforms (Affymetrix and Illumina respectively), making the findings of the current study platform-agnostic and more clinically applicable than insights from a single instrument platform. In general, confirming DEG consistency across different platforms presents a challenge due to several factors: a) difference in technology for measuring gene expression (bead versus microarray), b) probes targeting the same gene at different chromosome locations, c) alternate splicing. As an example, Li et al. could not confirm any of the ANM-identified DEGs in the ADNI dataset [14]. A novel processing pipeline is proposed here to address platform differences and avoid shortcomings associated with averaging probes targeting the same gene [15]. By showing the inconsistency of most DEGs identified with a single dataset analysis, we further show the benefit of using multiple datasets to identify and confirm DEGs associated with AD.

MATERIALS AND METHODS

Data used in the preparation of this article were obtained from the ADNI database (http://adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD.

Gene expression data processing

Whole blood samples were collected and analyzed on two different platforms, with ADNI and ANM data having been acquired on Affymetrix Human Genome U219 and Illumina HumanHT-12 (v3 and v4) Expression BeadChip platforms, respectively. In order to facilitate comparison between these platforms, only genes targeted on both platforms ($N = 14{,}498$) have been included in the subsequent analysis (Fig. 1). Focusing on the shared genes reduced the number of probes from 49,386 to 38,947 for ADNI, from 38,324 to 29,485 for ANM1, from 32,051 to 20,177 for ANM2 datasets. Age and the number of *APOE* $\varepsilon 4$ alleles are known risk factors for AD onset and were accounted for with a General Linear Model (GLM) [16] in ADNI and ANM datasets independently. Additional annotation information for the DEGs was obtained through the National Center for Biotechnology Information database [17].

Classification of AD versus CN using DEGs

The ADNI training set for this classification task included 162 (38 AD) cases, and the test set had 74 (21 AD) cases, according to the 67/33% random split (Table 1). The ANM1 training and testing sets consisted of 105 (56 AD) cases and 45 (24 AD) cases, respectively. The ANM2 training and testing sets included 82 (44 AD) and 35 (19 AD) cases, respectively. The three datasets were simultaneously evaluated to find the DEGs for classifying CN versus AD, regardless of the platform. Cross-validation and test ROC AUC metrics were compared for fitting the logistic regression model within each individual dataset. The stopping criteria for adding new DEGs to the final feature set was defined as having less than 0.015 improvement in cross-validation ROC AUC in any one of the three datasets.

Gene expression data for MCI-to-AD conversion

Due to the limitations within the available data, conversion to AD was examined within 2 years in ADNI and 1 year in ANM1. The ADNI MCI pop-

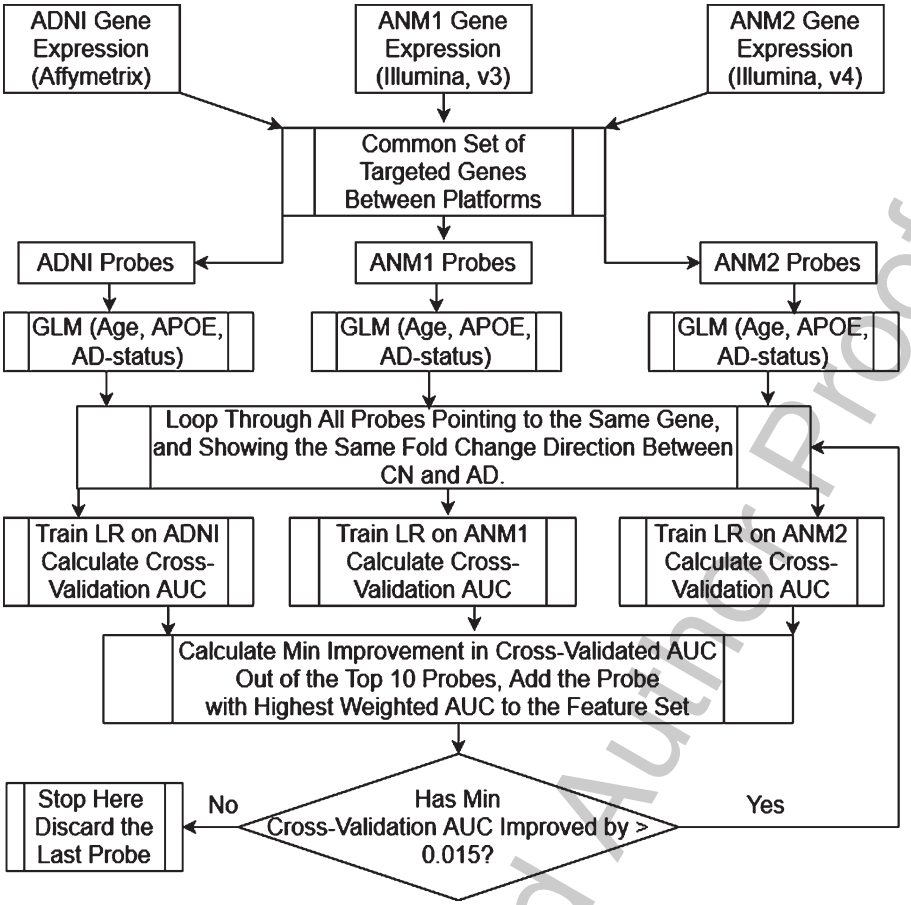

Fig. 1. Workflow for processing and eventually combining CN versus AD DEGs identified in the three datasets available, acquired on Affymetrix and Illumina platforms.

Table 1
Participant characteristics from the two independent datasets, ADNI and ANM (part 1 and 2), included in the current study. Age and MMSE are shown as mean ± SD, whereas the APOE ε4 carriers field shows percentage of participants with at ≥ 1 APOE ε4 risk allele

	ADNI				ANM1				ANM2	
	CN	AD	sMCI	pMCI	CN	AD	sMCI	pMCI	CN	AD
Number of Cases	177	59	193	39	70	80	55	23	54	63
Age	76.2 ± 6.7	76.8 ± 7.2	71.1 ± 7.6	73.7 ± 7.5	73.7 ± 7.5	76.0 ± 6.6	74.5 ± 5.3	74.9 ± 5.9	75.7 ± 6.1	78.6 ± 5.6
Gender, % Males	48.0%	61.0%	52.3%	56.4%	37.1%	30.0%	47.3%	34.8%	31.5%	34.9%
APOE ε4 Carriers	26.0%	74.6%	37.3%	74.4%	35.7%	56.3%	27.3%	56.5%	18.5%	54.0%
MMSE	29.2 ± 1.1	21.9 ± 3.3	28.2 ± 1.5	27.0 ± 1.8	29.1 ± 1.1	20.9 ± 4.7	27.1 ± 2.0	26.4 ± 1.7	28.4 ± 1.7	19.6 ± 4.6

ulation at baseline ($N = 279$) was separated into the training and testing sets according to a 67/33% random split, yielding 188 (32 pMCI) and 91 (13 pMCI) cases in each set, respectively. Similar train/test partitioning resulted in 60 (21 pMCI) and 25 (8 pMCI) assigned to the train/test sets in ANM1, and 39 (7 pMCI) and 16 (2 pMCI) assigned to the train/test sets in ANM2. Similar to the CN versus AD classification, the impact of age and the number of *APOE* ε4 alleles on transcriptional profile was accounted for with a linear regression [16] in ADNI and ANM datasets independently (Fig. 2).

Imaging data for MCI-to-AD conversion

For both ADNI and ANM datasets, T1-weighted images were acquired on a 1.5T MRI scanner. Structural imaging metrics included cortical volume and regional thickness estimates from FreeSurfer software [18] for ADNI (FreeSurfer v5.1) and ANM1

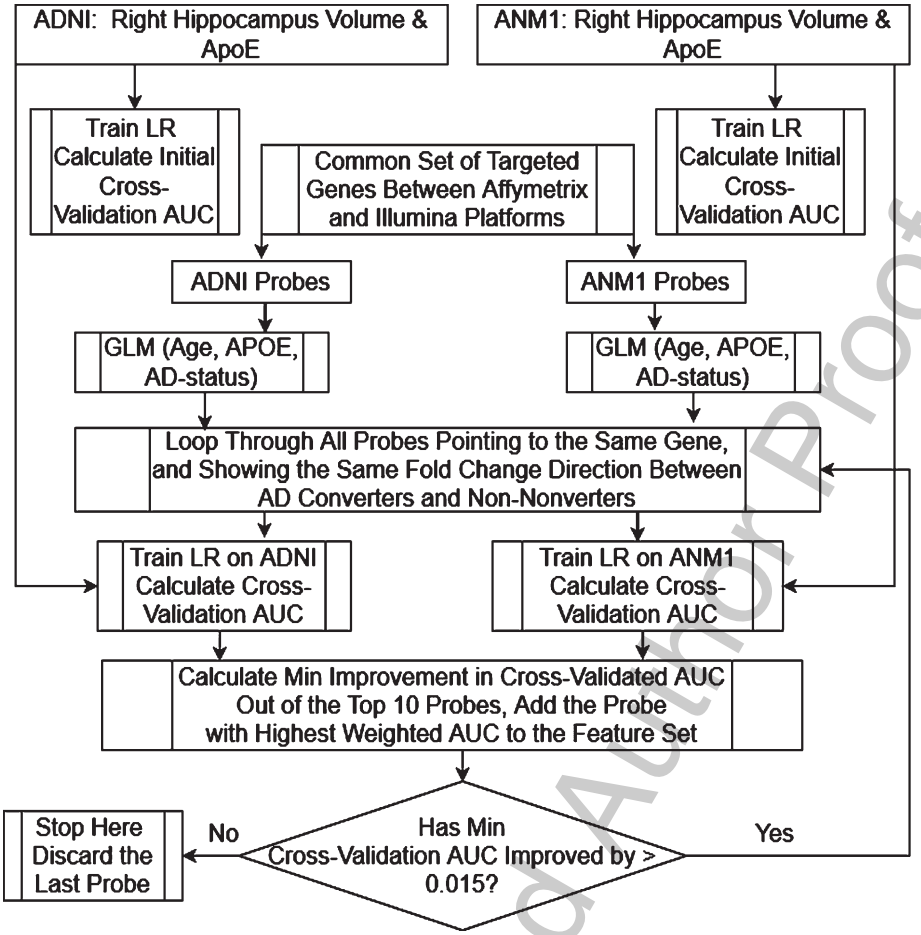

Fig. 2. Workflow for combining gene expression and imaging biomarkers for predicting MCI-to-AD conversion.

(FreeSurfer v5.3). Only brain scans with passing quality control status were included for subsequent analysis in both ADNI and ANM datasets. The logistic regression model integration of imaging and gene expression markers required eliminating subjects with incomplete imaging or gene expression data. In the ANM2 dataset, only 22 subjects had both gene expression and imaging data, and this dataset was excluded from further imaging analysis. The final ADNI training/test sets included 156 (28 pMCI) and 76 (11 pMCI) subjects, respectively, whereas the final ANM1 dataset consisted of 56 (18 pMCI) and 22 (5 pMCI) training/test subjects, respectively (Table 1). To account for head size variation, FreeSurfer computed intracranial volume was regressed out from other volume estimates. Following the methodology proposed by Lebedev et el., 24 structural MRI metrics for predicting MCI-to-AD conversion were used as a starting point [4]. Due to a significantly higher number of subjects in the ADNI dataset compared to the ANM1, the selection of imaging and demographic markers for MCI-to-AD conversion was performed exclusively in the ADNI dataset. Age, gender, education, and *APOE* information were initially combined with the 24 structural MRI metrics, yielding 28 features, in the LR model to predict MCI-to-AD conversion. To improve the predictive model's sparsity and interpretability, without the loss of performance, we used a sequential backward feature selection, initialized with the full 28-feature dataset, and eliminated redundant features one-at-a-time.

Enhancing imaging marker performance with gene expression for predicting MCI-to-AD conversion

Baseline cross-validation AUC was calculated separately for ADNI (denoted as $AUC_{0,ADNI}$) and ANM1 (denoted as $AUC_{0,ANM1}$) using solely imaging-demographic features. Augmenting

imaging and demographic markers with gene expression data involved first finding all the probe combinations between ADNI and ANM1 pointing to the same gene (Fig. 2). Next, the fold change between AD converters and non-converters was calculated for each probe combination and compared between ADNI and ANM1 datasets. The probe combinations with fold change trending in the same direction were simultaneously added to the initial imaging-demographic feature sets In ADNI and ANM1, one-at-a-time. A pair of candidate LR models was subsequently trained independently in the ADNI and ANM1 datasets, yielding a set of new cross-validation AUCs, $AUC_{New,ADNI}$ and $AUC_{New,ANM1}$, for ADNI and ANM1 respectively. We estimated the relative AUC improvement in each dataset from baseline as a vector [$AUC_{New,ADNI} - AUC_{0,ADNI}$, $AUC_{New,ANM1} - AUC_{0,ANM1}$]. To capture the overall improvement in both datasets with a single metric, we calculated minimum improvement in cross-validation AUC as min([$AUC_{New,ADNI} - AUC_{0,ADNI}$, $AUC_{New,ANM1} - AUC_{0,ANM1}$]). Probes were sorted in the descending order of minimum cross-validation AUC improvement, with the most promising probes at the top of the list. To further select the "winning" probe, individual cross-validation improvements from ADNI and ANM1 were weighted by the respective dataset size. The resulting metric had 73.6% of the weight determined by ADNI ($N = 156$ subjects), and the remaining 26.4% of the weight determined by ANM1 ($N = 56$ subjects). Final probe selection consisted of picking the probe with highest weighted cross-validation AUC improvement from the top 10 candidates in the original list. Following the selection of the first probe, other probes were evaluated in the same fashion, with the only difference being that they were evaluated against the previous best feature set, which included the last best probe. The stopping criteria for adding new DEGs to the final feature set was defined as having less than 0.015 improvement in cross-validation AUC in any one of the two datasets.

Predictive model interpretability

Logistic regression represents a special case of generalized additive models, where the outcome is modelled as a linear combination of each factor's contribution to the log-odds ratio. Deviance in logistic regression describes the discrepancy between the estimated maximum likelihood value and the observed value, and has been previously used to estimate feature importance for predicting a clinical endpoint [19]. To assess individual feature contributions in the final model predictive performance, change in deviance was calculated by iteratively eliminating one of the features in the model, re-fitting the reduced model, and comparing the resulting deviance of the reduced model to that of the full model. A large increase in deviance suggests that the eliminated variable plays an important role in the model. Since change in deviance for eliminating a single factor from the model follows a chi-squared distribution with one degree of freedom, p-values can be iteratively estimated for all included factors.

RESULTS

Classification of AD versus CN using DEGs

For the task of classifying AD from CN based on transcriptional profile, the current pipeline selected three DEGs with consistent fold change across all three datasets. The final set of DEGs consisted of NGDN, DFFB, and NDUFS5. Fitting a LR model independently to each of the three training datasets yielded the following performance as measured by cross-validation set ROC AUC: 0.718, 0.858, and 0.873 for the ADNI, ANM1, and ANM2 datasets respectively (Fig. 3). Models were further evaluated on the three test sets (one for each data subset), yielding 0.659, 0.903, and 0.737 ROC AUC for the ADNI, ANM1, and ANM2 datasets, respectively.

We used the change in deviance to estimate individual feature contribution in the LR models described above, see Table 2. In the ADNI dataset two of the three selected gene probes, NGDN and DFFB, could account for most of the change in deviance and were also statistically significant at $p < 0.05$ level. On the other hand, in the ANM1 and ANM2, the third gene probe, NDUFS5, accounted for the largest change in deviance. This probe was statistically significant at $p < 0.05$ level in the ANM2 dataset, but not in the ANM1 dataset.

Predicting MCI-to-AD conversion by combining imaging and gene expression

Initial ADNI features used for developing an MCI-to-AD predictive model included 24 MRI-derived structural metrics, age, gender, education, and *APOE* information. LR model developed with these features yielded 0.782 and 0.759 AUC for the cross-validation

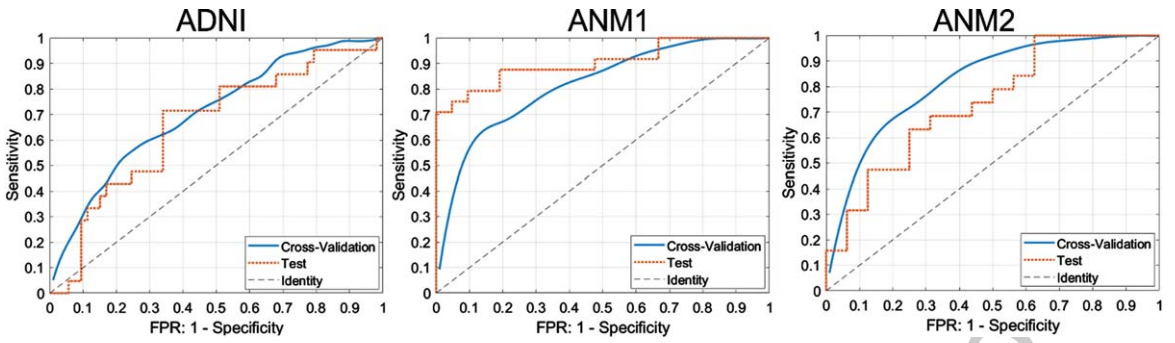

Fig. 3. ROC plots with LR model, trained on the 3 DEGs for classifying CN versus AD, after adjusting for age and *APOE* ε4 alleles. The following AUC metrics were calculated for five-fold cross-validation, and for internal test datasets (reported in parentheses): 0.718 (0.659), 0.858 (0.903), and 0.873 (0.737) for ADNI, ANM1, and ANM2 datasets respectively.

Table 2
Variable contributions to the CN versus AD model performance, characterized by the change in deviance and corresponding statistical significance from the Chi-Squared distribution. Values in bold highlight statistically significant results at p < 0.05

Variable-to-Eliminate	ADNI Change in Deviance	ADNI p	ANM1 Change in Deviance	ANM1 p	ANM2 Change in Deviance	ANM2 p
NGDN	**4.57**	**3.3E-2**	0.41	5.2E-1	0.71	4.0E-1
DFFB	**4.48**	**3.4E-2**	1.17	2.8E-1	3.68	5.5E-2
NDUFS5	2.44	1.2E-1	3.40	6.5E-2	**4.35**	**3.7E-2**

and test sets, respectively. Following step-wise backward feature selection, the final reduced model achieved 0.840 cross-validation and 0.794 test AUC, with the feature set consisting of the right hippocampus volume and the number of *APOE* risk alleles (Fig. 4a). When the same two features were used to derive LR model in the ANM1 dataset, the model had 0.808 cross-validation and 0.718 test AUC (Fig. 4b).

The search for gene probes to enhance imaging-*APOE* model performance yielded a single probe, associated with the DUSP13 gene, that 1) showed consistent fold change direction in all three datasets, 2) improved cross-validation AUC from 0.840 to 0.876, and from 0.808 to 0.832 in the ADNI and ANM1 datasets respectively, and 3) improved test set AUC from 0.794 to 0.820 in ADNI and from 0.718 to 0.741 in the ANM1 datasets. We performed a Monte Carlo simulation with 10,000 iterations to examine the statistical significance of these AUC increases. In each iteration, 25% of the ADNI and ANM1 training sets was randomly sampled without replacement and allocated for validation, whereas the remaining 75% of the training set was randomly sampled with replacement and used to train LR models. The results from 10,000 iterations yielded 95% confidence intervals of [0.0241, 0.0261] and [0.003,0.0078] for the increase in AUC in the ADNI and ANM1 respectively. Performing a t-test for the increase in AUC confirmed statistical significance at $p < 0.05$ for both datasets.

Figure 5 shows the transcriptional profile for the DUSP13 targeting probe in each of the three datasets. Interestingly, the expression level measured by this probe in ADNI is far lower than that in ANM1 and ANM2 datasets, yet the 0.22 ADNI fold change exceeds the 0.02-fold change in both ANM1 and ANM2 datasets. To further examine DUSP13 contribution to the final predictive model, we examined the change in deviance (Table 3). The DUSP13 contribution to deviance met statistical significance at $p < 0.05$ in the ADNI, but not in the ANM1. For the ANM1 dataset, the DUSP13's change in deviance of 2.67 exceeded that of *APOE* status (0.22), a known risk factor for AD. To a lesser extent, in ADNI, DUSP13's change in deviance (4.33) came close to that of *APOE* (4.88). Right hippocampus volume proved to have the greatest contribution to the predictive model.

Table 4 provides additional information related to the DEGs identified in the current study to be associated with AD disease progression. DEGs related to nervous system development (NGDN gene), apoptosis (DFFB gene), and mitochondrial function (NDUFS5 gene) were identified for CN versus AD classification. DUSP13 gene, identified for improving MCI-to-AD conversion, regulates cell proliferation and differentiation.

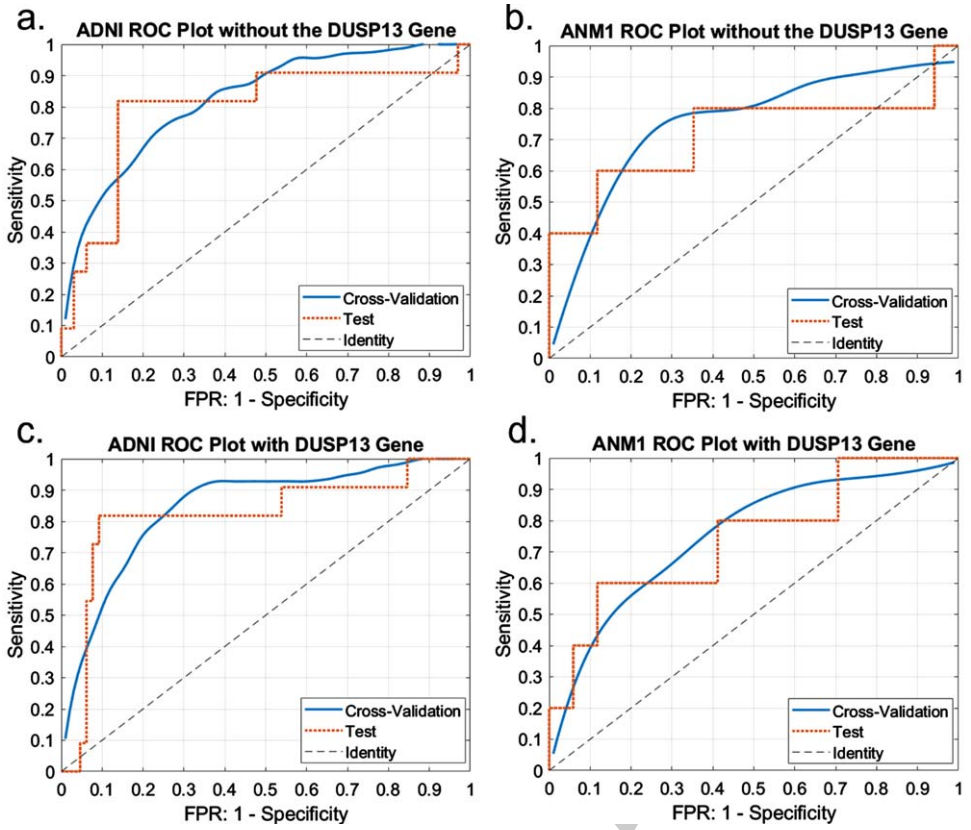

Fig. 4. ROC plots with LR model, trained on the right hippocampal volume and number of *APOE* alleles (a and b), augmented with DUSP13 gene expression level for predicting conversion from MCI to AD (c and d). The following AUC metrics were calculated for five-fold cross-validation, and for internal test datasets (reported in parentheses): 0.840 (0.794), 0.808 (0.718), 0.876 (0.820), and 0.832 (0.741) for a-d, respectively.

DISCUSSION

Classification of AD versus CN using DEGs

Despite the common perception that the gene expression data in the blood is noisy and may not be of sufficient diagnostic quality for AD prediction, the classifier performance presented here suggests that there is some diagnostic utility in considering blood gene expression data for this purpose. With only three DEGs, the LR model in the current study reached 0.718, 0.858, and 0.873 cross-validation AUC comparable to the previously reported 0.657, 0.874, and 0.804 for the ADNI, ANM1, and ANM2 datasets, respectively [8]. Interestingly the three DEGs identified in the current work did not overlap with the 334 DEGs reported by Lee et al. One of the genes (NDUFS5) that was used for AD versus CN classification in the current work has been previously reported, along with other 314 DEGs to be statistically significant ($p = 1.2E\text{-}8$) for CN versus AD classification, and along with 204 other DEGs for CN versus MCI conversion ($p = 9.8E\text{-}8$) [20]. Voyle et al. have also identified the NDUFS5 gene as a DEG in the classification of CN versus AD, with variable importance of 5.9 on the 4.7-11.9 scale [9]. With using 13 DEGs their predictive model achieved 0.724 test set AUC on the merged ANM dataset, which was lower than the 0.903 and 0.737 test set AUC for ANM1 and ANM2 respectively reported in the current work. The DFFB gene's association with AD reported in the current work is further supported by the earlier research showing that DFFB gene expression decreased in response to administration of neurotoxic fragment of amyloid-β protein $A\beta_{25–35}$ in the mouse model of AD [21]. Recently Madrid et al. reported NGDN as one of the 15 genes associated with AD irrespective of the *APOE* haplotype [22]. This way, all three genes found to be associated with CN versus AD classification in the current study have been implicated by prior research to be involved in AD pathology. To the best of our knowledge, this is the first time the three

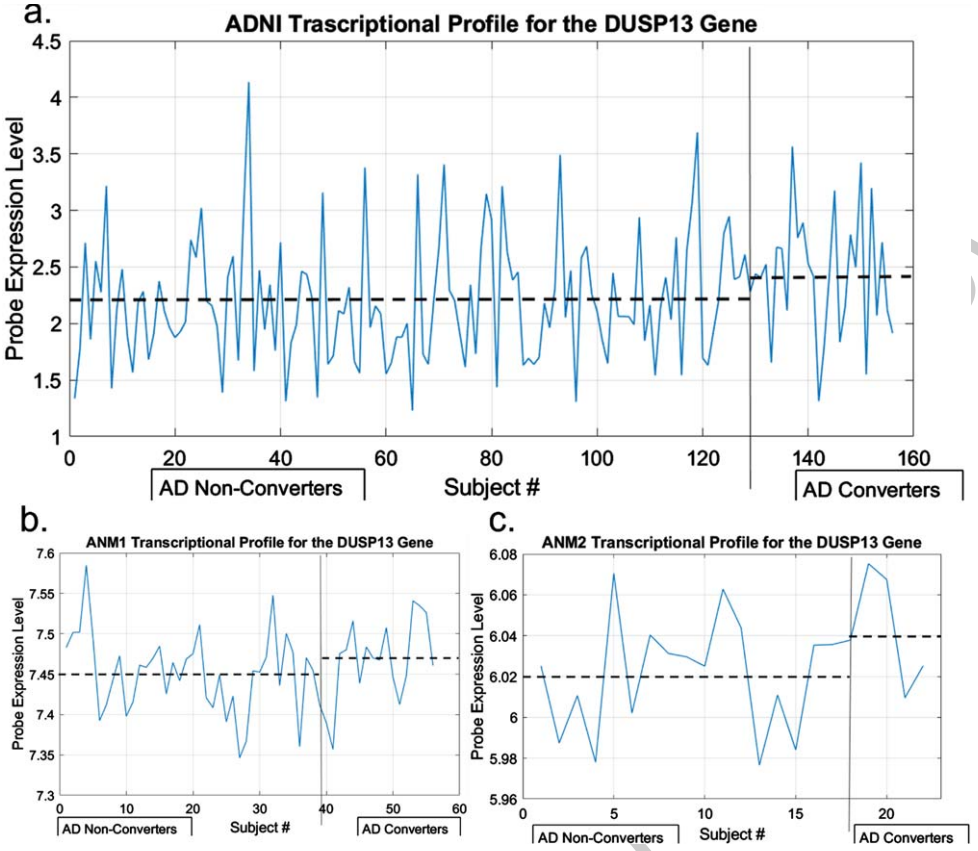

Fig. 5. Transcriptional profile for DUSP13 gene in ADNI (a), ANM1 (b), and ANM2 (c).

Table 3
Variable contributions to the MCI-to-AD model performance, characterized by the change in deviance and corresponding statistical significance from the Chi-Squared distribution. Values in bold highlight statistically significant results at $p < 0.05$

Variable-to-Eliminate	ADNI Change in Deviance	ADNI p	ANM1 Change in Deviance	ANM1 p
Right Hippocampus Volume	**16.29**	**5.5E-5**	**8.44**	**3.7E-3**
# *APOE* Risk Alleles	**4.88**	**2.7E-2**	0.22	6.4E-1
DUSP13 Gene Expression	**4.33**	**3.8E-2**	2.67	1.0E-1

DEGs above have been used together with Age and *APOE*, as part of the mathematical model to classify CN versus AD. Besides delivering competitive performance to the previously reported models with higher number of DEGs, using fewer gene expression probes offers several advantages, including better interpretability and lower analysis cost. As we show through the analysis of deviance in the logistic regression model, some of the DEGs are only identifiable by carrying out simultaneous DEG selection in all three datasets. The ability to assess each feature's contribution to the modelled outcome gives generalized additive models an interpretability advantage over other modeling techniques.

Predicting MCI-to-AD conversion by combining imaging and genetic variables

The current study explored whether a neuroimaging-*APOE* predictive model for MCI-to-AD conversion can be enhanced by including transcriptional profile features from the blood. Through the exhaustive search for the DEGs that would consistently improve MCI-to-AD conversion

Table 4
Summary information on the DEGs identified in the AD versus CN and MCI to AD conversion classification tasks (NCBI)

Classification task	Gene/Probe (Affymetrix Probe; Illumina v3 Probe; Illumina v4 Probe)	Summary
AD versus CN (Gene Expression Only)	NGDN (11743867_at; ILMN_2324998; ILMN_2324998)	Neuroguidin is an EIF4E (MIM 133440)-binding protein that interacts with CPEB (MIM 607342) and functions as a translational regulatory protein during the development of the vertebrate nervous system.
	DFFB (11745694_x_at; ILMN_1678962; ILMN_1678962)	DNA fragmentation factor (DFF), triggers both DNA fragmentation and chromatin condensation during apoptosis.
	NDUFS5 (11757665_x_at; ILMN_1776104; ILMN_1776104)	Encodes one of the iron-sulfur protein (IP) components of mitochondrial NADH:ubiquinone oxidoreductase (complex I).
MCI to AD Conversion (with Imaging)	DUSP13 (11728562_a_at; ILMN_1776240; ILMN_2373666)	Member of the protein-tyrosine phosphatase superfamily, cooperates with protein kinases to regulate cell proliferation and differentiation.

prediction in both ADNI and ANM1 datasets, only the DUSP13 gene met the criteria for minimum improvement in cross-validation AUC of >0.015. AD-converters in ADNI, ANM1, and ANM2 showed upregulated expression of the DUSP13 gene when compared to non-converters. While this gene is not commonly associated with AD pathology, research has linked the DUSP13 upregulation to adaptive response to oxidative stress [23]. Chronic inflammation, one of the characteristics of AD pathology, has been suggested to start with proinflammatory cytokines released by activated macrophages in the blood [24]. In this context, the DUSP13 upregulation in the AD patient group supports this suggestion and makes the DUSP13 gene expression a potentially valuable companion to imaging biomarkers for AD prognosis. Due to the short 1–2-year delay in MCI-to-AD conversion studied here, and the difference in the follow up period in ANM (1 year) and ADNI (2 years), the DUSP13 gene's role in predicting longer time window conversion remains unclear.

One of the limitations of the current study is that we treated each predictive biomarker in the predictive model as an independent feature, whereas a more complex relationship likely exists between AD-related biomarkers. Future research may benefit from using graphical theory tools to examine biomarker-to-biomarker connections in the context of AD pathology.

Another limitation of the current study stems from not considering other, likely more sensitive blood-based biomarkers such as p-tau species including p-tau181, p-tau217, p-tau231, and glial fibrillary acidic protein (GFAP) [25]. A promising future direction would be to explore the potential of combining DEGs and the p-tau species/GFAP biomarkers in the context of improving prognosis quality for MCI patients.

ACKNOWLEDGMENTS

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis

Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (http://www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

The results published here are in whole or in part based on data obtained from the AD Knowledge Portal (https://adknowledgeportal.org).

The AddNeuroMed data are from a public – private partnership supported by EFPIA companies and SMEs as part of InnoMed (Innovative Medicines in Europe), an Integrated Project funded by the European Union of the Sixth Framework program priority FP6 2004-LIFESCIHEALTH-5. Clinical leads responsible for data collection are Iwona Kłoszewska (Lodz), Simon Lovestone (London), Patrizia Mecocci (Perugia), Hilkka Soininen (Kuopio), Magda Tsolaki (Thessaloniki), and Bruno Vellas (Toulouse) and imaging leads are Andy Simmons (London), Lars-Olad Wahlund (Stockholm) and Christian Spenger (Zurich) and bioinformatics leads are Richard Dobson (London) and Stephen Newhouse (London).

The ANMerge update of the AddNeroMed data has received partial support from the Innovative Medicines Initiative Joint Undertaking "AETIONOMY" under grant agreement #115568, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in kind contribution. The update of the original AddNeuroMed data was conducted by: Colin Birkenbihl, Sarah Westwood, Liu Shi, Alejo Nevado-Holgado, EricWestman, Simon Lovestone and Martin Hofmann-Apitius.

Authors' disclosures available online (https://www.j-alz.com/manuscript-disclosures/21-5640r1).

REFERENCES

[1] CDC (2018) *U.S. burden of Alzheimer's disease, related dementias to double by 2060*. US Centers for Disease Control and Prevention.

[2] Ganguli M, Snitz BE, Saxton JA, Chang CC, Lee CW, Vander Bilt J, Hughes TF, Loewenstein DA, Unverzagt FW, Petersen RC (2011) Outcomes of mild cognitive impairment by definition: A population study. *Arch Neurol* **68**, 761-767.

[3] Duyckaerts C, Delatour B, Potier MC (2009) Classification and basic pathology of Alzheimer disease. *Acta Neuropathol* **118**, 5-36.

[4] Lebedev AV, Westman E, Van Westen GJ, Kramberger MG, Lundervold A, Aarsland D, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Lovestone S, Simmons A, Alzheimer's Disease Neuroimaging Initiative and the AddNeuroMed consortium (2014) Random Forest ensembles for detection and prediction of Alzheimer's disease with a good between-cohort robustness. *Neuroimage Clin* **6**, 115-125.

[5] Barnes M, Freudenberg J, Thompson S, Aronow B, Pavlidis P (2005) Experimental comparison and cross-validation of the Affymetrix and Illumina gene expression analysis platforms. *Nucleic Acids Res* **33**, 5914-5923.

[6] Irizarry RA, Warren D, Spencer F, Kim IF, Biswal S, Frank BC, Gabrielson E, Garcia JG, Geoghegan J, Germino G, Griffin C, Hilmer SC, Hoffman E, Jedlicka AE, Kawasaki E, Martinez-Murillo F, Morsberger L, Lee H, Petersen D, Quackenbush J, Scott A, Wilson M, Yang Y, Ye SQ, Yu W (2005) Multiple-laboratory comparison of microarray platforms. *Nat Methods* **2**, 345-350.

[7] Wong WC, Loh M, Eisenhaber F (2008) On the necessity of different statistical treatment for Illumina BeadChip and Affymetrix GeneChip data and its significance for biological interpretation. *Biol Direct* **3**, 23.

[8] Lee T, Lee H (2020) Prediction of Alzheimer's disease using blood gene expression data. *Sci Rep* **10**, 3485.

[9] Voyle N, Keohane A, Newhouse S, Lunnon K, Johnston C, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Lovestone S, Hodges A, Kiddle S, Dobson RJ (2016) A pathway based classification method for analyzing gene expression for Alzheimer's disease diagnosis. *J Alzheimers Dis* **49**, 659-669.

[10] Miller JB, Kauwe JSK (2020) Predicting clinical dementia rating using blood RNA levels. *Genes (Basel)* **11**, 706.

[11] Lunnon K, Sattlecker M, Furney SJ, Coppola G, Simmons A, Proitsi P, Lupton MK, Lourdusamy A, Johnston C, Soininen H, Kloszewska I, Mecocci P, Tsolaki M, Vellas B, Geschwind D, Lovestone S, Dobson R, Hodges A, dNeuroMed Consortium (2013) A blood gene expression marker of early Alzheimer's disease. *J Alzheimers Dis* **33**, 737-753.

[12] Roed L, Grave G, Lindahl T, Rian E, Horndalsveen PO, Lannfelt L, Nilsson C, Swenson F, Lonneborg A, Sharma P, Sjogren M (2013) Prediction of mild cognitive impairment that evolves into Alzheimer's disease dementia within two years using a gene expression signature in blood: A pilot study. *J Alzheimers Dis* **35**, 611-621.

[13] Kiddle SJ, Voyle N, Dobson RJB (2018) A blood test for Alzheimer's disease: Progress, challenges, and recommendations. *J Alzheimers Dis* **64**, S289-S297.

[14] Li X, Wang H, Long J, Pan G, He T, Anichtchik O, Belshaw R, Albani D, Edison P, Green EK, Scott J (2018) Systematic analysis and biomarker study for Alzheimer's disease. *Sci Rep* **8**, 17394.

[15] Consortium M, Shi L, Reid LH, Jones WD, Shippy R, Warrington JA, Baker SC, Collins PJ, de Longueville F, Kawasaki ES, Lee KY, Luo Y, Sun YA, Willey JC, Setterquist RA, Fischer GM, Tong W, Dragan YP, Dix DJ, Frueh FW, Goodsaid FM, Herman D, Jensen RV, Johnson CD, Lobenhofer EK, Puri RK, Schrf U, Thierry-Mieg J,

Wang C, Wilson M, Wolber PK, Zhang L, Amur S, Bao W, Barbacioru CC, Lucas AB, Bertholet V, Boysen C, Bromley B, Brown D, Brunner A, Canales R, Cao XM, Cebula TA, Chen JJ, Cheng J, Chu TM, Chudin E, Corson J, Corton JC, Croner LJ, Davies C, Davison TS, Delenstarr G, Deng X, Dorris D, Eklund AC, Fan XH, Fang H, Fulmer-Smentek S, Fuscoe JC, Gallagher K, Ge W, Guo L, Guo X, Hager J, Haje PK, Han J, Han T, Harbottle HC, Harris SC, Hatchwell E, Hauser CA, Hester S, Hong H, Hurban P, Jackson SA, Ji H, Knight CR, Kuo WP, LeClerc JE, Levy S, Li QZ, Liu C, Liu Y, Lombardi MJ, Ma Y, Magnuson SR, Maqsodi B, McDaniel T, Mei N, Myklebost O, Ning B, Novoradovskaya N, Orr MS, Osborn TW, Papallo A, Patterson TA, Perkins RG, Peters EH, Peterson R, Philips KL, Pine PS, Pusztai L, Qian F, Ren H, Rosen M, Rosenzweig BA, Samaha RR, Schena M, Schroth GP, Shchegrova S, Smith DD, Staedtler F, Su Z, Sun H, Szallasi Z, Tezak Z, Thierry-Mieg D, Thompson KL, Tikhonova I, Turpaz Y, Vallanat B, Van C, Walker SJ, Wang SJ, Wang Y, Wolfinger R, Wong A, Wu J, Xiao C, Xie Q, Xu J, Yang W, Zhang L, Zhong S, Zong Y, Slikker W, Jr. (2006) The MicroArray Quality Control (MAQC) project shows inter- and intraplatform reproducibility of gene expression measurements. *Nat Biotechnol* **24**, 1151-1161.

[16] Mathworks in *Statistics and Machine Learning Toolbox*, ed. Matlab Mathworks.

[17] National Center for Biotechnology Information. National Center for Biotechnology Information (NCBI). Bethesda (MD), https://www.ncbi.nlm.nih.gov/

[18] Fischl B (2012) FreeSurfer. *Neuroimage* **62**, 774-781.

[19] Dobromyslin VI, Megherbi DB (2020) Analysis of Interacting Factors Contributing to the Onset of Diabetes Type 2 in the Female Pima Population: A Step towards Understanding the Links. *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)* IEEE, pp. 1-10.

[20] Tao Y, Han Y, Yu L, Wang Q, Leng SX, Zhang H (2020) The predicted key molecules, functions, and pathways that bridge mild cognitive impairment (MCI) and Alzheimer's disease (AD). *Front Neurol* **11**, 233.

[21] Kolobov VV, Davydova TV, Zakharova IA, Gorbatov V, Fomina VG (2012) [Repressional effects of the glutamate antibodies on expression of Dffb gene in the brain of rats with experimental Alzheimer's disease]. *Mol Biol (Mosk)* **46**, 757-765.

[22] Madrid L, Moreno-Grau S, Ahmad S, González-Pérez A, de Rojas I, Xia R, Martino Adami PV, García-González P, Kleineidam L, Yang Q, Damotte V, Bis JC, Noguera-Perea F, Bellenguez C, Jian X, Marín-Muñoz J, Grenier-Boley B, Orellana A, Ikram MA, Amouyel P, Satizabal CL; Alzheimer's Disease Neuroimaging Initiative (ADNI)*; EADI consortium, CHARGE consortium, GERAD consortium, GR@ACE/DEGESCO consortium, Real LM, Antúnez-Almagro C, DeStefano A, Cabrera-Socorro A, Sims R, Van Duijn CM, Boerwinkle E, Ramírez A, Fornage M, Lambert JC, Williams J, Seshadri S; ADAPTED consortium, Ried JS, Ruiz A, Saez ME (2021) Multiomics integrative analysis identifies APOE allele-specific blood biomarkers associated to Alzheimer's disease etiopathogenesis. *Aging (Albany NY)* **13**, 9277-9329.

[23] Omata Y, Saito Y, Fujita K, Ogawa Y, Nishio K, Yoshida Y, Niki E (2008) Induction of adaptive response and enhancement of PC12 cell tolerance by lipopolysaccharide primarily through the upregulation of glutathione S-transferase A3 via Nrf2 activation. *Free Radic Biol Med* **45**, 1437-1445.

[24] Tweedie D, Sambamurti K, Greig NH (2007) TNF-alpha inhibition as a treatment strategy for neurodegenerative disorders: New drug candidates and targets. *Curr Alzheimer Res* **4**, 378-385.

[25] Palmqvist S, Tideman P, Cullen N, Zetterberg H, Blennow K, Alzheimer's Disease Neuroimaging Initiative, Dage JL, Stomrud E, Janelidze S, Mattsson-Carlgren N, Hansson O (2021) Prediction of future Alzheimer's disease dementia using plasma phospho-tau combined with other accessible measures. *Nat Med* **27**, 1034-1042.