

Machine Learning-Based Method for Personalized and Cost-Effective Detection of Alzheimer's Disease

Javier Escudero*, *Member, IEEE*, Emmanuel Ifeakor, *Member, IEEE*, John P. Zajicek, Colin Green, James Shearer, and Stephen Pearson, for the Alzheimer's Disease Neuroimaging Initiative

Abstract—Diagnosis of Alzheimer's disease (AD) is often difficult, especially early in the disease process at the stage of mild cognitive impairment (MCI). Yet, it is at this stage that treatment is most likely to be effective, so there would be great advantages in improving the diagnosis process. We describe and test a machine learning approach for personalized and cost-effective diagnosis of AD. It uses locally weighted learning to tailor a classifier model to each patient and computes the sequence of biomarkers most informative or cost-effective to diagnose patients. Using ADNI data, we

classified AD versus controls and MCI patients who progressed to AD within a year, against those who did not. The approach performed similarly to considering all data at once, while significantly reducing the number (and cost) of the biomarkers needed to achieve a confident diagnosis for each patient. Thus, it may contribute to a personalized and effective detection of AD, and may prove useful in clinical settings.

Index Terms—Alzheimer's disease (AD), classification, cost, machine learning, mild cognitive impairment (MCI), personalization.

Manuscript received March 14, 2012; revised July 6, 2012; accepted August 2, 2012. Date of publication August 8, 2012; date of current version December 14, 2012. This work was funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research Programme (Grant Reference Number RP-PG-0707-10124). Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Amofix Life Sciences Ltd.; AstraZeneca; Bayer HealthCare; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd. and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. 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 (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129 and K01 AG030514. *Asterisk indicates corresponding author.*

*J. Escudero is with the Signal Processing and Multimedia Communications Research Group, School of Computing and Mathematics, Plymouth University, Plymouth, PL4 8AA, U.K. (e-mail: javier.escudero@ieee.org).

E. Ifeakor is with the Signal Processing and Multimedia Communications Research Group, School of Computing and Mathematics, Plymouth University, Plymouth, PL 4 8AA, U.K. (e-mail: e.ifeakor@plymouth.ac.uk).

J. P. Zajicek is with the Clinical Neurology Research Group, Peninsula College of Medicine and Dentistry, Plymouth University, Plymouth, PL6 8BU, U.K. (e-mail: john.zajicek@nhs.net).

C. Green and J. Shearer are with the Health Economics Group, Peninsula College of Medicine and Dentistry, University of Exeter, Exeter, EX4 4QJ, U.K. (e-mail: colin.green@pms.ac.uk; james.shearer@pms.ac.uk).

S. Pearson is with Devon Partnership NHS Trust, Ivybridge, PL21 9QP, U.K. (e-mail: stephen.pearson@nhs.net).

Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.ucla.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.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

Digital Object Identifier 10.1109/TBME.2012.2212278

I. INTRODUCTION

RECENT advances in technology have enabled the recording of vast amounts of data. Machine learning methods have been proposed to aid in the interpretation of such data for clinical decision making and diagnosis [1]–[3]. However, most current applications of machine learning fail to mimic the personalized diagnostic process of real clinical settings [4].

In practice, the clinician decides which tests are most appropriate for each patient. If the results are conclusive, a diagnosis is established. Otherwise, the clinician orders other tests for clarification. All these decisions are tailored to the patient [4]. Instead, most machine learning approaches apply the same classification model to all patients with no tailoring of the diagnostic decisions and they assume that all biomarkers are readily available at once [5]–[7]. This is seldom the case and implies that patients would need to undergo a considerable number of clinical procedures, which may be costly and/or invasive, even though some tests may not be relevant for them. Thus, it is desirable to develop new approaches to support clinicians in the early, more effective (in terms of number of tests and/or cost), and personalized detection of disease.

Alzheimer's disease (AD) is the most common neurodegenerative disease in older people [8]. There is a considerable delay between the start of AD pathology and the clinical diagnosis of AD dementia, which can only be confirmed by autopsy [8], [9]. Thus, it is very difficult to detect AD early and accurately [9], and there is a need for intelligent means to support clinicians in the personalized diagnosis of this disease [3].

To address such challenges, we test a proof-of-concept personalized classifier for AD dementia and mild cognitive impairment (MCI) patients based on biomarkers [10]–[12]. We extend previous analyses [4] to AD, including new feature selection approaches, classifier, and measures of similarity between subjects suitable for continuous variables. Our aim is to support the clinician in the diagnosis process by providing him or her with information about the patient's probability of disease and which biomarkers may be more informative. This approach is

consistent with the emerging recommendations on diagnosis across the spectrum of AD [13].

II. MATERIALS

A. ADNI Database

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.ucla.edu>). The ADNI was launched in 2003 by the NIA, the NIBIB, the FDA, private pharmaceutical companies and non-profit organizations. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of MCI and early AD. For up-to-date information, see <http://www.adni-info.org>.

B. Retrieval of Variables

The ADNI database as of Dec. 14, 2011 was queried for basic data (gender, age, years of education, and body mass index) and biomarkers—MRI entorhinal cortical thickness [10]; average level of PET glucose uptake [11]; number of ApoE ϵ 4 alleles (measured from blood) [12]; levels of $A\beta_{42}$, total Tau, phosphorylated pTau_{181p}, 8-iso-PGF_{2a} and 8,12-iso-iPF_{2a}-VI isoprostanes, and homocysteine in the cerebrospinal fluid [12]; and level of $A\beta_{40}$ and $A\beta_{42}$ in plasma [12]—of cognitive normal (CN), MCI, and AD subjects. We included all ADNI subjects whose variables were simultaneously available and had passed the quality controls. Two sets were considered. The first one contains 41 AD patients and 45 CN subjects, while the second one is composed of 71 MCI subjects who have diagnosis information after one year. The MCI subjects are split into 12 converters (cMCI, MCI patients who were diagnosed with AD after 12 months) and 59 nonconverters (nMCI, MCI subjects who remained diagnosed as MCI at the 12-month follow up).

C. Estimated Cost of the Biomarkers

The personalized classifier is able to account for the number or cost of the biomarkers used for diagnosis [4]. The cost of each biomarker in a clinical setting has been retrieved from the “2010–2011 reference costs” for the National Health Service in U.K. [14] or estimated after consultation with specialist units in the U.K. The basic data (gender, age, years of education, and body mass index) are considered readily available and they have no cost. The MRI and PET features cost £163 and £844, in that order. £169 is the estimation for the ApoE ϵ 4 data, whereas the measurement of isoprostanes and homocysteine is priced at £26 each. The cost of each other biomarker is set at £132 each.

III. METHODS

A. General Overview

Fig. 1 illustrates the approach. When a new subject arrives, the immediate basic data are collected. These are the variables available at this stage. The approach first tries to classify the new subject before deciding which biomarker to order. To *person-*

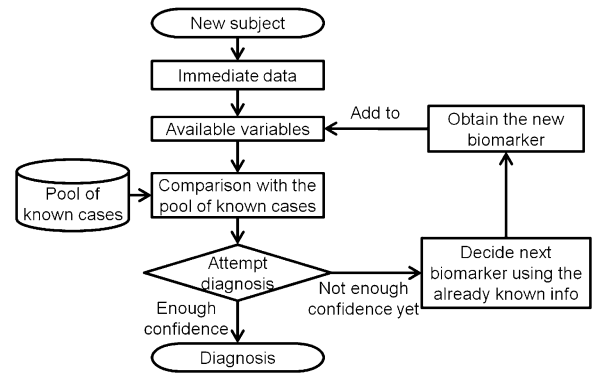


Fig. 1. Block diagram of the personalized and cost-effective classification.

alize the classifier, the available variables are compared against those of a *Pool* of already diagnosed people. This comparison establishes which cases in the *Pool* are most similar to the new subject. *Weights* are computed to reflect this: more similar subjects in the *Pool* are assigned larger weights, thus affecting the training of the classifier more. This approach is based on *locally weighted learning* [1], [4], [15].

Then, a diagnosis is attempted. If it can be established with enough confidence, then the process ends. Otherwise, the system determines which additional biomarker may contribute most to the diagnosis of the new subject by maximizing the diagnostic information or diagnostic information per unit of cost it provides. This process uses only the new subject's available variables and the *Pool* of known cases.

Once the system selects the next biomarker, it is acquired for the new subject and added to the set of available variables. New weights reflecting the similarity of this new set of available variables with those of the subjects in the *Pool* are calculated and the classification is reattempted. This iterative process ends when the confidence in the diagnosis exceeds a predefined threshold or no more biomarkers are available.

The next example illustrates our approach. An 81-year-old male with a body mass index of 24 and 18 years of education arrives at the clinic. The clinician uses the system to assist in the diagnosis. According to it, the probability of the AD based on the patient's basic data is 0.48. The clinician uses the system to check which biomarker may be most informative for this case. Comparing this subject with the *Pool*, the system recommends PET and the clinician orders it, which shows low glucose consumption. The system tries to classify the case and reports a probability of AD of 0.68. The clinician decides to order another biomarker. After comparing his data with the cases in the *pool*, the system recommends MRI. This is acquired and it shows clear evidence of cortical atrophy. With this information, the system reports a probability of AD of 0.99. Considering the biomarker evidence, the clinician decides to establish a diagnosis of AD.

The system implicitly refuses to label a subject when the confidence is not high enough as the classification output is the probability of disease, from which the crisp labels of “Positive/Negative” cases are derived only once predefined thresholds are exceeded.

B. Personalized Classification

Instance-based classifiers (e.g., k nearest neighbors), also known as memory-based classifiers, combine the information from the training set with that of the new subject to create an *ad hoc* classification model [1], [2], [4]. Other methods (e.g., neural networks) process all data in advance to create one-decision model that is applied to all cases [1], [2], [4]–[7]. The former may be sensitive to noisy data, whereas the latter are unable to do any personalization. An alternative is *locally weighted learning* [1], [4], [15], which weights each training instance by its distance to the new subject to reflect how relevant each case in the training set is with regard to the subject to be classified. Then, the weighted instances are used to train a classifier [1], [15]. Locally weighted learning is similar to memory-based classifiers because all computation is deferred to the moment the subject to be classified arrives so that an *ad hoc* decision rule is created. Thus, it allows us to tailor the classifier to each patient by allocating the most importance in training to the cases most similar to the patient.

In this letter, the weights for the locally weighted learning vary linearly with the normalized Euclidean distance between the new patient and the subjects in the Pool [1], ranging from 0 (totally dissimilar) to 1 (identical cases). Logistic regression is used as a proof-of-concept base classifier for the locally weighted learning [1]. Since the output of logistic regression ranges from 0 to 1, it can be seen as the probability Pr that the subject being classified is a positive case [1]. A confidence threshold α is set [4]. If Pr is high enough ($Pr > 1 - \alpha$, most surely the patient's diagnosis is positive) or low enough ($Pr < \alpha$, most probably the subject is a negative case) the diagnostic process ends. Otherwise ($\alpha < Pr < 1 - \alpha$), there is not enough confidence in the diagnosis, and more biomarkers are needed.

Using only the training set, all variables are normalized to the [0,1] range and synthetic minority oversampling technique [1] is used to equalize the frequency of the classes.

C. Selection of Additional Biomarkers

Feature selection is used to decide which variables are most informative for a task [1], [2], [7]. The feature selection considered in most articles is difficult to be directly applied in clinical settings because it assumes that all biomarkers are readily available and it is not personalized [4]. However, a personalized approach based on the distance between the new subject and known cases in the Pool may solve such problems.

The personalized approach selects the biomarkers one by one. It starts by taking the already available data and weights that account for the similarity between the subject and known cases in the Pool. Ten runs of a tenfold cross validation in the Pool of known cases are used to estimate the classification performance of this “old” setting. Then, in turn, each other potential biomarker available for the subjects in the Pool is added to the set of features to create a “new” set of variables. Ten runs of a tenfold cross validation are run again within the Pool so that the difference between the “old” and the “new” performance is computed for every biomarker. The one that maximizes this difference is selected. The “old” and “new”

classifications are computed using the weights of the previous diagnosis attempt because cases similar to the new subject will tend to remain similar when more biomarkers are added.

The performance in this selection procedure is measured with two criteria: 1) accuracy (ratio of correctly classified cases) and 2) area under the ROC curve (AUC, probability that the classifier ranks a randomly chosen positive instance above a randomly chosen negative one) [1]. Since this feature selection uses only subjects in the training set (Pool) who have all their data available and the already measured variables of the new case, it can be implemented in clinical settings.

This feature selection results in a sequence of biomarkers likely to lead to a confident subject classification with a small number of biomarkers. Alternatively, the biomarker costs can be considered. The difference in performance between “old” and “new” sets of variables is divided by the cost of the biomarker being considered, leading to a measure of increase in performance per unit of cost. Thus, the approach can operate in two modes: 1) maximizing the performance (which tries to minimize the *number* of biomarkers for diagnosis) or 2) maximizing the performance per unit of cost (which tries to minimize the *cost* of the diagnosis).

D. Evaluation

We evaluate our approach using leave-one-out cross-validation. One participant is considered the new subject and is left out for testing, while the remaining participants are considered the Pool of known cases (training set). This process is repeated as many times as subjects, leaving a different one out each time [1], [2]. The method is assessed against four criteria: final accuracy, final AUC, number of biomarkers to achieve a confident classification, and cost of such biomarkers. Two classification tasks are considered: 1) CN versus AD and 2) nMCI versus cMCI, while the system operates in two modes: minimization of number or cost of biomarkers. Three confidence thresholds α are studied (0.10, 0.15, and 0.20). These values are chosen because the diagnostic accuracy of AD in common clinical practice is about 80% [9], and it is expected that an ideal biomarker for AD should have sensitivity and specificity of no less than 80% [9].

For a fair comparison with bulk-data classifiers, a classical logistic regression based on all variables is applied to both tasks. We also compare our results with those of a system where logistic regression is combined with locally weighted learning, but no feature selection is used at all. This is an intermediate development where the decision rule is personalized but there is no feature selection whatsoever.

IV. RESULTS

Table I contains the results, computed with leave-one-out cross-validation, for the first task (CN–AD) in terms of final accuracy and AUC values, and mean \pm standard deviation of the number and cost (in £) of the biomarkers used in diagnosis. We tested several values of α and criteria to select the next biomarker (AUC or accuracy). For a fair comparison, we also classified the subjects using a logistic regression and a locally

TABLE I

RESULTS FOR THE CLASSIFICATION OF CN VERSUS AD SUBJECTS GIVEN AS FINAL AUC, FINAL ACCURACY, NUMBER, AND COST OF THE BIOMARKERS

CN vs. AD	Final AUC	Final Accuracy	Number of biomarkers	Cost of biomarkers
LR	0.900	0.849	11±0	£1914±0
LWL with LR	0.884	0.860	11±0	£1914±0
Personalized classifiers minimizing the number of biomarkers				
AUC with $\alpha=0.20$	0.887	0.826	1.94±1.89	£967±250
AUC with $\alpha=0.15$	0.877	0.837	2.24±2.30	£1003±288
AUC with $\alpha=0.10$	0.873	0.849	2.91±3.02	£1076±354
Accuracy with $\alpha=0.20$	0.882	0.826	2.01±2.24	£905±343
Accuracy with $\alpha=0.15$	0.879	0.837	2.34±2.60	£952±367
Accuracy with $\alpha=0.10$	0.872	0.849	3.03±3.22	£1049±409
Personalized classifiers minimizing the cost of biomarkers				
AUC with $\alpha=0.20$	0.731	0.744	2.38±1.70	£334±333
AUC with $\alpha=0.15$	0.774	0.767	2.79±1.89	£429±402
AUC with $\alpha=0.10$	0.816	0.802	3.40±2.18	£558±477
Accuracy with $\alpha=0.20$	0.758	0.756	2.35±1.52	£356±358
Accuracy with $\alpha=0.15$	0.782	0.756	2.77±1.94	£424±419
Accuracy with $\alpha=0.10$	0.837	0.791	3.47±2.45	£526±490

“LR” and “LWL” with LR” represent a logistic regression and a locally weighted learner with logistic regression as base classifier applied to all data. Other rows correspond to fully personalized classifiers with different values of α and criteria to select the sequence of biomarkers: AUC or accuracy.

TABLE II

RESULTS FOR THE CLASSIFICATION OF nMCI VERSUS cMCI PATIENTS GIVEN AS FINAL AUC, FINAL ACCURACY, NUMBER, AND COST OF THE BIOMARKERS

nMCI vs. cMCI	Final AUC	Final Accuracy	Number of biomarkers	Cost of biomarkers
LR	0.667	0.676	11±0	£1914±0
LWL with LR	0.653	0.732	11±0	£1914±0
Personalized classifiers minimizing the number of biomarkers				
AUC with $\alpha=0.20$	0.631	0.732	2.32±2.01	£760±465
AUC with $\alpha=0.15$	0.655	0.746	3.30±2.65	£903±484
AUC with $\alpha=0.10$	0.583	0.746	4.21±3.23	£1035±527
Accuracy with $\alpha=0.20$	0.638	0.746	2.27±2.05	£693±454
Accuracy with $\alpha=0.15$	0.668	0.718	3.41±2.95	£893±528
Accuracy with $\alpha=0.10$	0.586	0.746	4.65±3.54	£1075±556
Personalized classifiers minimizing the cost of biomarkers				
AUC with $\alpha=0.20$	0.496	0.648	3.18±2.29	£337±440
AUC with $\alpha=0.15$	0.597	0.690	4.25±2.73	£527±586
AUC with $\alpha=0.10$	0.582	0.746	5.32±2.82	£693±644
Accuracy with $\alpha=0.20$	0.473	0.648	3.45±2.30	£350±408
Accuracy with $\alpha=0.15$	0.624	0.676	4.56±2.94	£555±591
Accuracy with $\alpha=0.10$	0.535	0.704	5.38±2.91	£689±624

The abbreviations and acronyms are those of Table I.

weighted learner with logistic regression as base learner but no feature selection at all. Table II contains analogous results for the task cMCI–nMCI.

For illustration purposes, Fig. 2 shows the average, across subjects, of the cost of the biomarker selected at each iteration in the classifier of CN versus AD with $\alpha = 0.15$ when the system minimized the number or the cost of the biomarkers using the AUC as feature selection criterion.

V. DISCUSSION AND CONCLUSION

We tested a machine learning approach for personalized and cost-effective detection of AD based on: 1) locally weighted learning [1], [15] and 2) a sequential selection of biomarkers to reduce their number or cost for confident diagnosis [4]. Two classification tasks were addressed: CN–AD and cMCI–nMCI. The approach is closer to the clinical setting, where not all

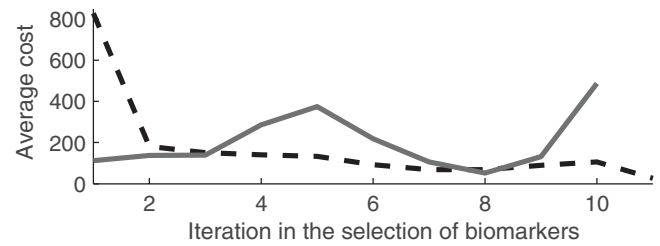


Fig. 2. Average cost (in GBP) of the biomarker included at each iteration for the personalized classification of CN versus AD subjects with $\alpha = 0.15$ and AUC as feature selection criterion for the modes minimizing the number (dashed black line) or cost (full gray line) of the biomarkers.

biomarkers are available at once. It also considers which previous cases are more relevant for the patient.

The personalized classifiers tried to minimize the number or cost of the biomarkers included in the process. In both modes, the classifications were considerably more cost-effective than those based on all variables as there were important reductions in the diagnosis cost. Minimizing the number of biomarkers led to classifiers with fewer but more expensive features. Fig. 2 suggests that expensive, but perhaps more informative, biomarkers tended to be selected in the first iterations of the process that minimized the number of tests. On the other hand, the system optimizing the cost tended to select inexpensive biomarkers first and only if these were not conclusive were more expensive tests chosen. The overall classification performance was better when the system tried to minimize the number of biomarkers. This improvement came with relatively modest additional cost. Nonetheless, both strategies still need a more detailed cost-effectiveness analysis.

We also carried out a preliminary inspection of the evolution of the criteria for biomarker selection with the number of iterations (results not shown due to space constraints). The results suggested that both accuracy and AUC are appropriate metrics to select the features since the improvement in performance tended to decrease monotonically from the first to the last iterations.

Our results are comparable to those computed in other cross-validated studies. Schemes that used MRI, biochemistry, and PET data simultaneously reached accuracies of about 0.93 and 0.75 for the tasks CN–AD and nMCI–cMCI, in that order [5], [6]. Another study classified progression from MCI to AD over a period of two years using MRI, biochemistry, and cognitive scores simultaneously with accuracy of 0.67 and AUC of 0.80 [7]. Yet, the results cannot be directly compared due to differences in the datasets and biomarkers. As a proof of concept, we considered only one marker from the MRI and PET [10], [11], but the inclusion of more advanced brain imaging features may improve the performance.

The *Pool of known cases* is a key part of the method. Interim analyses (not shown due to space constraints) indicate that the classification performance decreases when smaller subsamples of the Pool are considered. Thus, it is essential to include a large enough number of subjects in it. More extensive tests are needed to determine how many cases should be included in the Pool. In any case, in clinical practice, the Pool should be populated with

local data, which are more likely to reflect local life-style and environmental factors that might affect the disease.

According to current recommendations [13], the clinical diagnosis of AD and MCI should rely only on the patient's cognitive and behavioral symptoms. However, the biomarkers can increase or decrease the certainty that such symptoms are due to a pathological process of AD [8], [9], [13]. We conjecture that the personalized approach may also contribute to the diagnostic process, conveying supporting evidence that the patient's data match the profile of AD. This information is given as a probability, which summarizes the patient's data abnormalities, and shown to the clinician at every iteration, thus allowing him or her to monitor how the patient's data fit with known cases of the disease. The system assumes that not all biomarkers are available at once and it could be modified to account for other criteria than just cost when selecting the biomarkers (e.g., their level of invasiveness or risk of side effects). This could be done by deriving appropriate "modified costs" accounting for the relative effects of such factors (e.g., twice as much risk of side effects would lead to twice as much "modified cost"). The "modified costs" could also incorporate the clinician's prior expertise to guide the system towards specific biomarker combinations. We acknowledge that the classification performance of the system is not high enough to replace clinical diagnosis and that it is sometimes lower than that obtained considering all variables at once. However, this is not an inherent limitation of the method because its aim is to support, and not replace, the clinician, who must always make the final decision on clinical diagnosis.

Some limitations will be considered in future work. First, other classifiers [1] can be tested as base learners. Second, "modified cost" can be developed to account for additional factors in the selection of biomarkers. Finally, an independent validation set should be used to optimize α considering which values are clinically acceptable. Yet, the small number of ADNI subjects with all variables available limits our results and our ability to address such issues in this letter.

To sum up, the results are promising and might be used to support personalized diagnosis processes, while reducing the number or cost of the biomarkers needed for diagnosis. Future study is still needed but the framework presented in this letter could be readily extended to other biomarkers and diseases.

ACKNOWLEDGMENT

This article presents independent research commissioned by the NIHR under its Programme Grants for Applied Research

Programme. The views expressed in this article are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. The authors are grateful for the Associate Editor and Reviewers' feedback about the study.

REFERENCES

- [1] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed. New York: Elsevier, 2011.
- [2] S. Lemm, B. Blankertz, T. Dickhaus, and K.-R. Müller, "Introduction to machine learning for brain imaging," *NeuroImage*, vol. 56, no. 2, pp. 387–399, May 2011.
- [3] J. Ye, T. Wu, J. Li, and K. Chen, "Machine learning approaches for the neuroimaging study of alzheimer's disease," *Computer*, vol. 44, no. 4, pp. 99–101, Apr. 2011.
- [4] C.-L. Chi, W. N. Street, and D. A. Katz, "A decision support system for cost-effective diagnosis," *Artif. Intell. Med.*, vol. 50, no. 3, pp. 149–161, Nov. 2010.
- [5] C. Hinrichs, V. Singh, G. Xu, and S. C. Johnson, "Predictive markers for AD in a multi-modality framework: An analysis of MCI progression in the ADNI population," *NeuroImage*, vol. 55, no. 2, pp. 574–589, Mar. 2011.
- [6] D. Zhang, Y. Wang, L. Zhou, H. Yuan, and D. Shen, "Multimodal classification of Alzheimer's disease and mild cognitive impairment," *NeuroImage*, vol. 55, no. 3, pp. 856–867, Apr. 2011.
- [7] Y. Cui, B. Liu, S. Luo, X. Zhen, M. Fan, T. Liu, W. Zhu, M. Park, T. Jiang, J. S. Jin, and the Alzheimer's Disease Neuroimaging Initiative, "Identification of conversion from mild cognitive impairment to Alzheimer's disease using multivariate predictors," *PLOS ONE*, vol. 6, no. 7, p. e21896, Jul. 2011.
- [8] D. M. Holtzman, J. C. Morris, and A. M. Goate, "Alzheimer's disease: The challenge of the second century," *Sci. Transl. Med.*, vol. 3, no. 77, p. 77sr1, Apr. 2011.
- [9] K. Blennow, M. J. de Leon, and H. Zetterberg, "Alzheimer's disease," *Lancet*, vol. 368, no. 9533, pp. 387–403, Jul. 2006.
- [10] D. Holland, J. B. Brewer, D. J. Hagler, C. Fennema-Notestine, A. M. Dale, and the Alzheimer's Disease Neuroimaging Initiative, "Subregional neuroanatomical change as a biomarker for Alzheimer's disease," in *Proc. Nat. Acad. Sci. U.S.A.*, vol. 106, no. 49, pp. 20954–20959, Dec. 2009.
- [11] S. M. Landau, D. Harvey, C. M. Madison, R. A. Koeppe, E. M. Reiman, N. L. Foster, M. W. Weiner, and W. J. Jagust, "Associations between cognitive, functional, and FDG-PET measures of decline in AD and MCI," *Neurobiol. Aging*, vol. 32, no. 7, pp. 1207–1218, Jul. 2011.
- [12] J. Q. Trojanowski, H. Vandevertichele, M. Korecka, C. M. Clark, P. S. Aisen, R. C. Petersen, K. Blennow, H. Soares, A. Simon, P. Lewczuk, R. Dean, E. Siemers, W. Z. Potter, M. W. Weiner, C. R. Jack, Jr., W. Jagust, A. W. Toga, V. M.-Y. Lee, and L. M. Shaw, "Update on the biomarker core of the Alzheimer's Disease Neuroimaging Initiative subjects," *Alzheimers Dement.*, vol. 6, no. 3, pp. 230–238, May 2010.
- [13] C. R. Jack, M. S. Albert, D. S. Knopman, G. M. McKhann, R. A. Sperling, M. C. Carrillo, B. Thies, and C. H. Phelps, "Introduction to the recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease," *Alzheimers Dement.*, vol. 7, no. 3, pp. 257–262, May 2011.
- [14] D. of Health. (2011, Nov. 17). "2010-11 reference costs publication," [Online]. Available: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_131140. [Accessed: 18-Jun-2012]
- [15] C. G. Atkeson, A. W. Moore, and S. Schaal, "Locally weighted learning," *Artif. Intell. Rev.*, vol. 11, no. 1–5, pp. 11–73, Feb. 1997.