

Amyloid PET-Positive Predictability of Machine Learning Algorithm Based on MDS-OA β Levels

Young Chul Youn

Chung-Ang University Hospital

Jung-Min Pyun

Uijeongbu Eulji Medical Center

Hye Ryoun Kim

Chung-Ang University - First Campus: Chung-Ang University

Sungmin Kang

PeopleBio Inc.

Nayoung Ryoo

Seoul National University Bundang Hospital

Young Ho Park

Seoul National University Bundang Hospital

Hae-Won Shin

Chung-Ang University - Seoul Campus: Chung-Ang University

SangYun Kim (✉ neuroksy@snu.ac.kr)

Seoul National University Bundang Hospital <https://orcid.org/0000-0002-9101-5704>

Research

Keywords: Machine learning, Logistic regression, Oligomer, Amyloid β , Alzheimer's disease, Biomarker, Multimer Detection System, Blood, amyloid positron emission tomography

Posted Date: June 12th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-578834/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Abstract

Background: The Multimer Detection System-Oligomeric amyloid- β (MDS-OA β) level is a valuable blood-based biomarker for Alzheimer's disease (AD). We used machine learning algorithms trained using multi-center datasets to examine whether blood MDS-OA β values can predict AD-associated changes in the brain.

Methods: A logistic regression model using TensorFlow (ver. 2.3.0) was applied to data obtained from 163 participants (amyloid positron emission tomography [PET]-positive and -negative findings in 102 and 61 participants, respectively). Algorithms with various combinations of features (MDS-OA β levels, age, gender, and anticoagulant type) were tested 50 times on each dataset.

Results: The predictive accuracy, sensitivity, and specificity values of blood MDS-OA β levels for amyloid PET positivity were $78.16 \pm 4.97\%$, $83.87 \pm 9.40\%$, and $70.00 \pm 13.13\%$, respectively.

Conclusions: The findings from this multi-center machine learning-based study suggest that MDS-OA β values may be used to predict amyloid PET-positivity.

Introduction

Alzheimer's disease (AD) is a degenerative brain disease associated with the loss of independent living due to the deterioration of cognitive function and is linked to the gradual loss of cortical neurons.¹ Pathologically, it is characterized by cerebral amyloidosis that begins about 20 years before the onset of clinical symptoms.^{2,3} Early detection of AD is essential to patient outcomes and clinical trials of the disease modifying drugs.

Cerebral amyloidosis in AD has been evaluated based on the cerebrospinal fluid amyloid- β 1–42 levels and amyloid positron emission tomography (PET) imaging. However, these approaches are invasive, costly, and associated with interlaboratory variability, which limits their clinical use. Therefore, efforts have been made to develop blood-based amyloid-targeted biomarkers. The Multimer Detection System-Oligomeric amyloid- β (MDS-OA β) level is a valuable blood-based biomarker for AD; it is a modified sandwich immunoassay for measuring A β oligomerization in the plasma.^{4,5} This technique involves introducing synthetic A β into the plasma to trigger oligomerization of A β prior to antigen-antibody reactions and measuring the oligomerization tendency of plasma A β in AD patients.⁶ We have previously evaluated the role of MDS-OA β levels in differentiating AD and community-based healthy participants, showing high sensitivity and specificity of this approach.⁷ Meanwhile, we attempted to evaluate whether brain AD pathology can be predicted based on blood MDS-OA β levels in studies on the relationship between MDS-OA β , and magnetic resonance imaging or amyloid PET findings.^{8,9} Our previous studies have reported MDS-OA β cut-off, sensitivity, and specificity values in AD diagnosis and amyloid PET-positivity. In the present study, we aimed to use machine learning algorithms to examine amyloid PET-positivity prediction accuracy of MDS-OA β levels and factors that affect it using multi-center datasets.

Methods

This was an observational cross-sectional study to evaluate the role of MDS-OA β levels in predicting amyloid PET-positivity, using machine learning. This study was based on data obtained from the multi-center Alzheimer's Disease All Markers Study (ADAM), involving the Seoul National University Bundang Hospital and Chung-Ang University Hospital, and on data from the Dementia Overcoming Project in Korea (HI14C1251) (DOP), which is a clinical study on protein biomarker development and early diagnosis of dementia, involving five universities in Korea between May 2015 and December 2017. Patients who had simultaneously undergone amyloid PET and MDS-OA β assessments included 96 and 67 subjects from the ADAM and DOP studies, respectively. This study was approved by the institutional review board of the Seoul National University Bundang Hospital (B-2004-604-305) and Chung-Ang University Hospital (1792-012-300, C2012048[743]). This study was conducted in accordance with the Declaration of Helsinki. The need for written informed consent was waived due to the retrospective nature of the study.

Subjects consisted of 86 probable AD dementia patients diagnosed according to the criteria presented by National Institute on Aging-Alzheimer's Association,¹⁰ 41 mild cognitive impairment patients diagnosed according to this Association's criteria,¹¹ and 36 control subjects (28 of whom experienced subjective cognitive decline, according to guidelines by Jessen et al.,^{12,13} and 8 of whom experienced other neurodegenerative diseases, including 4 cases of frontotemporal dementia, 1 case of corticobasal syndrome, 1 case of Parkinson's disease dementia, and 2 cases of progressive supranuclear palsy).⁸

The amyloid PET ligands used in this study were [18F]Florbetaben (n = 82), [18F]Flutemetamol (n = 6), [18F]Florbetapir (n = 2), and [11C]Pittsburgh compound B (n = 2). Amyloid status was defined as positive or negative after visual assessment of scans by one nuclear medicine physicians and two neurologists.

Table 1

Demographic characteristics of the participants in the Alzheimer's Disease All Markers study (ADAM) and Dementia Overcoming Projects in Korea (DOP)

M: F	Age, years	Diagnosis	Amyloid PET	
		CTL: MCI: ADD	+	-
Total (n = 163)	59:104	68.20 \pm 9.45	36*:41:86	102 61
ADAM (n = 96)	42:54	69.39 \pm 9.45	15:27:54	68 28
DOP (n = 67)	17:50	66.49 \pm 8.83	21:14:32	34 33

*36 controls (5 disease controls, 27 cases of subjective cognitive decline). M, Male; F, Female; CTL, disease control; MCI, mild cognitive impairment; ADD, Alzheimer's disease dementia

Model training

Five variables were of interest in this study, including one target outcome
(<https://docs.google.com/spreadsheets/d/1Zig6hdCOpxjK7yfZ0AdIXtXrRJSVq1GeiwbsmZAHVLY/edit?>)

usp=sharing). One of the variables, MDS-OA β values, was based on raw data obtained from subjects' heparin or EDTA (ethylenediaminetetraacetic acid) anticoagulant plasma samples. Accordingly, we included the anticoagulant type as a variable. The other two variables were age and gender. The last variable, which was included as the target outcome, was the 'amyloid PET-positivity'. To distinguish the states of amyloid PET-positivity, according to the MDS-OA β levels in various combinations of variables, the machine learning algorithms were trained, using TensorFlow (<https://www.tensorflow.org>), a commonly used open-source software library for machine learning developed by Google based on Python.¹⁴ The following link will take the reader to the Python script for predicting amyloid positivity, given the values of MDS-OA β levels and the other five variables (<https://drive.google.com/file/d/1QS11f-1nJqb-emx0oc7r3q2J7gzkz1K/view?usp=sharing>).

In pre-processing, the dataset was randomly split into training (70%) and test (30%) datasets, using the train_test_split function from the scikit-learn library (<https://scikit-learn.org/>); the cardinality variables were normalized using mean and standard deviation. The feature (x_data) and outcome (y_data) variables were created in each dataset. The model was trained using the train dataset. The cost was calculated using logistic regression and minimized using the 'GradientDescentOptimizer'. Lastly, the accuracy, sensitivity, and specificity values of the amyloid PET prediction were calculated 50 times, using the randomly split test dataset with various combinations 'MDS-OA β ', 'MDS-OA β + age', 'MDS-OA β + gender', 'MDS-OA β + age + gender'.

Results

Logistic regression was used to assess the predictive accuracy of the MDS-OA β values alone and combined with other variables for amyloid PET-positivity findings. The predictive accuracy of the MDS-OA β levels was the highest at $78.16 \pm 4.97\%$; the corresponding sensitivity and specificity values were $83.87 \pm 9.40\%$ and $70.00 \pm 13.13\%$, respectively. The addition of gender feature to the MDS-OA β levels did not improve the model's predictive accuracy; meanwhile, the inclusion of age led to a reduction in the predictive accuracy to below 60%.

Table 2
Mean accuracies of logistic regression model predicting amyloid PET-positive of MDS-OA β on 50 trials

Features	MDS-OA β	MDS-OA β	MDS-OA β	MDS-OA β
	+ age	+ gender	+ age	+ gender
Accuracy	78.16 ± 4.97	56.70 ± 4.91	76.80 ± 5.55	57.65 ± 3.32
Sensitivity	83.87 ± 9.40	67.50 ± 16.22	77.90 ± 10.76	78.03 ± 16.46
Specificity	70.00 ± 13.13	41.43 ± 22.01	76.71 ± 11.56	27.71 ± 22.16

Discussion

Pyun et al.'s study using a conventional method of predicting positive amyloid PET findings based on MDS-OA β values showed 85% of sensitivity and specificity.⁸ Adding age and Mini-Mental Status Examination variables could change the sensitivity and specificity to 91% and 82%, respectively. In this study, the predictive accuracy of the machine learning algorithm was approximately 78%, and the sensitivity and specificity values were 83% and 70%, respectively. It is not a bad prediction accuracy considering the results from the data obtained from multi-center using different anti-coagulants.

It was observed that the inclusion of age and gender in the model did not improve its predictive accuracy. As the number of input features increases in a model based on a relatively small dataset, the model's predictive power may decrease due to overfitting.^{15,16} Fig. 1 illustrates MDS-OA β levels, stratified by gender, age, and anticoagulant type. Intuitively, we could not draw any useful decision lines to distinguish amyloid positivity by age and gender. Since there were only 163 subjects with amyloid PET and MDS-OA β values, including additional features in the model did not improve the model's predictive accuracy.

Pyun et al. used plasma MDS-OA β values estimated using heparin anticoagulants; however, the present study included both heparin-based ($n = 96$) and EDTA-based ($n = 67$) plasma measures because different centers use different anticoagulant types. As shown in Fig. 1, to predict amyloid PET-positivity, the decision line of EDTA plasma-based MDS-OA β values can be drawn at approximately 1.0 ng/ml; the corresponding heparin-based level is lower. Consequently, we included the anticoagulant type in the algorithm. Although the sensitivity and specificity values were somewhat lower than those previously reported, the present findings are acceptable, given the use of data obtained from different centers, using different anticoagulants.

We also examined other models, including three- or four-layer deep neural networks and light gradient boosting models; however, the corresponding accuracy values were below 70% and 76%, respectively, and none of the models showed performance superior to that of logistic regression.

Limitations

One limitation of this study is the use of imbalanced data; specifically, 62.58% of the participants were amyloid PET-positive; the ADAM participants showed a higher positive rate than the DOP participants. Due to the small number of participants, we could not create a random balanced dataset, precluding the creation of an algorithm based on such a dataset. Datasets obtained in a clinical setting are unlikely to be balanced. The imbalance observed in this study was deemed acceptable; however, it requires that the presented findings be interpreted with caution. However, our previous study did not show any differences in the accuracy of classification between imbalanced clinical and randomly selected balanced datasets.¹⁷ The other limitation of this study is that its retrospective design, including the use of data obtained by different projects, whereby PET examinations were performed at the discretion of the attending neurologist rather than using on a standardized protocol.

Conclusions

The machine learning algorithm using logistic regression and multi-center MDS-OA β values yielded satisfactory predictive accuracy, sensitivity, and specificity values suitable for the prediction of amyloid PET-positive findings. The MDS-OA β test is straightforward and cost-effective and allows for the evaluation of the oligomerization tendency of the plasma. This test can be used to screen patients for amyloid PET-positivity to identify if they are at a risk of developing AD. Further well-designed prospective studies are required to evaluate the amyloid PET-positive predictive values of MDS-OA β .

Abbreviations

AD

Alzheimer's disease; ADAM:Alzheimer's Disease All Markers Study; DOP:Dementia Overcoming Project in Korea; MDS-OA β :Multimer Detection System-Oligomeric amyloid- β ; PET:positron emission tomography; EDTA:ethylenediaminetetraacetic acid.

Declarations

Acknowledgements

The authors wish to thank Young Uk Park (Rowan Inc.), who contributed to the preprocessing of the data presented here. We also thank the staff at the Department of Neurology at Chung-Ang University Hospital and clinical research centers participating in the Dementia Overcoming Project (HI14C1251), Department of Neurology, Veterans Health Service Medical Center, Department of Neurology, Inha University School of Medicine, Department of Neurology, Ewha Womans University Mokdong Hospital, Department of Neurology, Dong-A University College of Medicine and Institute of Convergence Bio-Health, and Department of Neurology, Kyungpook National University School of Medicine.

Authors' contributions

YCY collected, analyzed, and interpreted the data, and drafted and revised the manuscript. SK measured MDS-OA β levels. JP analyzed and interpreted the data. HS, YHP, NR, and HK revised the manuscript. SYK designed and conceptualized the study and revised the manuscript. All authors read and approved the final manuscript.

Funding

The costs of manuscript publication, study design, data management, and manuscript writing were supported by the Ministry of Education of the Republic of Korea and National Research Foundation of Korea (NRF-2017S1A6A3A01078538).

Availability of data and materials

Some of the data were obtained from one of the Dementia Overcoming Projects in Korea, funded by the Ministry of Health & Welfare from May 2015 to December 2017 (HI14C1251). The whole dataset used in this study is accessible via the following link:

(<https://docs.google.com/spreadsheets/d/1Zig6hdCOpxjK7yfZ0AdIXtXrRJSVq1GeiwbsmZAHVLY/edit?usp=sharing>).

Ethics approval and consent to participate

Anonymous amyloid PET and MDS-OA β datasets were used retrospectively and provided by the Alzheimer's Disease All Markers study; one dataset was acquired from the Dementia Overcoming Projects in Korea. This study was approved by the institutional review board of the Seoul National University Bundang Hospital (B-2004-604-305) and Chung-Ang University Hospital (1792-012-300, C2012048(743)). The written informed consent requirement was waived due to the retrospective nature of this study.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Kang SH, Park YH, Lee D, et al. The Cortical Neuroanatomy Related to Specific Neuropsychological Deficits in Alzheimer's Continuum. *Dement Neurocogn Disord* 2019;18:77-95.
2. Hardy JA, Higgins GA. Alzheimer's disease: the amyloid cascade hypothesis. *Science* 1992;256:184-185.
3. Bateman RJ, Xiong C, Benzinger TL, et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. *N Engl J Med* 2012;367:795-804.
4. Wang MJ, Yi S, Han JY, et al. Oligomeric forms of amyloid-beta protein in plasma as a potential blood-based biomarker for Alzheimer's disease. *Alzheimers Res Ther* 2017;9:98.
5. Pyun JM, Kang MJ, Ryoo N, et al. Amyloid Metabolism and Amyloid-Targeting Blood-Based Biomarkers of Alzheimer's Disease. *J Alzheimers Dis* 2020;75:685-696.
6. An SSA, Lee BS, Yu JS, et al. Dynamic changes of oligomeric amyloid beta levels in plasma induced by spiked synthetic Abeta42. *Alzheimers Res Ther* 2017;9:86.
7. Youn YC, Lee BS, Kim GJ, et al. Blood Amyloid-beta Oligomerization as a Biomarker of Alzheimer's Disease: A Blinded Validation Study. *J Alzheimers Dis* 2020;75:493-499.

8. Pyun JM, Ryu JS, Lee R, et al. Plasma Amyloid-beta Oligomerization Tendency Predicts Amyloid PET Positivity. *Clin Interv Aging* 2021;16:749-755.
9. Youn YC, Kang S, Suh J, et al. Blood amyloid-beta oligomerization associated with neurodegeneration of Alzheimer's disease. *Alzheimers Res Ther* 2019;11:40.
10. McKhann GM, Knopman DS, Chertkow H, 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. *Alzheimers Dement* 2011;7:263-269.
11. Albert MS, DeKosky ST, Dickson D, 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. *Alzheimers Dement* 2011;7:270-279.
12. Jessen F. Subjective and objective cognitive decline at the pre-dementia stage of Alzheimer's disease. *Eur Arch Psychiatry Clin Neurosci* 2014;264 Suppl 1:S3-7.
13. Jessen F, Amariglio RE, van Boxtel M, et al. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimers Dement* 2014;10:844-852.
14. Rampasek L, Goldenberg A. TensorFlow: Biology's Gateway to Deep Learning? *Cell Syst* 2016;2:12-14.
15. Ying X. An Overview of Overfitting and its Solutions. *Journal of Physics: Conference Series* 2019;1168:022022.
16. Chicco D. Ten quick tips for machine learning in computational biology. *BioData Min* 2017;10:35.
17. Kang MJ, Kim SY, Na DL, et al. Prediction of cognitive impairment via deep learning trained with multi-center neuropsychological test data. *BMC Med Inform Decis Mak* 2019;19:231.

Figures

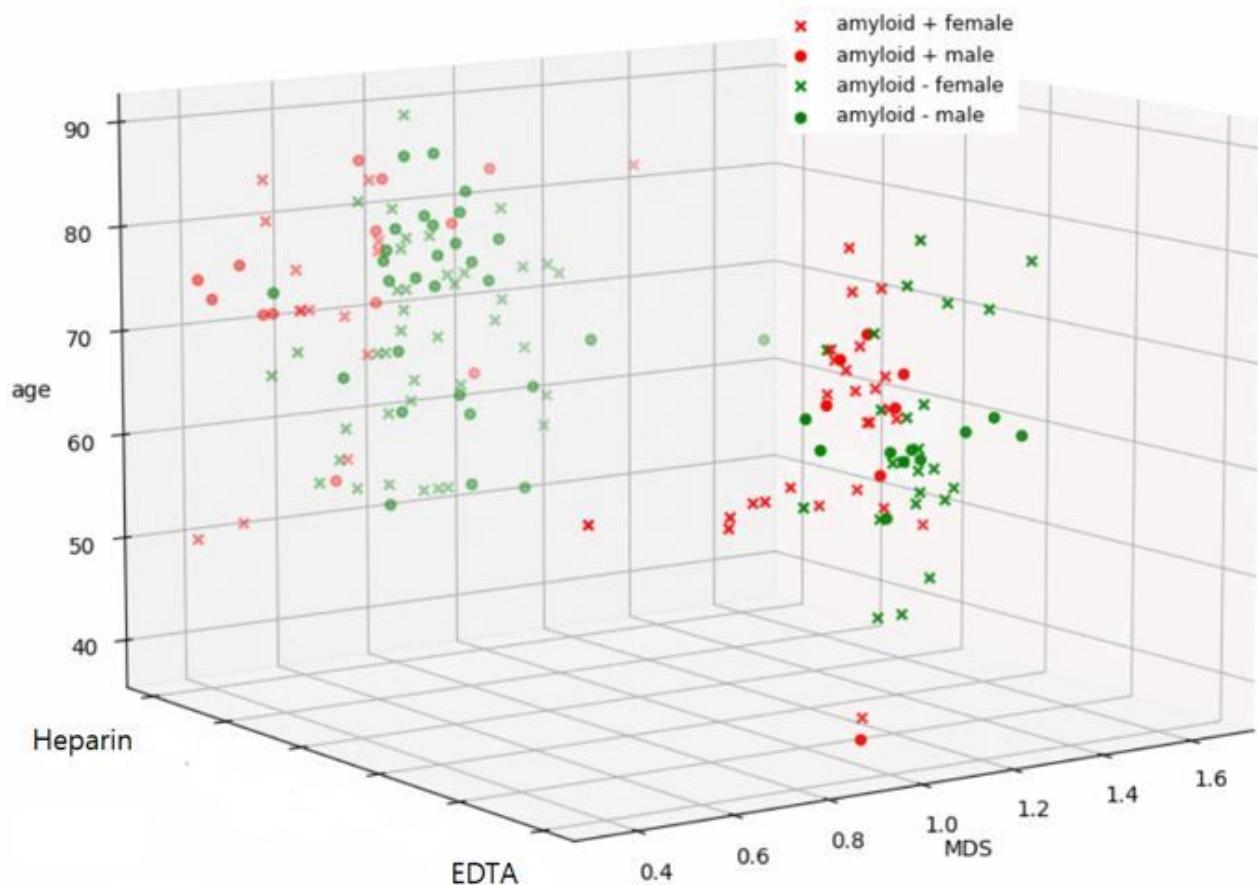


Figure 1

Levels of MDS-OA β , stratified by gender, age, and anticoagulant type
[\(<https://drive.google.com/file/d/14U-rKXLyHUVTKhRAtSoZHON-5YsjmnDn/view?usp=sharing>\)](https://drive.google.com/file/d/14U-rKXLyHUVTKhRAtSoZHON-5YsjmnDn/view?usp=sharing). amyloid +, amyloid PET-positive; amyloid -, amyloid PET-negative; EDTA, ethylenediaminetetraacetic acid.