



# Eyetracking Metrics in Young Onset Alzheimer's Disease: A Window into Cognitive Visual Functions

Ivanna M. Pavisic<sup>1\*</sup>, Nicholas C. Firth<sup>2</sup>, Samuel Parsons<sup>3</sup>, David Martinez Rego<sup>3</sup>, Timothy J. Shakespeare<sup>1</sup>, Keir X. X. Yong<sup>1</sup>, Catherine F. Slattery<sup>1</sup>, Ross W. Paterson<sup>1</sup>, Alexander J. M. Foulkes<sup>1</sup>, Kirsty Macpherson<sup>1</sup>, Amelia M. Carton<sup>1</sup>, Daniel C. Alexander<sup>2</sup>, John Shawe-Taylor<sup>3</sup>, Nick C. Fox<sup>1</sup>, Jonathan M. Schott<sup>1</sup>, Sebastian J. Crutch<sup>1</sup> and Silvia Primativo<sup>1\*</sup>

<sup>1</sup>Dementia Research Centre, Department of Neurodegenerative Diseases, Institute of Neurology, University College London, London, United Kingdom, <sup>2</sup>Centre for Medical Image Computing, Department of Computer Science, University College London, London, United Kingdom, <sup>3</sup>Centre for Computational Statistics and Machine Learning, Faculty of Engineering Science, Department of Computer Science, University College London, London, United Kingdom

## OPEN ACCESS

### Edited by:

Alessandra Rufa,  
University of Siena, Italy

### Reviewed by:

Jayasri Das Sarma,  
Indian Institute of Science Education  
and Research-Kolkata, India  
Christopher Charles Glisson,  
Michigan State University,  
United States

### \*Correspondence:

Ivanna M. Pavisic  
ivanna.pavisic.15@ucl.ac.uk;  
Silvia Primativo  
s.primativo@ucl.ac.uk

### Specialty section:

This article was submitted to  
Neuro-Ophthalmology,  
a section of the journal  
Frontiers in Neurology

Received: 06 April 2017

Accepted: 17 July 2017

Published: 07 August 2017

### Citation:

Pavisic IM, Firth NC, Parsons S,  
Rego DM, Shakespeare TJ,  
Yong KXX, Slattery CF, Paterson RW,  
Foulkes AJM, Macpherson K,  
Carton AM, Alexander DC,  
Shawe-Taylor J, Fox NC, Schott JM,  
Crutch SJ and Primativo S (2017)  
Eyetracking Metrics in Young Onset  
Alzheimer's Disease: A Window  
into Cognitive Visual Functions.  
Front. Neurol. 8:377.  
doi: 10.3389/fneur.2017.00377

Young onset Alzheimer's disease (YOAD) is defined as symptom onset before the age of 65 years and is particularly associated with phenotypic heterogeneity. Atypical presentations, such as the clinic-radiological visual syndrome posterior cortical atrophy (PCA), often lead to delays in accurate diagnosis. Eyetracking has been used to demonstrate basic oculomotor impairments in individuals with dementia. In the present study, we aim to explore the relationship between eyetracking metrics and standard tests of visual cognition in individuals with YOAD. Fifty-seven participants were included: 36 individuals with YOAD ( $n = 26$  typical AD;  $n = 10$  PCA) and 21 age-matched healthy controls. Participants completed three eyetracking experiments: fixation, pro-saccade, and smooth pursuit tasks. Summary metrics were used as outcome measures and their predictive value explored looking at correlations with visuo-perceptual and visuo-spatial metrics. Significant correlations between eyetracking metrics and standard visual cognitive estimates are reported. A machine-learning approach using a classification method based on the smooth pursuit raw eyetracking data discriminates with approximately 95% accuracy patients and controls in cross-validation tests. Results suggest that the eyetracking paradigms of a relatively simple and specific nature provide measures not only reflecting basic oculomotor characteristics but also predicting higher order visuo-spatial and visuo-perceptual impairments. Eyetracking measures can represent extremely useful markers during the diagnostic phase and may be exploited as potential outcome measures for clinical trials.

**Keywords:** young onset Alzheimer's disease, eye movements, eyetracking metrics, cognitive visual functions, machine learning, classification model

## INTRODUCTION

Alzheimer's disease (AD) is the most common major neurodegenerative dementia type (1). While characterized by gradual and progressive episodic memory impairment, it is also associated with other cognitive impairments such as executive dysfunction, language, praxis, and complex visual processing deficits (2–5). Several phase three clinical trials have recently failed and there are no

disease-modifying treatments available for AD (6, 7). Sensitive and sensible markers are needed to facilitate earlier diagnosis and to serve as outcome measures in clinical trials.

The focus of neuropsychological investigations in AD has previously been directed toward the study of anterograde episodic memory, and attentional and executive processes as primary consequences of AD and sources of functional impairment (8); by contrast, cognitive visual impairment has been widely overlooked (9). More recently, the presence of both low and high-level visual processing impairments has received more attention [for a review see Ref. (10)]. Deficits related to both the ventral and dorsal processing pathways have been described; impairments in object and facial recognition and color and pattern processing have been reported (11–13); abnormal performance has also been shown for tasks investigating visuospatial processing and motion perception (13–15). The presence of visual impairments has been associated with the severity of the disease (11, 13, 16, 17), leading to the possibility that visual testing could provide a method of screening and tracking AD. Studies suggest that cognitive visual deficits are more marked in young vs. late onset AD (18) and there may also be qualitative differences in the nature of the deficits. While some studies have highlighted the prominence of both ventral and dorsal stream deficits in late onset AD (13, 19), disproportionately impaired visuospatial ability has been described in young onset AD (18). Furthermore, the so called “visual variant” of AD—posterior cortical atrophy (PCA)—in which visual symptoms predominate (20–23) exhibits commonalities in cortical thinning with typical, amnesic AD particularly within temporoparietal regions (24), suggesting a continuum of visual impairment between typical AD (tAD) and PCA.

Posterior cortical atrophy is a progressive neurodegenerative syndrome mainly caused by AD pathology and characterized by progressive visuospatial and visuoperceptual dysfunction with relatively preserved memory, insight, and judgment (20). Individuals with PCA often manifest some or all of the features of Balint's syndrome such as simultanagnosia, oculomotor apraxia, optic ataxia, and environmental agnosia (21–23, 25–28). Not only differences but also similarities have been described between PCA and tAD in terms of visual processing deficits (29), emphasizing the need for further study on the cognitive visual deficits in AD.

Recently, eyetracking technology has become more widely available and the simplicity of the instruments needed to collect good quality eyetracking data has enabled the application of the methodology to clinical populations (29–32). The resulting literature has demonstrated the presence of oculomotor impairment in AD patients as compared to age-matched healthy controls: longer saccade latencies in pro-saccade tasks and lower accuracy than controls in anti-saccade tasks have been reported in AD (33–37), together with abnormalities in saccadic accuracy (38–41).

The study by Shakespeare et al. (29) represents, to our knowledge, the only direct comparison between the basic oculomotor characteristics of individuals with PCA, tAD, and healthy controls. PCA patients showed increased time to saccadic target fixation, increased first major saccade latency, and decreased saccadic amplitude as compared to both tAD and controls. The

patients with PCA also showed more frequent large intrusive saccades and lower longest period of fixation in the fixation stability task and lower pursuit and more saccades per trial in the smooth pursuit task as compared to controls. The authors also described impaired performance in the fixation stability test in tAD patients, who produced a large proportion of square wave jerks as compared to controls and individuals with PCA. As in PCA, tAD patients also had shorter maximum fixation period than the healthy controls. In the smooth pursuit task, both PCA and tAD patients showed significantly lower gain than the control group (29).

In cognitive psychology, eyetracking metrics have frequently been used to study higher order cognitive functions (42, 43). However, few studies in the field of dementia have utilized this potential to date. Crutcher and colleagues (44) and Richmond and colleagues (45) used a visual paired-comparison task and showed that eye movement metrics, such as number of fixations and fixation duration, can be indicative of short-term memory difficulties in a group of patients with mild cognitive impairment (MCI) as compared to age-matched controls. Fernández et al. (46) have investigated the semantic, working, and retrieval memory deficits in individuals with young onset AD by looking at differences with controls when reading high- and low-predictability sentences.

Despite the above-reported evidence of visual cognitive processing impairments and oculomotor deficits in both AD and PCA, little has been done so far to exploit eyetracking metrics as a route to explore deficits in visual cognition.

There are multiple potential advantages of using eyetracking metrics for studying visual cognitive processing in dementia. In contrast to many traditional neuropsychological assessments, eye movement recording does not require additional behavioral responses, such as button presses or verbal responses to make inferences about psychological changes. Eyetracking is also non-invasive and does not have contraindications, making it particularly well suited for patient studies. Moreover, modern eyetracking systems have excellent recording frames of up to 1,000 Hz, enabling the building of very large datasets (time series of  $x$  and  $y$  coordinates) in a relatively short amount of time (e.g., a 1,000 Hz system generates 600,000  $x$ - $y$  data coordinates for a 10-min recording session). Such qualities represent incentives to fully explore the possible contribution of eyetracking metrics to an accurate and sensitive diagnosis and as outcome markers for clinical trials. An increasingly practical approach to take advantage of the volume of eyetracking datasets involves the application of machine-learning methods, in which automatically generated feature vectors of individual participants may be used to assign categories to each participant.

In the present study, we explored the relationship between eyetracking and standard visual cognitive tests in individuals with young onset AD (both tAD and PCA). We extracted standard eyetracking summary metrics and used them to test the hypothesis according to which such metrics well correlate with visuoperceptual and visuospatial metrics derived from standard cognitive tests. We then applied a machine-learning approach to a proportion of the data to explore the possibility of automatically discriminating patients and healthy individuals

on the basis of raw eyetracking metrics only (i.e., time series of  $x$  and  $y$  coordinates). Therefore, our secondary hypothesis was that machine-learning classifiers would offer the discriminative power (47) for the diagnosis of young onset AD among healthy controls based on oculomotor profiles during a discrete task.

## MATERIALS AND METHODS

### Participants

The study was approved by the local Research Ethics Committee and all participants provided written informed consent according to guidelines established by the Declaration of Helsinki.

Data were collected from 36 individuals with young onset Alzheimer's disease (YOAD) (26 patients with tAD and 10 patients with PCA) and 21 age-matched healthy controls. Patients with PCA fulfilled standard criteria for PCA (21, 22). Patients with AD had a clinical diagnosis of probable AD and fulfilled the NIA (National Institute of Aging) clinical criteria (5). Healthy controls and patients with YOAD did not differ in terms of age at assessment (two sample  $t$ -test,  $t = 0.09$ ,  $p = 0.93$ ) or years of education (Wilcoxon Mann-Whitney  $U$ -test,  $z = 1.57$ ,  $p = 0.12$ ) (Table 1). Healthy controls and YOAD patients differed in terms of Mini-Mental State Examination (MMSE) scores (Wilcoxon Mann-Whitney  $U$ -test,  $z = 6.21$ ,  $p < 0.0001$ ). Within the YOAD patients, tAD and PCA were matched in terms of disease duration [tAD: 5.0 (2.8) years and PCA: 5.6 (3.4) years, Wilcoxon Mann-Whitney  $U$ -test,  $z = -0.50$ ,  $p = 0.62$ ] and MMSE scores [tAD: 20.1 (0.8) and PCA: 23.1 (1.5), Wilcoxon Mann-Whitney  $U$ -test,  $z = -1.78$ ,  $p = 0.07$ ].

All participants had a detailed neuropsychology assessment investigating memory, language, executive function, and vision. The battery included six standard visual tasks, which were the focus of the subsequent correlational analysis between eyetracking and traditional neuropsychological metrics. Early visual processing was examined using the shape detection subtest from the Visual and Object Space Perception battery (VOSP) (48) where individuals were presented with 20 patterns, 11 of which contained a faint cross. Participants were asked to express a judgment as to whether a faint cross was present or not with score ranging from 10 (chance) to 20. Visuo-perceptual processing was assessed using the fragmented letters and object decision VOSP subtests (48). In the former, individuals were asked to identify 20-fragmented capital letters presented one at a time with score ranging from 0 to 20. In the latter, they were asked to identify, among four silhouettes, which one represented a real life object [score ranged from 5 (chance) to 20]. Visuospatial processing was assessed using the dot-counting VOSP subtest (48) and letter cancellation test (49). In the dot-counting test, individuals were presented with 10 pages containing from 5 to 9 dots in different positions and asked to identify the number of dots without using their fingers (score range: 0–10). In the "A" cancellation task, participants were presented with an A4 sheet and asked to mark all the letters "A" embedded among 69 distractors (other letters) within 90 s. Last, single word recognition was assessed using the National Adult Reading Test (NART) (50) where individuals were asked to read 50 words aloud (score ranged from 0 to 50).

### Equipment

The experiment was run on a Dell 2120 desktop computer with a 23-inch screen at a viewing distance of 60 cm. Eye movements were recorded at 250 Hz using a head-mounted infrared video-based eye tracker (Eyelink II; SR Research). A chin rest was used to provide stability and maintain the viewing distance. The Eyelink system considered saccades using standard velocity and acceleration thresholds ( $30^\circ/s$  and  $8,000^\circ/s^2$ ) and automatically identified periods with no saccadic movement as fixations. A 9-point calibration and validation were performed prior to each experiment. All the data were obtained from recordings with an average Cartesian prediction error of  $<1^\circ$  during the validation. A drift correction procedure was used before each individual trial.

### Procedure

Three eyetracking experiments were performed:

*Fixation stability:* a red cross subtending at  $0.5^\circ$  of visual angle was presented in the center of the screen for 10 s. There was a practice trial followed by three test trials and participants were instructed to "look as closely as possible at the red cross without blinking for 10 s" (29, 51).

*Pro-saccade:* participants were initially presented with a black circle (subtending  $0.4^\circ$  of visual angle) having a white inner circle ( $0.1^\circ$ ) in the center of the screen lasting 500 ms. A blank screen was then displayed for 200 ms. After this, a target (black circle having a diameter of  $0.75^\circ$  and an inner white circle subtending  $0.25^\circ$ ) was shown. The target remained on the screen until a fixation of minimum 250 ms duration was made within an area of  $1.5^\circ$  of visual angle from the center of the target (interest area) or after 5,000 ms from target onset. The participant's task was to look at the target as quickly and accurately as possible when it appeared. The target appeared in one of 10 possible locations:  $5^\circ$ ,  $10^\circ$ , or  $15^\circ$  either on the left or on the right,  $5^\circ$  or  $10^\circ$  either up or down. There were four trials for each location, giving a total of 40 test trials. Trials were split into two blocks ( $n = 20$  each) with target locations randomized and balanced within each block. Four practice trials were used. The target positions were pseudo-randomized but their order was kept constant for all participants.

*Smooth pursuit:* the target was a red circle subtending  $0.5^\circ$  of visual angle in diameter. Twelve trials of target sinusoidal movement followed (horizontal:  $n = 6$ ; vertical:  $n = 6$ ). Two target velocities were used ( $10^\circ$  and  $20^\circ$  of visual angle/second). The frequency of the target oscillation was set at 0.25 Hz for a target speed of  $10^\circ/s$  and 0.5 Hz for a target speed of  $20^\circ/s$ . Each trial lasted 10 s. The task started with two practice trials. Participants were instructed to follow the target as closely as possible with their eyes.

### Eyetracking Summary Metrics

All eyetracking recordings were visually inspected using Data Viewer and trials and/or participants were excluded if there was a significant signal loss that would have interfered with the data analysis and interpretation of results. Overall, 5.4% of the trials

**TABLE 1** | Mean and SD demographic information and neuropsychology scores for the 36 patients with young onset Alzheimer's disease (YOAD) and 21 age-matched healthy controls.

	Max score	Controls (N = 21)	YOAD (N = 36)	N (%) below 5% <sup>a</sup>	Normative mean (SD)
<b>Demographics</b>					
Gender M:F		11:10	17:19	NA	NA
Age (years)		61.0 (5.3)	60.9 (5.2)	NA	NA
Education (years)		16.5 (3.2)	15.3 (2.7)	NA	NA
Disease duration (years)		NA	5.2 (2.9)	NA	NA
<b>Background neuropsychology</b>					
MMSE	30	29.5 (0.7)	20.9 (4.4)	NA	29.0 (1.3)
Visual acuity: Snellen <sup>b</sup>	6/9	NA	6/9	NA	NA
WASI vocabulary	80	69.0 (8.5)	53.4 (18.3)	1 (2.8%)	NA
WASI matrices	32	26.7 (2.7)	8.1 (7.1)	8 (22.2%)	NA
Digit span forward (max)	8	7.3 (1.2)	5.4 (1.5)	11 (30.6%)	NA
Digit span backward (max)	7	5.4 (0.9)	3.2 (1.5)	10 (27.8%)	NA
RMT for faces	25	24.7 (0.8)	19.5 (4.4)	11 (30.6%)	22.8 (1.9)
RMT for words	25	24.4 (1.4)	17.5 (3.2)	27 (75.0%)	23.7 (1.8)
GDA: oral	24	13.8 (6.6)	2.9 (4.7)	25 (69.4%)	11.95 (5.1)
<b>Early visual processing</b>					
Shape detection (VOSP)	20	19.5 (0.8)	18.0 (1.4)	NA	19.5 (0.7)
<b>Visuo perceptual processing</b>					
Object decision (VOSP)	20	18.2 (1.4)	14.7 (3.9)	12 (33.3%)	17.7 (1.9)
Fragmented letters (VOSP)	20	19.5 (0.7)	11.2 (7.2)	23 (63.9%)	18.8 (1.4)
<b>Visuospatial processing</b>					
Dot counting: n correct	10	9.9 (0.3)	7.6 (2.9)	16 (44.4%)	9.9 (0.2)
A cancellation time (s)	90	21.1 (6.0)	54.0 (22.8)	27 (75.0%)	20.5 (6.5)
<b>Word recognition</b>					
NART: number of errors	50	11.5 (8.0)	20.1 (11.0)	NA	NA

Mini-Mental State Examination (MMSE) (67); Cortical Vision Screening Test (CORVIST) (68); Wechsler Abbreviated Scale of Intelligence (WASI) (69); Wechsler Memory Scale Revised-digit span forwards and backwards (70); Matrices and Vocabulary subtest; Short Recognition Memory Test (RMT) for words and faces (71); oral Graded Difficulty Arithmetic (GDA) subtests addition and subtraction (72); Visual Object and Space Perception battery subtest object decision, shape detection, fragmented letters and dot counting (48); A cancellation (49); National Adult Reading Test (NART) (50). NA, not available.

<sup>a</sup>Percentage of scores below 5% percentile are shown for the YOAD group (the performance of controls did not reach this level).

<sup>b</sup>Median value is reported.

were excluded from the fixation stability task, 3.6% from the pro-saccade and 4% from the smooth pursuit tasks.

Blinks were identified and removed using Eyelink's automated blink detection and practice trials were discarded from the analysis. Vision was binocular but eye movements from the right eye were recorded. If a problem was detected (i.e., poor eyesight, watery, or dry eye) recordings were performed using the left eye.

Statistical analyses were carried out using Stata (v. 12.1).

### Fixation Stability

All participants performed the fixation stability task but two controls were excluded from analysis because of failure in signal detection. The relevant eyetracking metrics for the fixation stability task were large intrusive saccades, square wave jerks, and maximum fixation duration (29).

#### Number of Large Intrusive Saccades

The number of saccades with an amplitude greater than 2° of visual angle were identified as large intrusive saccades (29, 35).

#### Number of Square Wave Jerks

Square wave jerks were identified as saccades smaller than 2° in amplitude which took the gaze away from the target position, were followed within 300 ms by another saccade with a similar

amplitude (difference in amplitude <0.75°) and took the gaze back to the target position (29, 52).

#### Maximum Fixation Duration

The longest period of fixation on the target (length of time between saccades) was measured for each participant (29, 35).

#### Pro-Saccade

All participants performed the pro-saccade task but two individuals (a YOAD patient and a control) were excluded from the analysis due to a failure in signal detection. For this task, the following variables were taken into account: accuracy, time to fixate the target, and number of saccades necessary to fixate the target. For these metrics, fixations made within an area of 1.5° from the center of the target (interest area) were considered to have met the target.

#### Accuracy

This metric was defined as the ability of the participant to fixate the target (within 1.5° from its center) while it was presented on the screen.

#### Time to Fixate the Target

The time between the target onset and the first fixation reaching the target was calculated. Negative values due to anticipatory



saccades were either corrected if another fixation reaching the interest area was detected (0.09%) or removed if no fixation reached the interest area after target onset (0.14%).

### Number of Saccades Necessary to Fixate the Target

The minimum number of saccades necessary to fixate the target was calculated for each trial.

### Smooth Pursuit Task

All but three patients performed the smooth pursuit task. A control and a patient were excluded from the analysis due to a signal failure. The following variables were extracted: pursuit gain and proportion of time pursuing the target.

### Pursuit Gain

This was defined by the ratio between the eye and the target velocity (in the relevant direction). Saccades and blinks were excluded and only a ratio greater than 0.5 was considered as pursuit gain. This cut-off was applied to dismiss eye movements happening after anticipatory saccades and turnaround points in each trial.

### Proportion of Time Pursuing the Target

The proportion of time the participant spent pursuing the target during the trial was reported. This was calculated taking into account the number of samples considered as pursuit gain and multiplying this value by four (recordings were made at 250 Hz).

## Statistical Analysis

Differences in eyetracking metrics between the YOAD group and healthy controls were evaluated using linear regression models (clustered by participants) with robust SEs adjusted for repeated measures. Gender and age were considered as covariates for all metrics and stimulus distance and direction were considered as additional covariates for the pro-saccade task as well as target direction and velocity for the smooth pursuit task.

Normal distribution was assessed using the Shapiro–Wilk normality test. As the data were not normally distributed, a non-parametric measure of correlation (Spearman's correlations) was performed and the coefficients reported. Correlations were explored between all oculomotor metrics and six standard visual cognitive tests including measures of early visual processing (shape detection); visuo-perceptual processing (fragmented letters and object decision); visuospatial processing (dot counting and “A” cancelation); and single word recognition tests (NART).

## Machine Learning Classification Model

A machine-learning classification model is presented here as a proof of concept. As the statistical model aims to model movements in gaze location, the smooth pursuit experiment provided the most suitable data. The fixation stability and saccade experiments are designed to elicit 0 and one-gaze movements, respectively, and as such, their data did not provide enough information to discriminate between diagnostic classes on the basis of gaze movements. For this reason, the data from the smooth pursuit task were used in the pilot automatic classification procedure for the present study.

The automated classification procedure used the eyetracking data from the smooth pursuit task and modeled the movements in gaze location as the target moved. The gaze movements of each individual were used along with the statistical model to automatically generate feature vectors. These feature vectors were then used in a classification procedure that could predict the diagnoses of unseen individuals. The procedure, therefore, consisted of three components: (a) a fitted statistical model of each individual's data, (b) the generation of feature vectors for each individual *via* the fitted model, and (c) the classification of individuals *via* their feature vectors.

A hidden Markov model (HMM) (53) was the statistical model used. This considers movements in gaze location, and assumes that each gaze movement has an underlying “intended” movement direction. We have assumed that each gaze movement was a noisy application of one of the following possibilities: “no movement,” “left,” “right,” “up,” or “down.” The HMM is slightly non-standard, in that fixed transformations of each gaze movement are applied before they are passed to the sub-model associated with each underlying intended movement direction. Furthermore, knowledge of where the gaze “should” be moving to was incorporated *via* the location of the target at each time. Intended movement directions that were more aligned to the target direction were given a higher likelihood. This implemented the natural assumption that individuals would follow the target as long as they were able to.

If  $Y_{1:T}$  denotes the gaze movements for an individual over the course of one trial,  $U_{1:T}$  denotes the direction of the target from the current gaze location at each time, and  $X_{1:T}$  denotes the (unknown) underlying intended movement directions, then

$$\begin{aligned} \Pr(Y_{1:T}, U_{1:T} | X_{1:T}, \theta) &= \prod_{t=1}^T N(Y_t^{(X_t)} | \mu_{X_t}, \Sigma_{X_t}) f(U_t | X_t) \\ \Pr(X_{1:T} | \theta) &= \pi_1(X_1) \prod_{t=2}^T P(X_t | X_{t-1}) \end{aligned} \quad (1)$$

where  $Y_t^{(X_t)}$  are the transformed data and  $\theta = (\mu_{X_t}, \Sigma_{X_t}, \pi_1, P)$  are the parameters of the model.  $N(\mu, \Sigma)$  indicates the normal distribution with mean  $\mu$  and covariance matrix  $\Sigma$ , and  $f(U_t | X_t)$  are parameter-free logarithmic distributions that go to 0 as the target direction diverges from the intended direction  $X_t$ .

The model could be fit to the data using the EM algorithm (54). For every trial in the experiment, we fit the model to the data for one control. The individual that could follow the dot most accurately was subjectively chosen.

Once a model for each trial in the smooth pursuit experiment had been fitted, this could be used to generate feature vectors for each individual. The feature vectors are composites, made as the sum of “Fisher” feature vectors from the fitted models for each trial. Fisher feature vectors were computed as the gradient vectors of the data log-likelihood, evaluated at the fitted parameter settings. If  $Z^{(j)}$  is the Fisher feature vector for an individual in trial  $j$ ,  $Y_{1:T}$  are the gaze movement data for that trial, and  $\hat{\theta}$  is the fitted parameter vector, then

$$Z^{(j)} = \nabla_{\theta} \log \Pr(Y_{1:T} | \theta) \Big|_{\theta=\hat{\theta}} \quad (2)$$

which is a vector with as many dimensions as the model has parameters. The sum over trials gave the un-normalized feature vectors for each individual:

$$\hat{Z} = \sum_j Z^{(j)} \tag{3}$$

Once all un-normalized feature vectors for an experiment had been computed, they were normalized element-wise by their SDs (over individuals):

$$Z_i = \frac{\hat{Z}_i}{std\{\hat{Z}_i\}} \tag{4}$$

where  $Z_i$  and  $\hat{Z}_i$  are the  $i^{\text{th}}$  elements of the normalized and un-normalized feature vectors, respectively.

These feature vectors can be used in many classification algorithms. We chose to use logistic regression to classify individuals. As there are approximately as many individuals in the dataset as there are dimensions to the feature vectors, the logistic regression classifier required some regularization. We used the Bayesian methodology to regularize the classifier, placing a sparsity prior on the weights in the classifier. If  $d(Z)$  is the diagnosis of the patient associated with feature vector  $Z$  (with 0 meaning control and 1 meaning either tAD or PCA), and  $w$  are the weights of the model, then

$$\begin{aligned} Pr(d(Z) = 1 | w) &= \sigma(Zw) \\ Pr(w | \alpha) &= \prod_i N(w_i | 0, \alpha_i^{-1}) \end{aligned} \tag{5}$$

where  $\sigma()$  is the sigmoid function, and  $\alpha$  are the Bayesian hyper-parameters of the model.

The performance of the classifier was assessed through cross-validation tests. Leave-one-out, leave-two-out, and leave-half-out tests were performed. In each of these tests, the data was partitioned multiple times into training and test sets. The classifier was trained on the training sets, and its predictions for the test sets were compared to the true diagnoses. If the proportion of correct predictions for each group class was high, then the classifier had high predictive power on the data within the dataset.

## RESULTS

### Eyetracking Summary Metrics

Mean and SD performance metrics for fixation stability, pro-saccade and smooth pursuit tasks are shown in **Table 2**.

#### Fixation Stability

Results from the fixation stability task are represented in **Figure 1** and in **Table 2**.

#### Number of Large Intrusive Saccades

Young onset Alzheimer’s disease patients made a statistically significant higher number of large intrusive saccades compared to healthy controls [YOAD: 2.5 (4.3), healthy controls: 0.7 (1.7),  $t = 2.5, p = 0.02$ ].

**TABLE 2** | Mean and SD of fixation stability, pro-saccade, and smooth pursuit metrics for young onset Alzheimer’s disease (YOAD) patients and age-matched healthy controls.

		YOAD	Controls
		Mean (SD)	Mean (SD)
<b>Fixation stability</b>			
Number of large intrusive saccades		2.5 (4.3)*	0.7 (1.7)*
Number of square wave jerks		0.9 (1.6)	0.9 (1.6)
Maximum fixation duration (ms)		1,950.7 (1,352.8)*	2,908.5 (2,062.1)*
<b>Pro-saccade</b>			
Accuracy	Overall	0.85 (0.35)*	0.94 (0.24)*
	5°	0.90 (0.31)	0.96 (0.19)
	10°	0.84 (0.37)*	0.92 (0.27)*
	15°	0.81 (0.39)*	0.93 (0.26)*
	Up	0.86 (0.35)	0.90 (0.30)
	Down	0.86 (0.32)*	0.93 (0.22)*
	Right	0.83 (0.38)*	0.97 (0.18)*
	Left	0.86 (0.35)	0.93 (0.25)
	Time taken to reach the target (ms)	Overall	538.7 (682.3)*
5°		437.1 (583.0)*	306.1 (366.6)*
10°		613.8 (777.3)*	337.4 (333.4)*
15°		609.0 (650.4)*	358.8 (222.9)*
Up		537.8 (648.5)*	365.6 (468.4)*
Down		518.5 (571.2)*	329.2 (163.4)*
Right		591.0 (827.1)*	300.6 (180.7)*
Left		501.9 (613.8)*	333.6 (413.8)*
Saccades made to reach the target		Overall	3.1 (2.3)*
	5°	2.7 (1.9)*	2.0 (1.2)*
	10°	3.4 (2.7)*	2.1 (1.1)*
	15°	3.4 (2.4)*	2.2 (1.2)*
	Up	3.3 (2.4)*	2.3 (1.3)*
	Down	3.1 (2.5)*	2.1 (1.1)*
	Right	3.0 (2.1)*	2.1 (1.2)*
	Left	3.0 (2.2)*	2.0 (1.1)*
	<b>Smooth pursuit</b>		
Pursuit gain	Overall	1.4 (0.4)	1.4 (0.4)
	10°/s	1.4 (0.3)	1.4 (0.4)
	20°/s	1.3 (0.4)	1.4 (0.4)
	Horizontal	1.3 (0.3)	1.3 (0.3)
	Vertical	1.5 (0.4)	1.4 (0.4)
Prop. of time pursuing the target	Overall	0.4 (0.2)*	0.6 (0.2)*
	10°/s	0.5 (0.2)*	0.6 (0.1)*
	20°/s	0.3 (0.2)*	0.5 (0.2)*
	Horizontal	0.5 (0.2)*	0.7 (0.1)*
	Vertical	0.3 (0.1)*	0.5 (0.1)*

Statistically significant differences are highlighted in blue and marked with an asterisk (\*) (for specific  $p$  values see text).

#### Number of Square Wave Jerks

Healthy controls and YOAD patients did not show a statistically significant difference in terms of the average number of square

wave jerks [YOAD: 0.9 (1.6); healthy controls 0.9 (1.6);  $t = 0.4$ ,  $p = 0.60$ ].

### Maximum Fixation Duration

The longest period of fixation was significantly shorter for YOAD patients as compared to healthy controls [YOAD: 1,950.7 (1,352.8) ms; healthy controls: 2,908.5 (2,062.1) ms;  $t = -2.3$ ,  $p = 0.02$ ].

### Pro-Saccade

Results from the pro-saccade task are represented in **Figure 2** and in **Table 2**.

### Accuracy

Young onset Alzheimer's disease patients had an overall significantly lower accuracy compared to healthy controls ( $z = -2.1$ ,  $p = 0.04$ ). The effect of stimulus distance was statistically significant: the greater the distance from the center, the lower the accuracy ( $z = -3.5$ ,  $p < 0.001$ ). When looking at the accuracy for specific stimulus distances YOAD patients showed a trend toward lower accuracy than controls at  $10^\circ$  and  $15^\circ$  ( $z = -1.81$ ,  $p = 0.07$ ) but not at  $5^\circ$  ( $z = -1.64$ ,  $p = 0.10$ ). The effect of stimulus direction was not statistically significant when comparing the accuracy of YOAD patients and healthy controls ( $z = 0.79$ ,  $p = 0.40$ ).

### Time Taken to Fixate the Target

Young onset Alzheimer's disease patients took significantly longer to fixate the target compared to healthy controls ( $t = 3.7$ ,  $p = 0.001$ ). The effect of the stimulus distance was significant on the time taken to fixate the target: the time increased with stimulus distance ( $t = 4.2$ ,  $p < 0.001$ ). A statistical trend was observed for the effect of stimulus direction ( $t = -1.9$ ,  $p = 0.06$ ). YOAD patients took more time to reach the target at all stimulus distances (all  $p < 0.01$ ) as well as all stimulus directions (all  $p < 0.01$ ) compared to controls.

### Number of Saccades Necessary to Fixate the Target

Young onset Alzheimer's disease patients made a statistically higher number of saccades in order to fixate the target ( $t = 3.65$ ,  $p = 0.001$ ). The effect of the stimulus distance ( $t = -3.7$ ,  $p < 0.001$ ) and the stimulus direction ( $t = 4.9$ ,  $p < 0.001$ ) both had a significant effect on the number of saccades necessary to reach the target. The number of saccades increased with stimulus distance and the greatest number of saccades for both groups was made when the stimulus moved upwards. YOAD patients made a greater number of saccades to reach the target for all stimulus distances (all  $p < 0.001$ ) and stimulus directions (all  $p < 0.001$ ) compared to controls.

### Smooth Pursuit

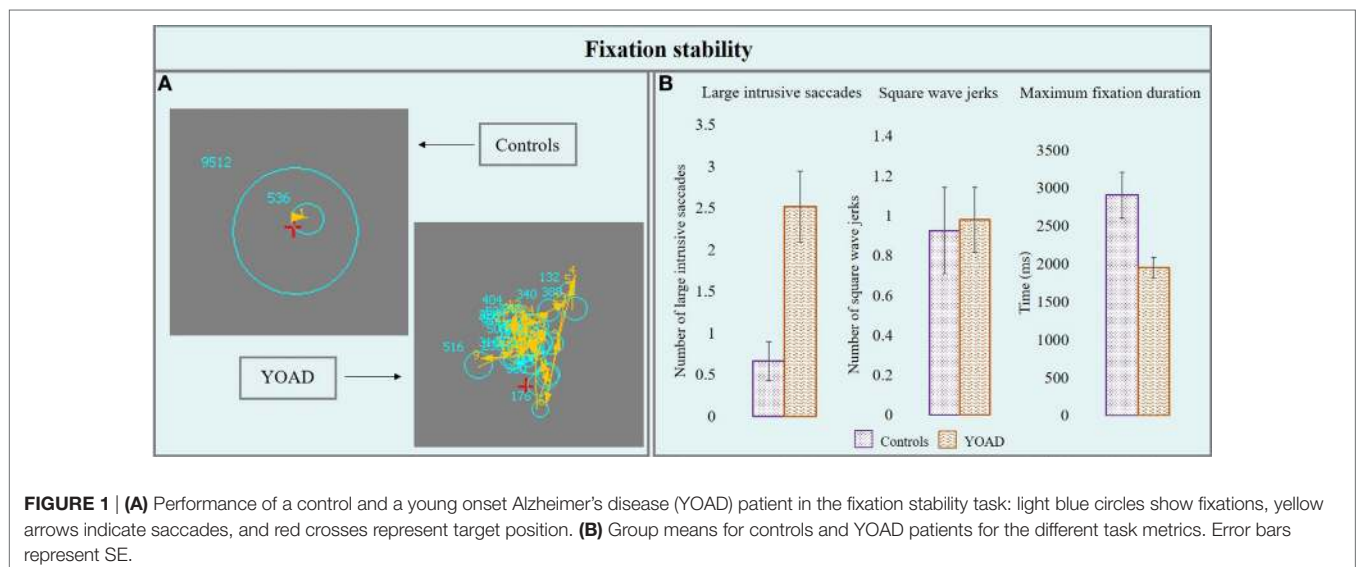
Results from the smooth pursuit task are represented in **Figure 3** and are reported in **Table 2**.

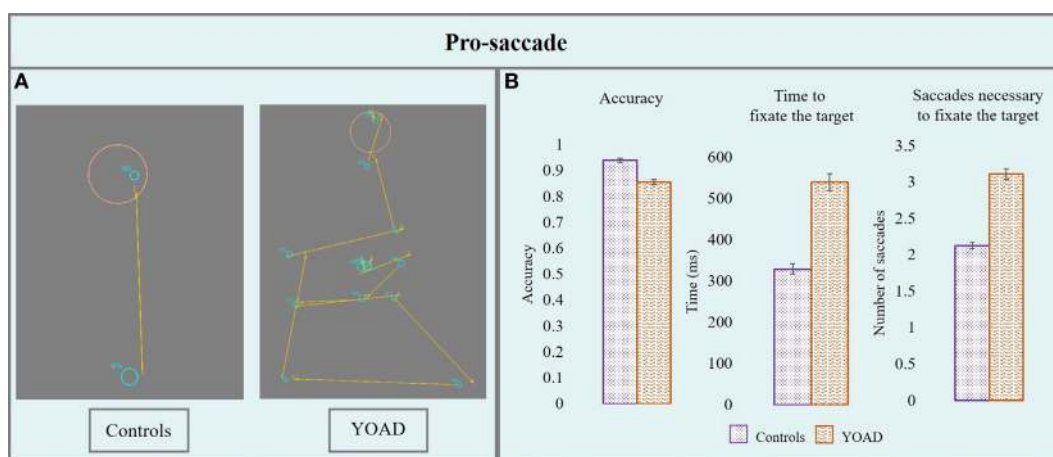
### Pursuit Gain

Young onset Alzheimer's disease patients and healthy controls did not differ in terms of pursuit gain ( $t = 0.52$ ,  $p = 0.60$ ). Stimulus velocity did not have a statistically significant effect ( $t = -0.95$ ,  $p = 0.30$ ) but stimulus direction did ( $t = 4.06$ ,  $p < 0.001$ ). Pursuit gain was closer to one (one: eye velocity = target velocity, eyes moving at the exact target's velocity) when the target moved horizontally (as opposed to vertically) for both groups.

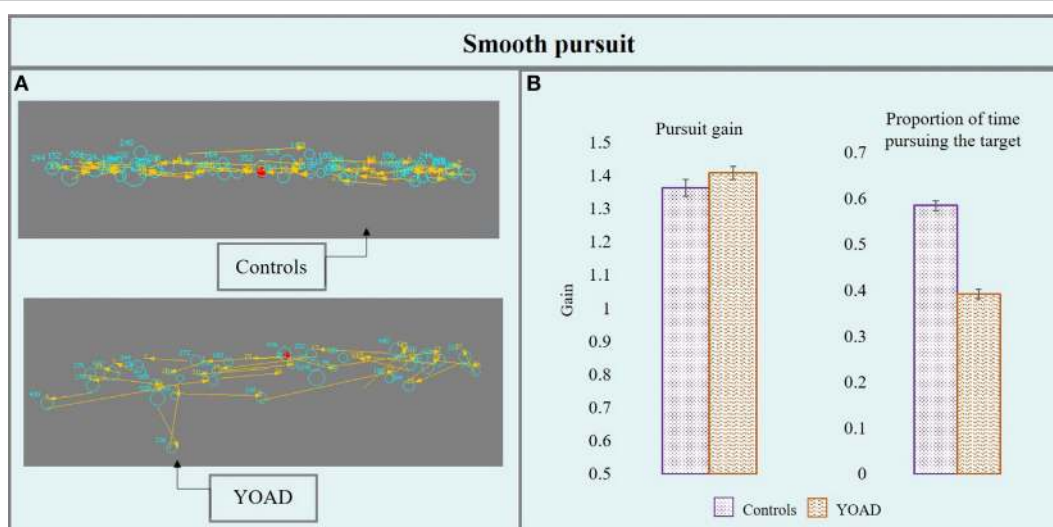
### Proportion of Time Pursuing the Target

Young onset Alzheimer's disease patients spent significantly less time pursuing the target compared to controls ( $t = -5.5$ ,  $p < 0.001$ ). Stimulus direction ( $t = -10.31$ ,  $p < 0.001$ ) and velocity ( $t = -10.84$ ,  $p < 0.001$ ) were both statistically significant: the proportion of time pursuing the target was greater when the target moved at  $10^\circ/s$  (as opposed to  $20^\circ/s$ ) and horizontally (compared to vertically). YOAD patients spent less time pursuing the target when it was presented either horizontally ( $t = -4.29$ ,  $p < 0.001$ ) or vertically ( $t = -5.79$ ,  $p < 0.001$ ) as well as at both stimulus





**FIGURE 2 | (A)** Performance of a control and a patient with young onset Alzheimer's disease (YOAD) in the pro-saccade task: light blue circles show fixations, yellow arrows indicate saccades, red crosses represent target position and orange circles outline the interest area ( $1.5^\circ$  from the center of the target). **(B)** Group means for controls and YOAD patients for the different task metrics. Error bars represent standard error.



**FIGURE 3 | (A)** Performance of a control and a young onset Alzheimer's disease (YOAD) patient in the smooth pursuit task: light blue circles show fixations, yellow arrows indicate saccades, and red crosses represent target position. **(B)** Group means for controls and YOAD patients for the different task metrics. Error bars represent SE.

velocities ( $10^\circ/s$ :  $t = -5.02$ ,  $p < 0.001$ ,  $20^\circ/s$ :  $t = -5.59$ ,  $p < 0.001$ ) when compared to controls.

### Comparisons between Eyetracking Metrics in tAD and PCA Patients

Given the aims of the present study, all the described analyses were conducted by examining individuals with tAD and PCA as part of the same group of individuals with YOAD. There were statistically significant differences between the two groups of patients in only three out of the eight eyetracking metrics all of which showed a poorer performance of PCA compared to tAD: maximum fixation duration (controls: 2,908.5 ms, tAD:

2,183.6 ms, PCA: 1,342.4 ms,  $t = 2.85$ ,  $p = 0.006$ ); time to fixate the target (controls: 328.7 ms, tAD: 428.4 ms, PCA: 924.4 ms,  $t = 5.13$ ,  $p < 0.001$ ) and number of saccades necessary to reach the target (controls = 2.13, tAD: 2.59, PCA: 4.86,  $t = 5.13$ ,  $p < 0.001$ ). PCA and tAD individual group performance was worse than that of healthy controls on all metrics for the three tasks (all  $p < 0.001$ ).

No statistically significant differences were observed between tAD and PCA in the remaining five-eyetracking metrics. There was no difference in the number of square wave jerks (controls: 0.93, tAD: 1.04, PCA: 0.82,  $t = 0.11$ ,  $p = 0.91$ ) or pursuit gain (controls: 1.36, tAD: 1.35, PCA: 1.62,  $t = 1.50$ ,  $p = 0.14$ ) between



tAD and PCA nor was there an effect of the phenotype in general ( $t = 0.36, p = 0.72$  and  $t = 1.50, p = 0.14$ , respectively). No differences were observed between tAD and PCA in the number large intrusive saccades (controls: 0.67, tAD: 1.81, PCA: 4.39,  $t = -1.56, p = 0.12$ ) and accuracy in the pro-saccade task (controls: 0.94, tAD: 0.89, PCA: 0.75,  $z = 1.55, p = 0.12$ ). For these two metrics patients performed worse than controls but only PCA were statistically worse ( $t = 2.38, p = 0.02$  and  $z = -2.62, p = 0.01$ , respectively). No difference was observed between the two groups of patients in the proportion of time pursuing the target (controls: 0.58, tAD: 0.40, PCA: 0.34,  $t = 1.11, p = 0.27$ ) and the two performed statistically worse compared to controls (both  $p < 0.001$ ).

## Relationship between Oculomotor Metrics and Standard Visual Cognitive Tests

In Table 3 coefficients and  $p$  values of correlations between estimates of visual cognitive processing and eyetracking metrics for the fixation stability, pro-saccade and smooth pursuit tasks are reported.

### Fixation Stability

Statistically significant negative correlations were observed between the number of large intrusive saccades and the scores on the following visual cognitive tests: object decision, fragmented letter, and dot counting. The association between the number of square wave jerks and the score in the object decision test approached statistical significance ( $p = 0.06$ ) as did the association between maximum fixation duration and dot-counting task ( $p = 0.07$ ) (see Figure 4).

### Pro-Saccade

Statistically significant negative correlations were reported between the time taken to fixate the target and the following visual cognitive tests: object decision, fragmented letters, and dot counting. The association between the time taken to fixate the target and the scores corresponding to the “A” cancelation task approached statistical significance ( $p = 0.07$ ).

Statistically significant negative correlations were found between the number of saccades necessary to fixate the target and the shape detection, object decision, fragmented letters, and dot-counting tests as well as a positive correlation with the “A” cancelation test scores (see Figure 5).

### Smooth Pursuit

Statistically significant correlations were observed between visual cognitive scores and both the pursuit gain and the proportion of time spent pursuing the target during the trial. In particular, the pursuit gain scores for YOAD patients negatively correlated with the object decision and dot-counting tests. The proportion of time spent pursuing the target positively correlated with the fragmented letters, dot counting and negatively with the “A” cancelation test scores (see Figure 6).

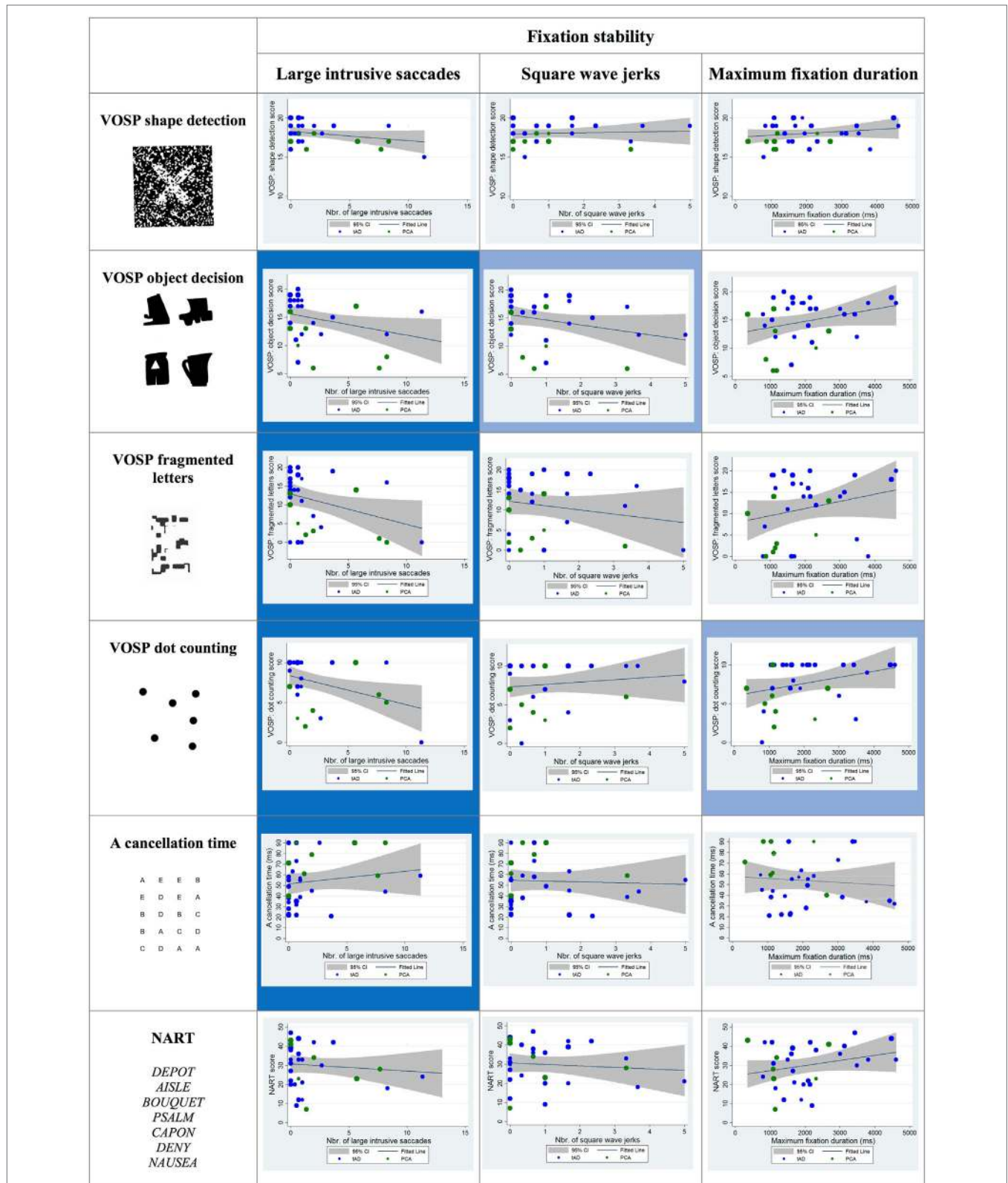
## Machine-Learning Classification Model

Fitting the HMM to each trial of the smooth pursuit experiment resulted in fitted parameters that conformed to expectations. In particular, the fitted model placed significantly more likelihood on the movement directions that followed the target than on any other direction. The results of the logistic regression classifier using the automatically generated feature vectors were able to discriminate with 95% accuracy patients and controls. Feature

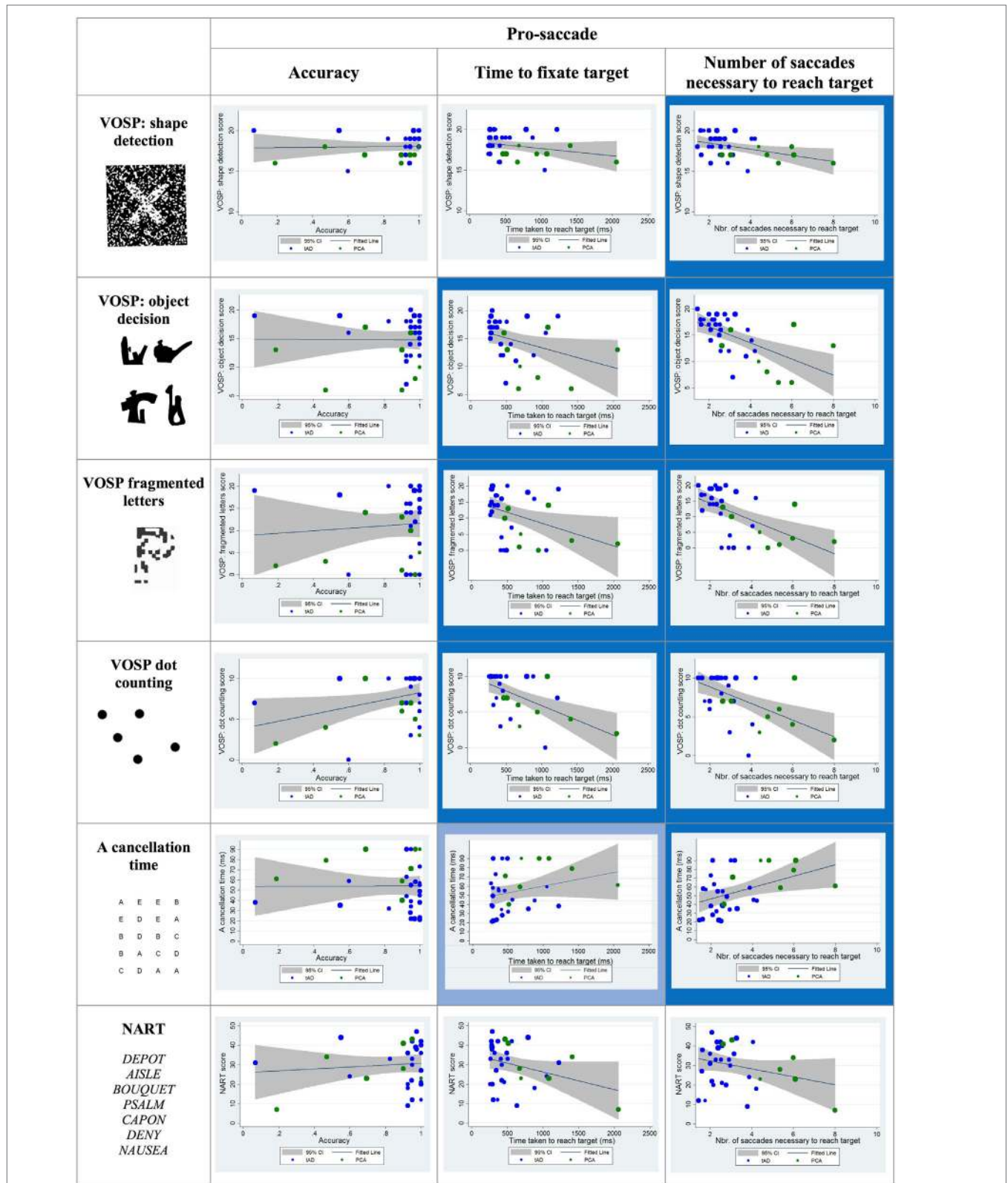
**TABLE 3** | Spearman's rank coefficient (Spearman's rho) and  $p$  values for correlations between visual cognitive tests and eyetracking metrics for the fixation stability, pro-saccade and smooth pursuit tasks.

	Eyetracking metrics								
	Fixation stability			Pro-saccade			Smooth pursuit		
	No. of large intrusive saccades	No. of square wave jerks	Max. fixation duration (ms)	Accuracy	Time to reach the target (ms)	Saccades made to fixate the target	Pursuit gain	Prop. of time pursuing the target	
Visual Cognitive Test	VOSP shape detection	$r = -0.16$ $p = 0.38$	$r = 0.04$ $p = 0.81$	$r = 0.19$ $p = 0.28$	$r = -0.26$ $p = 0.14$	$r = -0.19$ $p = 0.71$	$r = -0.38$ $p = 0.03^*$	$r = -0.09$ $p = 0.61$	$r = 0.04$ $p = 0.81$
	VOSP object decision	$r = -0.49$ $p = 0.003^*$	$r = -0.32$ $p = 0.06$	$r = 0.26$ $p = 0.13$	$r = 0.09$ $p = 0.61$	$r = -0.44$ $p = 0.009^*$	$r = -0.64$ $p < 0.001^*$	$r = -0.39$ $p = 0.03^*$	$r = 0.29$ $p = 0.11$
	VOSP fragmented letters	$r = -0.41$ $p = 0.02^*$	$r = -0.16$ $p = 0.36$	$r = 0.26$ $p = 0.14$	$r = 0.14$ $p = 0.44$	$r = -0.40$ $p = 0.02^*$	$r = -0.61$ $p < 0.001^*$	$r = -0.23$ $p = 0.22$	$r = 0.41$ $p = 0.02^*$
	VOSP dot counting	$r = -0.48$ $p = 0.005^*$	$r = 0.07$ $p = 0.70$	$r = 0.32$ $p = 0.07$	$r = 0.18$ $p = 0.32$	$r = -0.60$ $p = 0.002^*$	$r = -0.54$ $p < 0.001^*$	$r = -0.46$ $p = 0.01^*$	$r = 0.66$ $p < 0.001^*$
	A cancelation time	$r = 0.29$ $p = 0.10$	$r = -0.15$ $p = 0.42$	$r = -0.10$ $p = 0.57$	$r = -0.04$ $p = 0.79$	$r = 0.32$ $p = 0.07$	$r = 0.45$ $p = 0.009^*$	$r = 0.25$ $p = 0.19$	$r = -0.49$ $p = 0.006^*$
	National adult reading test	$r = -0.23$ $p = 0.18$	$r = -0.06$ $p = 0.72$	$r = 0.14$ $p = 0.44$	$r = 0.08$ $p = 0.64$	$r = 0.18$ $p = 0.32$	$r = -0.08$ $p = 0.66$	$r = -0.14$ $p = 0.48$	$r = 0.03$ $p = 0.88$

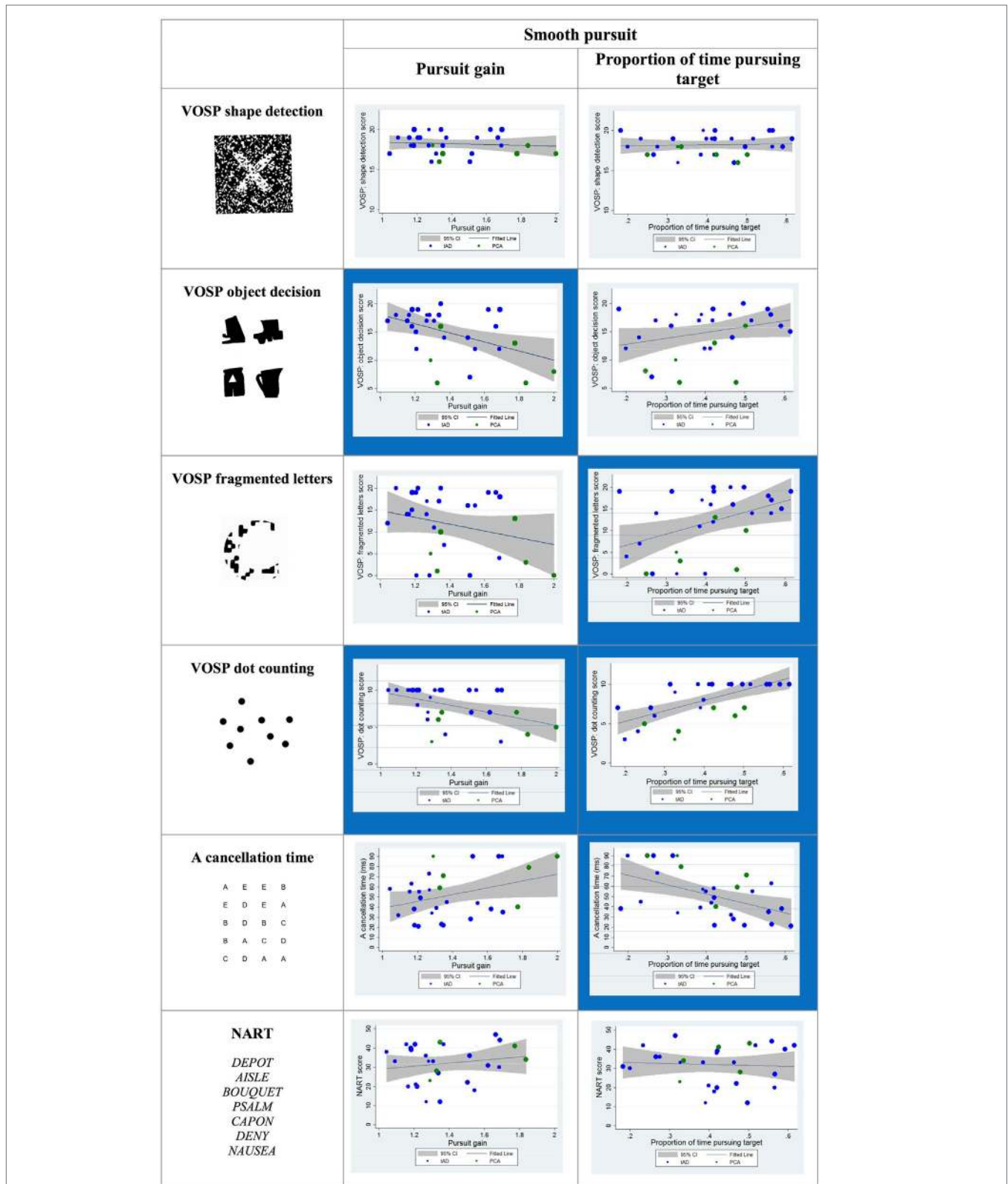
Statistically significant correlations are highlighted in blue and their  $p$  values marked with an asterisk (\*).



**FIGURE 4** | Scatter plots showing correlations between fixation stability metrics and visual cognitive scores for young onset Alzheimer’s disease patients. Statistically significant correlations are marked with a blue background and statistical trends ( $p > 0.05$  and  $p < 0.10$ ) in a lighter blue. Best fit line and 95% CI are shown. Each data point corresponds to a participant and its size is proportionate to the Mini-Mental State Examination score (indication of disease severity). typical AD (tAD) patients are shown in blue and posterior cortical atrophy (PCA) in green.



**FIGURE 5** | Scatter plots showing correlations between pro-saccade metrics and visual cognitive scores for young onset Alzheimer's disease patients. Statistically significant correlations are marked with a blue background and statistical trends ( $p > 0.05$  and  $p < 0.10$ ) in a lighter blue. Best fit line and 95% CI are shown. Each data point corresponds to a participant and its size is proportionate to the Mini-Mental State Examination score (disease severity), typical AD (tAD) individuals are shown in blue and posterior cortical atrophy (PCA) in green.



**FIGURE 6** | Scatter plots showing correlations between smooth pursuit metrics and visual cognitive scores for young onset Alzheimer’s disease patients. Statistically significant correlations are marked with a blue background and statistical trends ( $p > 0.05$  and  $p < 0.10$ ) in a lighter blue. Best fit line and 95% CI are shown. Each data point corresponds to a participant and its size is proportionate to the Mini-Mental State Examination score (disease severity). typical AD (tAD) individuals are shown in blue and posterior cortical atrophy (PCA) in green.



**TABLE 4** | Table showing the results of cross-validation tests for the predictive power of the Bayesian logistic regression classifier.

Test	Actual diagnosis	Predicted diagnosis	
		Control	Young onset Alzheimer's disease (YOAD) patient
L-1-O	Control	0.95	0.05
	YOAD patient	0.03	0.97
L-2-O	Control	0.95	0.05
	YOAD patient	0.03	0.97
L-H-O	Control	0.94	0.06
	YOAD patient	0.04	0.96

*Leave-one-out (L-1-O) tests take each individual in turn and train the classifier on all other individuals' feature vectors. The diagnostic status of that individual is predicted by the classifier and compared to the actual diagnosis (YOAD patient vs. control). Leave-two-out (L-2-O) tests take each possible pair of one control individual and one patient, train on all other individuals, and then predict the diagnosis of the original pair. The leave-half-out (L-H-O) test takes 500 random partitions of the data, with half of each diagnostic class in each partition, trains on one partition and predicts the diagnoses of the other. The columns of the table represent predicted diagnostic classes, and the rows represent actual diagnoses.*

vectors generated as described above are effective at separating the patients in the dataset into their diagnostic classes. The results of each test for the experiment are shown in **Table 4**. As can be seen, the predicted diagnoses show at least 95% accuracy for both diagnostic classes (patients vs. controls). While these results are very promising, they do only relate to the data from a relatively small number of individuals on a single test. An expanded experimental set-up, with a larger sample size, would be able to further verify the utility of this methodology for accurately predicting the diagnoses of previously unseen individuals.

## DISCUSSION

In the present study, we examined basic oculomotor metrics in individuals with YOAD and explored the link between such metrics and measures of visual cognition.

Results confirmed that patients have abnormal eye movement patterns in fixation stability, saccade and smooth pursuit tasks as compared to age-matched healthy controls. In the fixation stability task, YOAD patients showed a larger number of large intrusive saccades and shorter fixation duration. In the pro-saccade task, they were less accurate, required a longer amount of time and greater saccadic movements to fixate the target. Finally, in the smooth pursuit task, YOAD patients spent a shorter amount of time pursuing the target and made more interruptive saccades compared to controls. Results also demonstrated that these established basic control and movement metrics were significantly associated with measures of higher order visual cognition. In the fixation stability task the number of large intrusive saccades negatively correlated with performance in the object decision, fragmented letters and dot-counting tests. Pro-saccade metrics such as the time taken to reach the target and the number of saccades made in the process negatively correlated with object decision, fragmented letters, and dot-counting test scores. Additionally the "number of saccades necessary to reach the target" also correlated with shape detection scores and as the

time taken to complete the "A" cancellation task. The eyetracking metrics extracted from the smooth pursuit task correlated with object decision, fragmented letters, dot counting and time to complete the "A" cancellation tests.

To our knowledge, this is the first time that the systematic impairment of basic oculomotor functions is reported in patients with young onset dementia as a single group and that the relationship between such impairment and visual cognition is explored.

Our data underline the extent of visual cognitive impairments in individuals with YOAD (18). Awareness of dementia-related visual dysfunction in tAD is increasing (11, 12, 55) but, especially in the early stages of the disease, sensitive measures are required in order to highlight subtle changes that can potentially be discriminated from normal aging. For example, a recent study has shown that an eyetracking behavioral task can predict the conversion from cognitively normal to MCI and from MCI to AD up to three years prior to a change in clinical diagnosis (56). The presence of both eye movement deficits and impairments in visual processing is in accordance with the neuroimaging and neuropathological literature showing that in AD multifocal neuronal degeneration affects visual areas in the occipital, temporal and parietal lobes (57–59) and subcortical regions such as pulvinar (60) that process visual information and orient eye movements accordingly (61, 62).

The present study provides preliminary evidence suggesting the potential use of eyetracking metrics as markers of high-level vision and other cognitive domains. Examining the pattern of correlations, it should be noted that some eyetracking metrics, especially the "number of saccades needed to fixate the target," correlated with most of the visual cognitive tests, possibly reflecting non-specific associations with disease severity or the ubiquity of pro-saccade generation deficits in YOAD patients. However, not all eyetracking metrics had such widespread associations. Particularly evident were the impact of "large intrusive saccades" and "time to reach a pro-saccade target" upon the visuospatial tests (object decision and fragmented letters) and the visuospatial dot-counting test. All three of these tests require scanning over relatively small visual areas and across multiple discrete perceptual items, all of which are relevant to the task demands. By contrast, neither of these eyetracking metrics was correlated significantly with the "A" cancellation test, in which visual attention must be deployed over a much wider visual area and across items, only a small proportion of which constitute task-relevant targets. Also possibly noteworthy are the significant correlations between the "proportion of time pursuing target" metric from the smooth pursuit task and the fragmented letter, dot counting and "A" cancellation tests. Unlike the other tests with which these correlations were not observed, these three tests all require participants to trace a specific continuous visual route through separated stimuli, whether that route pertains to the shape of a large letter (fragmented letter), the path through a group of dots that permits them to be counted (and not accidentally re-counted) in an efficient manner (dot counting), or the line-by-line orderly searching for target "A"s among other distractor letters (A cancellation).

Of equal note is the relative absence of correlations between eyetracking metrics and either shape detection or reading. In the case of shape detection, which was only correlated with the

number of saccades required to reach the pro-saccade target, this may relate in part to the limited dynamic range of the test (all patient scores between 15 and 20 out of 20). However, for single word reading, with which there were no significant correlations and which with small print can be achieved in a single fixation, disordered eye movements such as large saccadic intrusions appear to have relatively little impact on accuracy (though note that reading latencies, if recorded, may have elicited a different result).

Naturally, such a qualitative examination of patterns of association has inherent limitations in determining causal relations between cognitive functions and observed behaviors. However, the current data arguably provide a useful starting point for generating testable hypotheses regarding the ability of certain eye movement patterns and paradigms to index specific cognitive abilities and deficits among dementia patients and other clinical populations.

Eyetracking-based measures of cognition may offer certain advantages over traditional pen and paper-based cognitive tests in some dementia contexts. Eyetracking data by definition do not suffer from ceiling and floor effects, which are instead common problems when exploring cognitive performance in patients (floor effect) and comparing it with performance in age-matched healthy controls (ceiling effect). Tasks such as fixation stability, saccade generation and smooth pursuit require minimal verbal instructions. Eye movement metrics derived from appropriate test designs may also be less vulnerable to the practice effects normally observed in standard cognitive testing, allowing for re-testing in the context of longitudinal assessments or before/after a trial phase.

One further potential advantage of eyetracking-based measures of visual cognition and other cognitive capacities is the type and scale of data generated. The large datasets that can be extracted in terms of time series of  $x$  and  $y$  coordinates open up new avenues of statistical analysis on individual trials. This may contribute to the design of shorter, less stressful cognitive assessments for patients. In the current study, we have provided proof of principle evidence for the feasibility of using an eyetracking dataset as the input for a machine-learning classification model to discriminate YOAD patients and controls on the basis of eye movements alone. While eyetracking in isolation is unlikely to ever be a primary determinant of clinical decision making, the application of a machine-learning approach to such examinations may add value in detecting change in at risk and presymptomatic individuals, monitoring progression over time (especially in the context of clinical trials), improving the discrimination and characterization disease and syndromic phenotypes.

The study had a number of potential limitations. First, the sample size was relatively small, and thus we have not been able to clarify whether the eye movement impairment is more widely a consequence of the disease severity rather than being a direct indication of visual cognition. Nonetheless, previous studies exploring eyetracking metrics in individuals affected by different types of dementia but matched for disease severity have shown that eye movement deficits can be disease-specific (33, 63). Second, the study did not include markers of focal and sustained attention, deficits in which may have contributed to

both eyetracking metrics and visual cognition estimates. Third, we only found limited evidence of differences between PCA and tAD in terms of eyetracking metrics. This result can have several explanations. The PCA group size was very small and the individual variability within each group was very large, as is frequently described within this literature (28, 64, 65). This might reflect the biological reality of YOAD where a greater prevalence of visual deficits across the population has already been suggested (18) and/or the phenotypic continuum of visual impairment across the tAD-PCA spectrum (64, 66). To address further the issue of individual phenotypic differences, rather than a binary PCA/tAD diagnostic category, future studies involving a larger sample size should take into account the possibility of using a quantitative continuous metric of visual cognitive impairment, such as a ratio of memory to perceptual and/or spatial scores.

In conclusion, we have demonstrated that basic oculomotor metrics can provide information about not only the oculomotor system and its functionality *per se*, but also about high-level visual cognition. We have also shown that such metrics can be used in a machine-learning approach to discriminate between YOAD patients and healthy controls. Visual deficits represent a common feature in AD and eyetracking metrics may have potential as sensitive markers, particularly as outcome measures for clinical trials.

## ETHICS STATEMENT

The study was approved by the local Research Ethics Committee and all participants provided written informed consent according to guidelines established by the Declaration of Helsinki.

## AUTHOR CONTRIBUTIONS

IP and SPrimativo: conception of the work, analysis and interpretation of data, drafting, final approval of the manuscript, and agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are investigated and resolved. SPArsons and DR: analysis of the data, drafting, final approval of the manuscript, and agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are investigated and resolved. TS, CS, RP, KM, AC, KY, NFirth, AF, DA, JT, NFox, and JS: contribution to the conception and design of the study, collection of the data, revising it, final approval of the manuscript, and agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are investigated and resolved. SC: conception of the work, interpretation of data, contributed to the drafting of the manuscript, final approval, and agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of the work are investigated and resolved.

## ACKNOWLEDGMENTS

This work was supported by Alzheimer's Research UK (ARUK 2012-6-ICE) and was undertaken at UCLH/UCL, which received

a proportion of funding from the Department of Health's NIHR Biomedical Research funding scheme. The Dementia Research Centre is supported by Alzheimer's Research UK, Brain Research Trust, and The Wolfson Foundation. This work was supported by the NIHR Queen Square Dementia Biomedical Research Unit and by an Alzheimer's Research UK Senior Research Fellowship

(ART-SRF2010-3) and ESRC/NIHR (ES/L001810/1) and EPSRC (EP/M006093/1) grants to SC. TS was supported by Alzheimer's Research UK Fellowship (ARUK-RF2013-5). The authors would also like to thank patients and controls for taking part in the study and Sam Hutton for his helpful suggestions in terms of data analysis.

## REFERENCES

- Alzheimer's Association. 2015 Alzheimer's disease facts and figures. *Alzheimer's Dement* (2015) 11(3):332. doi:10.1016/j.jalz.2015.02.003
- Dubois B, Feldman HH, Jacova C, Dekosky ST, Barberger-Gateau P, Cummings J, et al. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *Lancet Neurol* (2007) 6(8):734–46. doi:10.1016/S1474-4422(07)70178-3
- Dubois B, Feldman HH, Jacova C, Cummings JL, Dekosky ST, Barberger-Gateau P, et al. Revising the definition of Alzheimer's disease: a new lexicon. *Lancet Neurol* (2010) 9(11):1118–27. doi:10.1016/S1474-4422(10)70223-4
- Dubois B, Feldman HH, Jacova C, Hampel H, Molinuevo JL, Blennow K, et al. Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria. *Lancet Neurol* (2014) 13(6):614–29. doi:10.1016/S1474-4422(14)70090-0
- McKhann GM, Knopman DS, Chertkow H, Hyman BT, Jack CR, Kawas CH, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the national institute on aging-Alzheimer's association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement* (2011) 7(3):263–9. doi:10.1016/j.jalz.2011.03.005
- Doody RS, Thomas RG, Farlow M, Iwatsubo T, Vellas B, Joffe S, et al. Phase 3 trials of solanezumab for mild-to-moderate Alzheimer's disease. *N Engl J Med* (2014) 370(4):311–21. doi:10.1056/NEJMoa1312889
- Gauthier S, Feldman HH, Schneider LS, Wilcock GK, Frisoni GB, Hardlund JH, et al. Efficacy and safety of tau-aggregation inhibitor therapy in patients with mild or moderate Alzheimer's disease: a randomised, controlled, double-blind, parallel-arm, phase 3 trial. *Lancet* (2016) 388(10062):2873–84. doi:10.1016/S0140-6736(16)31275-2
- Weintraub S, Wicklund AH, Salmon DP. The neuropsychological profile of Alzheimer disease. *Cold Spring Harb Perspect Med* (2012) 2(4):a006171. doi:10.1101/cshperspect.a006171
- Hodges JR. Alzheimer's centennial legacy: origins, landmarks and the current status of knowledge concerning cognitive aspects. *Brain* (2006) 129(11):2811–22. doi:10.1093/brain/awl275
- Kirby E, Bandelow S, Hogervorst E. Visual impairment in Alzheimer's disease: a critical review. *J Alzheimer's Dis* (2010) 21(1):15–34. doi:10.3233/JAD-2010-080785
- Cronin-Golomb A, Corkin S, Rizzo JF, Cohen J, Growdon JH, Banks KS. Visual dysfunction in Alzheimer's disease: relation to normal aging. *Ann Neurol* (1991) 29(1):41–52. doi:10.1002/ana.410290110
- Mendola JD, Cronin-Golomb A, Corkin S, Growdon JH. Prevalence of visual deficits in Alzheimer's disease. *Optom Vis Sci* (1995) 72(3):155–67. doi:10.1097/00006324-199503000-00003
- Rizzo M, Anderson SW, Dawson J, Nawrot M. Vision and cognition in Alzheimer's disease. *Neuropsychologia* (2000) 38(8):1157–69. doi:10.1016/S0028-3932(00)00023-3
- Kurylo DD, Corkin S, Rizzo JF, Growdon JH. Greater relative impairment of object recognition than of visuospatial abilities in Alzheimer's disease. *Neuropsychology* (1996) 10:74–81. doi:10.1037/0894-4105.10.1.74
- Rizzo M, Nawrot M. Perception of movement and shape in Alzheimer's disease. *Brain* (1998) 121(12):2259–70. doi:10.1093/brain/121.12.2259
- Uhlmann RF, Larson EB, Koepsell TD, Rees TS, Duckert LG. Visual impairment and cognitive dysfunction in Alzheimer's disease. *J Gen Intern Med* (1991) 6(2):126–32. doi:10.1007/BF02598307
- Cronin-Golomb A, Gilmore GC, Neargarder S, Morrison SR, Laudate TM. Enhanced stimulus strength improves visual cognition in aging and Alzheimer's disease. *Cortex* (2007) 43(7):952–66. doi:10.1016/S0010-9452(08)70693-2
- Fujimori M, Imamura T, Yamashita H, Hirono N, Ikejiri Y, Shimomura T, et al. Age at onset and visuo-cognitive disturbances in Alzheimer disease. *Alzheimer Dis Assoc Disord* (1998) 12(3):163–6. doi:10.1097/00002093-199809000-00007
- Cronin-Golomb A, Sugiura R, Corkin S, Growdon JH. Incomplete achromatopsia in Alzheimer's disease. *Neurobiol Aging* (1993) 14(5):471–7. doi:10.1016/0197-4580(93)90105-K
- Benson D, Davis R, Snyder BD. Posterior cortical atrophy. *Arch Neurol* (1988) 45(7):789–93. doi:10.1001/archneur.1988.00520310107024
- Mendez MF, Ghajarania M, Perryman KM. Posterior cortical atrophy: clinical characteristics and differences compared to Alzheimer's disease. *Dement Geriatr Cogn Disord* (2002) 14(1):33–40. doi:10.1159/000058331
- Tang-Wai DF, Graff-Radford NR, Boeve BF, Dickson DW, Parisi JE, Crook R, et al. Clinical, genetic, and neuropathologic characteristics of posterior cortical atrophy. *Neurology* (2004) 63(7):1168–74. doi:10.1212/01.WNL.0000140289.18472.15
- Crutch SJ, Schott JM, Rabinovici GD, Murray M, Snowden JS, van der Flier WM, et al. Consensus classification of posterior cortical atrophy. *Alzheimer's Dement* (2017). doi:10.1016/j.jalz.2017.01.014
- Ridgway GR, Lehmann M, Barnes J, Rohrer JD, Warren JD, Crutch SJ, et al. Early-onset Alzheimer disease clinical variants. *Neurology* (2012) 79(1):80–4. doi:10.1212/WNL.0b013e31825dce28
- Renner JA, Burns JM, Hou CE, McKeel DW, Storandt M, Morris JC. Progressive posterior cortical dysfunction: a clinicopathologic series. *Neurology* (2004) 63(7):1175–80. doi:10.1212/01.WNL.0000140290.80962.BF
- Charles RF, Hillis AE. Posterior cortical atrophy: clinical presentation and cognitive deficits compared to Alzheimer's disease. *Behav Neurol* (2005) 16(1):15–23. doi:10.1155/2005/762569
- McMonagle P, Deering F, Berliner Y, Kertesz A. The cognitive profile of posterior cortical atrophy. *Neurology* (2006) 66(3):331–8. doi:10.1212/01.wnl.0000196477.78548.db
- Lehmann M, Barnes J, Ridgway GR, Wattam-Bell J, Warrington EK, Fox NC, et al. Basic visual function and cortical thickness patterns in posterior cortical atrophy. *Cereb Cortex* (2011) 21(9):2122–32. doi:10.1093/cercor/bhq287
- Shakespeare TJ, Kaski D, Yong KX, Paterson RW, Slattery CF, Ryan NS, et al. Abnormalities of fixation, saccade and pursuit in posterior cortical atrophy. *Brain* (2015) 138(7):1976–91. doi:10.1093/brain/awv103
- Michaelis JR, McConnell DS, Smither JA. Attention deficit/hyperactivity disorder's effects on individuals' scan paths during a simulated drive. *Work* (2012) 41(Suppl 1):5833–5. doi:10.3233/WOR-2012-0967-5833
- Sasson NJ, Elison JT. Eye tracking young children with autism. *J Vis Exp* (2012) 61:3675. doi:10.3791/3675
- Molitor RJ, Ko PC, Ally BA. Eye movements in Alzheimer's disease. *J Alzheimer's Dis* (2015) 44(1):1–12. doi:10.3233/JAD-141173
- Boxer AL, Garbutt S, Seeley WW, Jafari A, Heuer HW, Mirsky J, et al. Saccade abnormalities in autopsy-confirmed frontotemporal lobar degeneration and Alzheimer disease. *Arch Neurol* (2012) 69(4):509–17. doi:10.1001/archneurol.2011.1021
- Fletcher WA, Sharpe JA. Saccadic eye movement dysfunction in Alzheimer's disease. *Ann Neurol* (1986) 20(4):464–71. doi:10.1002/ana.410200405
- Bylsma FW, Rasmussen DX, Rebok GW, Keyl PM, Tune L, Brandt J. Changes in visual fixation and saccadic eye movements in Alzheimer's disease. *Int J Psychophysiol* (1995) 19(1):33–40. doi:10.1016/0167-8760(94)00060-R
- Yang Q, Wang T, Su N, Liu Y, Xiao S, Kapoula Z. Long latency and high variability in accuracy-speed of prosaccades in Alzheimer's disease at mild to moderate stage. *Dement Geriatr Cogn Dis Extra* (2011) 1(1):318–29. doi:10.1159/000333080
- Yang Q, Wang T, Su N, Xiao S, Kapoula Z. Specific saccade deficits in patients with Alzheimer's disease at mild to moderate stage and in patients with amnesic mild cognitive impairment. *Age (Dordr)* (2013) 35(4):1287–98. doi:10.1007/s11357-012-9420-z
- Garbutt S, Matlin A, Hellmuth J, Schenk AK, Johnson JK, Rosen H, et al. Oculomotor function in frontotemporal lobar degeneration, related disorders

- and Alzheimer's disease. *Brain* (2007) 131(5):1268–81. doi:10.1093/brain/awn047
39. Abel LA, Bowman EA, Velakoulis D, Fahey MC, Desmond P, Macfarlane MD, et al. Saccadic eye movement characteristics in adult Niemann-pick type C disease: relationships with disease severity and brain structural measures. *PLoS One* (2012) 7(11):e50947. doi:10.1371/journal.pone.0050947
  40. Crawford TJ, Higham S, Renvoize T, Patel J, Dale M, Suriya A, et al. Inhibitory control of saccadic eye movements and cognitive impairment in Alzheimer's disease. *Biol Psychiatry* (2005) 57(9):1052–60. doi:10.1016/j.biopsych.2005.01.017
  41. Shafiq-Antonacci R, Maruff P, Masters C, Currie J. Spectrum of saccade system function in Alzheimer disease. *Arch Neurol* (2003) 60(9):1272–8. doi:10.1001/archneur.60.9.1272
  42. Pereira MLF, Camargo MV, Aprahamian I, Forlenza OV. Eye movement analysis and cognitive processing: detecting indicators of conversion to Alzheimer's disease. *Neuropsychiatr Dis Treat* (2014) 10:1273–85. doi:10.2147/NDT.S55371
  43. Forssman L, Ashorn P, Ashorn U, Maleta K, Matchado A, Kortekangas E, et al. Eye-tracking-based assessment of cognitive function in low-resource settings. *Arch Dis Child* (2016) 102(4):301–2. doi:10.1136/archdischild-2016-310525
  44. Crutcher MD, Calhoun-Haney R, Manzanares CM, Lah JJ, Levey AI, Zola SM. Eye tracking during a visual paired comparison task as a predictor of early dementia. *Am J Alzheimers Dis Other Demen* (2009) 24(3):258–66. doi:10.1177/1533317509332093
  45. Richmond J, Sowerby P, Colombo M, Hayne H. The effect of familiarization time, retention interval, and context change on adults' performance in the visual paired-comparison task. *Dev Psychobiol* (2004) 44(2):146–55. doi:10.1002/dev.10161
  46. Fernández G, Manes F, Rotstein NP, Colombo O, Mandolesi P, Politi LE, et al. Lack of contextual-word predictability during reading in patients with mild Alzheimer disease. *Neuropsychologia* (2014) 62:143–51. doi:10.1016/j.neuropsychologia.2014.07.023
  47. Costa L, Gago MF, Yelshyna D, Ferreira J, David Silva H, Rocha L, et al. Application of machine learning in postural control kinematics for the diagnosis of Alzheimer's disease. *Comput Intell Neurosci* (2016) 2016:15. doi:10.1155/2016/3891253
  48. Warrington EK, James M. *Visual Object and Space Perception Battery (VOSP)*. Pearson Assessment (2017). Available from: [http://www.pearson-clinical.co.uk/Psychology/AdultCognitionNeuropsychologyandLanguage/AdultPerceptionandVisuomotorAbilities/VisualObjectandSpacePerception-Battery\(VOSP\)/VisualObjectandSpacePerceptionBattery\(VOSP\).aspx](http://www.pearson-clinical.co.uk/Psychology/AdultCognitionNeuropsychologyandLanguage/AdultPerceptionandVisuomotorAbilities/VisualObjectandSpacePerception-Battery(VOSP)/VisualObjectandSpacePerceptionBattery(VOSP).aspx)
  49. Willison JR, Warrington EK. Cognitive retardation in a patient with preservation of psychomotor speed. *Behav Neurol* (1992) 5(2):113–6. doi:10.3233/BEN-1992-5207
  50. Nelson HE. *National Adult Reading Test*. Windsor: NFER-Nelson (1982).
  51. Crossland MD, Rubin GS. The use of an infrared eyetracker to measure fixation stability. *Optom Vis Sci* (2002) 79(11):735–9. doi:10.1097/00006324-200211000-00011
  52. Zee DS, Leigh RJ. *The Neurology of Eye Movements*. Oxford: Oxford University Press (2006).
  53. Baum LE, Petrie T. Statistical inference for probabilistic functions of finite state Markov chains. *Ann Math Stat* (1966) 37(6):1554–63. doi:10.1214/aoms/1177699147
  54. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. *J R Stat Soc Ser B Methodol* (1977) 39:1–38.
  55. Cronin-Golomb A, Corkin S, Growdon JH. Visual dysfunction predicts cognitive deficits in Alzheimer's disease. *Optom Vis Sci* (1995) 72(3):168–76. doi:10.1097/00006324-199503000-00004
  56. Zola SM, Manzanares CM, Clopton P, Lah JJ, Levey AI. A behavioral task predicts conversion to mild cognitive impairment and Alzheimer's disease. *Am J Alzheimers Dis Other Demen* (2013) 28(2):179–84. doi:10.1177/1533317512470484
  57. Beach TG, McGeer EG. Cholinergic fiber loss occurs in the absence of synaptophysin depletion in Alzheimer's disease primary visual cortex. *Neurosci Lett* (1992) 142(2):253–6. doi:10.1016/0304-3940(92)90385-K
  58. Buck BH, Black SE, Behrmann M, Caldwell C, Bronskill MJ. Spatial- and object-based attentional deficits in Alzheimer's disease. Relationship to HMPAO-SPECT measures of parietal perfusion. *Brain* (1997) 120(7):1229–44. doi:10.1093/brain/120.7.1229
  59. Mielke R, Kessler J, Fink G, Herholz K, Heiss W-D. Dysfunction of visual cortex contributes to disturbed processing of visual information in Alzheimer's disease. *Int J Neurosci* (1995) 82(1–2):1–9. doi:10.3109/00207459508994285
  60. Kuljis RO. Lesions in the pulvinar in patients with Alzheimer's disease. *J Neuropathol Exp Neurol* (1994) 53(2):202–11. doi:10.1097/00005072-199403000-00012
  61. Ogren MP, Mateer CA, Wyler AR. Alterations in visually related eye movements following left pulvinar damage in man. *Neuropsychologia* (1984) 22(2):187–96. doi:10.1016/0028-3932(84)90061-7
  62. Robinson DL, Petersen SE. Responses of pulvinar neurons to real and self-induced stimulus movement. *Brain Res* (1985) 338(2):392–4. doi:10.1016/0006-8993(85)90176-3
  63. Boxer AL, Garbutt S, Rankin KP, Hellmuth J, Neuhaus J, Miller BL, et al. Medial versus lateral lobe contributions to voluntary saccade control as revealed by the study of patients with frontal lobe degeneration. *J Neurosci* (2006) 26(23):6354–63. doi:10.1523/JNEUROSCI.0549-06.2006
  64. Crutch SJ, Lehmann M, Schott JM, Rabinovici GD, Rossor MN, Fox NC. Posterior cortical atrophy. *Lancet Neurol* (2012) 11(2):170–8. doi:10.1016/S1474-4422(11)70289-7
  65. Primativo S, Yong KX, Shakespeare TJ, Crutch SJ. The oral spelling profile of posterior cortical atrophy and the nature of the graphemic representation. *Neuropsychologia* (2017) 94:61–74. doi:10.1016/j.neuropsychologia.2016.11.021
  66. Migliaccio R, Agosta F, Rascovsky K, Karydas A, Bonasera S, Rabinovici GD, et al. Clinical syndromes associated with posterior atrophy early age at onset AD spectrum. *Neurology* (2009) 73(19):1571–8. doi:10.1212/WNL.0b013e3181c0d427
  67. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res* (1975) 12(3):129–138.
  68. Warrington E, James M, Plant GT. *CORVIST: Cortical Vision Screening Test: Manual & Test Materials*. Thames Valley Test Company. (2001).
  69. Wechsler D. *Wechsler Abbreviated Scale of Intelligence (WASI) Manual*. The Psychological Corporation (1999).
  70. Wechsler D. A standardized memory scale for clinical use. *J Psychol* (1945) 19(1):87–95.
  71. Warrington EK. *The Camden Memory Tests*. Psychology Press (1996). 140 p.
  72. Jackson M, Warrington EK. Arithmetic skills in patients with unilateral cerebral lesions. *Cortex* (1986) 22(4):611–20.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2017 Pavisić, Firth, Parsons, Rego, Shakespeare, Yong, Slattery, Paterson, Foulkes, Macpherson, Carton, Alexander, Shawe-Taylor, Fox, Schott, Crutch and Primativo. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.