

A comparative analysis of analytic approaches for predicting performance on a putative behavioural marker for Alzheimer's disease

Joe Butler^{1,2} | Dennis Boateng³ | Mario A A Parra Sr.⁴ | Tamlyn J Watermeyer^{5,6,7} | Samuel O. Danso^{8,9} | Binding in Neuropsychiatric Disorders (B.I.N.D) Group

¹School of Psychology, University of Sunderland, Sunderland, United Kingdom

²National Institute of Health Applied Research Collaboration North East & Cumbria, Newcastle-Upon-Tyne, England, United Kingdom

³Global Statistical Consult, Accra, Greater Accra Region, Ghana

⁴University of Strathclyde, Glasgow, Glasgow, United Kingdom

⁵National Institute of Health Applied Research Collaboration North East & Cumbria, Newcastle-upon-Tyne, England, United Kingdom

⁶Edinburgh Dementia Prevention, Edinburgh, United Kingdom

⁷University of Northumbria, Newcastle upon tyne, England, United Kingdom

⁸School of Computer Science and Engineering, University of Sunderland, Sunderland, England, United Kingdom

⁹University of Edinburgh, Edinburgh, United Kingdom

Correspondence

Joe Butler, School of Psychology, University of Sunderland, Sunderland, United Kingdom.

Email: joe.butler@sunderland.ac.uk

Abstract

Background: Remote neurocognitive assessment tools, such as the Visual Short-Term Memory Binding Task (VSTMBT), offer promising opportunities for assessing Alzheimer's Disease (AD) risk and identifying new population-level risk factors. State-of-the-art machine learning methods can enhance these tools, providing insights into mechanisms driving risk. This study evaluates three models to examine the impact of self-reported variables on VSTMBT performance

Method: We categorised participants as strong-binders (SB – indicative of no pathology; 85.9% of sample) or weak-binders (WB – indicative of pathology; 14.1%) based on binding-cost (Parra et al., 2024). Bivariate and multivariate regression analyses were conducted to identify variables with 5% significance and 95% confidence intervals. Continuous variables were summarised using means and standard deviations, while categorical variables were presented as proportions with *p*-values. Machine Learning models (SVM, RF, DT) were also developed as an alternative analytical approach for validation. Models were trained with an 80/20% split between training and testing.

Result: Both groups shared similar demographics (mean BMI = 24.1kg/m²; WB mean age = 28.1, range = 24.8 - 31; SB mean age = 28.6, range = 27.3 - 30.0). Regression models showed significant positive associations of PDQ5 score with higher odds of WB effect (AOR = 1.143; 95% CI: 1.012 - 1.291; *p* = 0.031). Significant positive associations with higher odds of WB effect were observed for multi-morbidity (AOR = 1.863; 95% CI: 1.082 - 3.208; *p* = 0.025); MHC (AOR = 1.03; 95% CI: 1.082 - 3.208; *p* = 0.049) and SART (AOR = 1.863; 95% CI: 1.000 - 1.020; *p* = 0.075). Consistent sleep quality was marginally associated with lower odds of WB. There were slightly significant lower odds (AOR=0.50, 95% CI: 0.247 - 1.011; *p* = 0.054) See Table 1. Further attribute

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Alzheimer's Association. *Alzheimer's & Dementia* published by Wiley Periodicals LLC on behalf of Alzheimer's Association.

importance analysis revealed overlapping risk factors identified by our regression models to have high significant odds ratio also being ranked among the top five attributes by the best RF model (Accuracy=81.0, Precision = 86.0; Recall = 93.0; F1 score=90.0; AUROC = 60.0) in Table 3.

Conclusion: Our analysis highlights the value of integrating machine learning with regression approaches to identify VSTMBT predictors and their role in risk stratification. Machine learning complements regression in population-based AD assessments. Future research should apply these models for individual prediction and early detection.

Table 1: Association of individual characteristics with Weak Binding

Variables	Bivariate analysis			Multivariate analysis		
	OR	95% CI	P-value	AOR	95% CI	P-value
Corsi level	0.960	(0.829 - 1.126)	0.602	0.989	(0.838 - 1.166)	0.891
PDQ5	1.044	(0.970 - 1.127)	0.262	1.143	(1.012 - 1.291)	0.031**
MFIS	1.005	(0.924 - 1.073)	0.875	0.924	(0.829 - 1.031)	0.156
Multimorbidity	1.676	(1.029 - 2.673)	0.0313**	1.863	(1.082 - 3.208)	0.025**
MHC	1.017	(0.992 - 1.042)	0.191	1.03	(1.000 - 1.061)	0.049**
BMI (kg/m ²)	0.966	(0.910 - 1.014)	0.221	0.95	(0.889 - 1.015)	0.130
SART	1.009	(0.999 - 1.019)	0.058*	1.009	(0.999 - 1.02)	0.075*
Sleep Quality						
0				1		
1	0.577	(0.302 - 1.109)	0.095	0.5	(0.247 - 1.011)	0.054*
2	0.573	(0.157- 1.664)	0.344	0.397	(0.115 - 1.37)	0.144
3	2.150	(0.097- 23.535)	0.540	2.138	(0.166 - 27.522)	0.56
Smoking Status						
No				1		
Yes	2.914	(0.577 - 53.116)	0.304	2.711	(0.337 - 21.795)	0.348
Gender						
Male				1		
Female	1.050	(0.514 - 2.328)	0.898	0.842	(0.375 - 1.891)	0.678
Covid Status						
No				1		
Yes	1.123	(0.558 - 2.158)	0.735	1.142	(0.563 - 2.318)	0.713
Education Level						
1				1		
2	1.000	(0.307- 4.506)	1.000	1.286	(0.287 - 5.749)	0.742
3	1.634	(0.262 - 10.274)	0.585	2.511	(0.362 - 17.405)	0.351
4	0.923	(0.279 - 4.189)	0.905	1.146	(0.263 - 4.994)	0.856

Odd Ratio (OR); Adjusted Odd Ratio (AOR); Confidence Interval (CI)

*** $p < .01$, ** $p < .05$, * $p < .1$

Table 2: Confusion Matrix

Confusion matrix		Logistic		SVM		Decision Trees		Random Forest	
		Predicted		Predicted		Predicted		Predicted	
		WB	SB	WB	SB	WB	SB	WB	SB
Observed	WB	7	3	5	5	1	9	1	9
	SB	23	37	16	44	10	50	4	56
		%		%		%		%	
Accuracy		63.0		70.0		73.0		81.0	
Precision		93.0		90.0		85.0		86.0	
Recall		62.0		73.0		83.0		93.0	
F1 score		74.0		81.0		84.0		90.0	
AUROC		66.5		63.0		46.7		60.0	

Figure 1: Feature importance

