Agreement Between Amyloid Status Prediction from Visual Read, SUVR Thresholding, and AI-Based Interpretation of Amyloid PET

Jane Waugh[1], Luis Peraza[1], Richard Joules[1], Richard Mohs[2], Douglas Beauregard[2], John Dwyer[2], Lynne Hughes[1,2], Robin Wolz[1]

1: IXICO, London, UK

2: Global Alzheimer's Platform Foundation, Washington, DC, USA.

Amyloid-β positivity is a core biomarker of Alzheimer's disease. Visual read of amyloid PET by expert readers remains the clinical gold standard, but interpretation can be subjective. Quantitative approaches such SUVRs and centiloid banding can support visual assessment. Recent AI methods aim to capture complex spatial uptake patterns directly from imaging data and may offer an alternative classification.

Despite these advances, discrepancies between visual reads, SUVR thresholding, and AI-based classifications persist, underscoring the need to better understand agreement and complementarity among these methods.

We analysed amyloid PET ([^{18}F]florbetapir) imaging data from n=827 participants enrolled in the Global Alzheimer's Platform (GAP) Bio-Hermes trial [1].

Population	Female	Mean age	Visual Aβ+
827	476	71.7	287

PET image amyloid status were classified with a range of methods, as described below, each method resulted in a label for positive or negative amyloid status.

Method	Label	
	Amyloid Positive	Amyloid Negative
Visual Read	V+	V-
GCA SUVR Thresholding	T+	T-
Embedding AI Prediction	EP+	EP-

PET imaging data was processed with the MIM software [2] to extract the global cortical average **(GCA) SUVR**. Quantative assignments of amyloid status were assigned through thresholding the MIM SUVR output values at 1.12. Images were additionally processed with a PET only research pipeline to extract regional SUVR values based on the AAL3 parcellation schema.

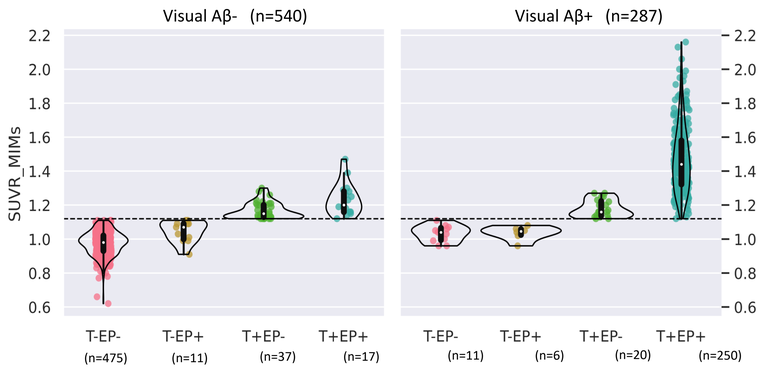
Automatic assessment of amyloid status was performed with an inhouse deep learning workflow, referred to as **Embedding AI Prediction**, utilising PET images directly to predict amyloid status, as defined based on visual read using direct input of PET images

This workflow employs siamese (dual encoder) network with a two-stage encoder pipeline: the first stage pretrained and frozen 3D convolutional autoencoder providing stable, low-level anatomical representations learned via unsupervised reconstruction. The second stage is a trainable autoencoder operating on the latent representation of the first stage, compressing features into a compact embedding, projected through a fully connected layer to produce a dimensionally reduced embedding. The network is trained using a contrastive loss based on the formulation of Hadsell et al. (2006) the distance between embeddings is minimized for pairs from the same class (visual amyloid status) and constrained to exceed a margin for pairs from different classes. Image embeddings were used as a feature set for a Catboost [3] classifier to predict image amyloid. All training was conducted in an 8-fold cross validated framework (test/train = 12.5% / 87.5%) to reduce bias.

Both AI image prediction and SUVR thresholding reported similar accuracy in prediction of visually defined amyloid state.

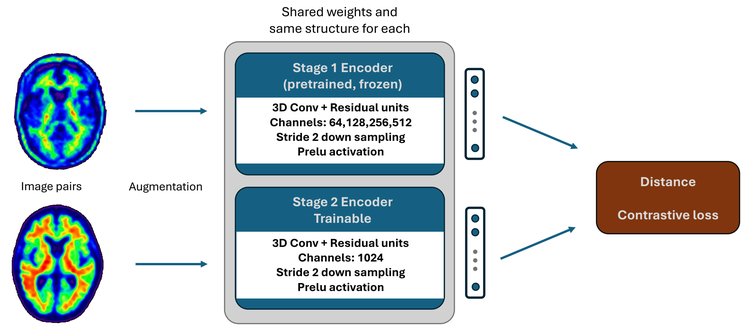
N = 827	MIM SUVR	Embedding AI
Balanced Accuracy	0.899	0.922
precision	0.941	0.892
recall	0.833	0.901
specificity	0.966	0.943
f1_score	0.883	0.897
False negative rate (nFN)	0.167 (54)	0.098 (28)
False positive rate (nFP)	0.034 (17)	0.057 (31)

Discrepancies between visual read outcome and algorithmic methods occur in a band around the SUVR threshold.

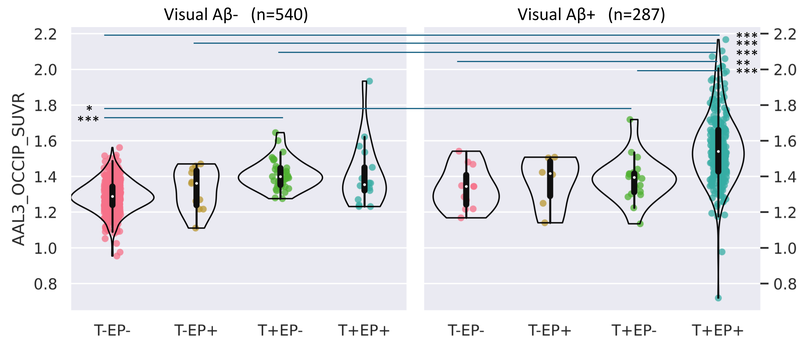


The visual read outcome differed from the SUVR threshold or AI approach predictions for 102 cases (~12%) with algorithmic approaches agreeing in 37 of those 102 cases (~36%).

Differences in the GCA SUVR outcome to the visual read may be driven by multiple factors such as SUVR computation (ROI placement and definition), WM/GM contrast and focal uptake patterns which may impact visual read and SUVR differentially.



Occipital regions, excluded from the GCA composite but visible during visual read and AI assessment, may contribute to the discordance in amyloid status prediction. Generally Occipital SUVR increases monotonically with multi-method agreement on amyloid positivity. Significant effects were observed broadly between V+T+EP+ and most V+ discordant groups as we well as between V-T-EP- and V-T+EP- (g=1.3, p<<.001).

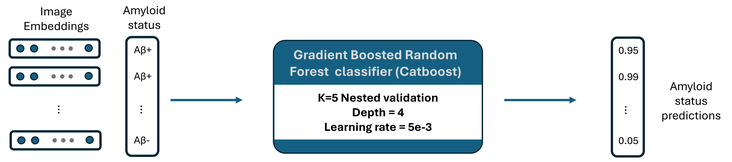


Cortical SUVR values were z-scored, and a region of focal uptake defined as an ROI with z > 1.5 to capture cortical ROIs with local extreme of uptake relative to the cortical baseline. When considering both counts and mean SUVR of focal uptake areas, no significant differences were observed. When considering focal ROI counts or intensity alone, multiple significant differences were observed for increased focal ROIs in V+T-EP+ compared to other discordant groups. Further assessment defining focal uptake relative to WM may reveal further drivers of discordance.

This preliminary analysis indicates strong agreement between visual read, global cortical SUVR thresholding, and AI-based amyloid PET interpretation. However, systematic discrepancies emerge in cases clustered around the SUVR positivity threshold, highlighting that these methods assess amyloid burden through partially distinct and complementary mechanisms.

In routine clinical practice, global SUVR thresholding provides a reproducible and scalable quantitative anchor, while visual read remains essential for contextual interpretation of regional uptake patterns. AI-based assessment shows promise as an additional, consistent reader that may highlight spatial features not reflected in global metrics.

For cases with clear agreement across methods, a single-read paradigm supported by quantitative SUVR may be sufficient. However, in borderline or discordant cases, particularly those falling within a narrow SUVR band around the positivity threshold, the results support the value of dual visual reads and/or AI-assisted review to increase diagnostic confidence.



[1] Mohs et al; Alzheimers Dement. 2024 Apr;20(4):2752-2765. doi: 10.1002/alz.13722
[2] https://www.mimsoftware.com/nuclear-medicine/mim-encore
[3] https://catboost.ai/