Exploratory Assessment of K-means Clustering to Classify [18F]Flutemetamol Brain PET as Positive or Negative

Abstract

Rationale: We evaluated K-means clustering to classify amyloid brain PETs as positive or negative.

Methods: Sixty-six participants (31 men, 35 women; age range 52-81 years) were recruited through a multicenter observational study: 19 cognitively normal (CN), 25 mild cognitive impairment (MCI), and 22 dementia (11 Alzheimer's disease, 3 subcortical vascular cognitive impairment and 8 Parkinson-Lewy Body spectrum disorder). As part of the neurocognitive and imaging evaluation each participant had a [18F]Flutemetamol (Vizamyl^{TM}, GE Healthcare) brain PET. All studies were processed using Cortex ID software (General Electric Company, Boston Massachusetts) to calculate standardized uptake value ratios (SUVr) in 19 regions of interest (ROIs), and clinically interpreted by two dual certified radiologists/ nuclear medicine physicians, using MIM software (MIM Software Inc., Cleveland, Ohio), blinded to the quantitative analysis, with final interpretation based on consensus. K-means clustering was retrospectively used to classify the studies from the quantitative data.

Results: Based on clinical interpretation, 46 brain PETs were negative and 20 were positive for amyloid deposition. Of 19 CN participants, 1 (5%) had a positive [18F]Flutemetamol brain PET. Of 25 participants with MCI, 9 (36%) had a positive [18F]Flutemetamol brain PET. Of 22 participants with dementia, 10 (45%) had a positive [18F]Flutemetamol brain PET; 7 of 11 participants with AD (64%), 1 of 3 participants with VCI (33%) and 2 of 8 participants with PD/DLB (25%) had a positive [18F]Flutemetamol brain PET. Using clinical interpretation as the gold standard, K-means clustering (K=2), gave sensitivity=95%, specificity=98%, and accuracy=97%.

Conclusion: K-means clustering may be a powerful algorithm for classifying amyloid brain PET.

Keywords: Machine learning, unsupervised learning, K-means clustering, brain imaging, dementia

Introduction:

It is estimated that over 50 million people have dementia worldwide, with nearly 10 million new cases diagnosed annually (1). Alzheimer's disease (AD), vascular cognitive impairment (VCI) and Parkinson-Lewy Body spectrum disorder (PD/LBD) are among the most common types of dementia and current diagnosis is based on clinical evaluation, with imaging playing a key supportive role (2). The pathology of AD is characterized by amyloid plaques and neurofibrillary tangles; however, amyloid deposition can also be seen in individuals with VCI and PD/DLB, possibly in part related to the high occurrence of mixed pathology. Although not routinely done, positron emission tomography (PET) can help assess the presence and extent of amyloid deposition in the brain (3). Clinically, PET scans are visually interpreted in a binary fashion as positive or negative for amyloid deposition. However, evaluation of the extent of amyloid deposition may be further augmented with quantitative imaging analysis in the form of standardized uptake values (SUVs) within specific regions of interest (ROIs) in the brain normalized to a reference region of the brain (SUVrs).

Machine learning (ML) has been used in medical applications for decades (4). However, its use has become ubiquitous in the past few years due, in part, to advances in hardware, software and access to training datasets. Typically, the ML algorithm chosen depends on the task at hand and the available data for training. K-means clustering is an unsupervised ML algorithm (5-7) that assigns data points from the dataset to a cluster based on commonalities. Being unsupervised, the algorithm does not require tagged data for training (e.g. datasets that have been clinically interpreted by a physician). Rather, in K-means clustering numerical values for a set of pre-determined features are used to assign each data point in the training dataset to one of K clusters through an iterative process, where K is chosen by the programmer.

The aim of this paper is to explore the feasibility of K-means clustering for the classification of amyloid brain PET scans as positive or negative for amyloid deposition. To our knowledge, this is the first time this ML algorithm has been used in this context.

Materials and Methods:

The first 66 participants (31 men, 35 women, age 52-81 years) recruited through a prospective multicenter observational study (8) were included in the analysis: 19 cognitively normal (CN), 25 with mild cognitive impairment (MCI), and 22 with dementia. Of the 22 participants with dementia: 11 had AD, 3 had subcortical VCI and 8 had PD/DLB spectrum disorder, based on clinical assessment (9-11). As part of the neurocognitive evaluation each participant had a [18F]Flutemetamol PET scan of their brain using scanners available at two of the participating sites (GE Discovery MI and Philips Gemini Big Bore PET/CT scanners). Images were acquired for 20 minutes approximately 90 minutes following intravenous administration of 185MBq of [^{18}F]flutemetamol; consistency of data was maintained by adherence to a

standard quality assurance program and use of a common imaging protocol. All brain PET studies were clinically interpreted as positive or negative for amyloid deposition by two dual certified radiologists/ nuclear medicine physicians, using MIM software (MIM Software Inc., Cleveland, Ohio), blinded to the quantitative analysis, with final interpretation based on consensus. Clinical interpretation was performed using a subjective visual approach according to the manufacturer's reader training program specific to the radiopharmaceutical. Table 1 provides a summary of the study participants. In addition, all brain PET scans were retrospectively processed using Cortex ID software (General Electric Company, Boston Massachusetts) to calculate the SUVr in each of 19 ROIs; subsequently converted to Z-scores using the GE database (12). ROIs included the: left and right prefrontal cortices, left and right anterior cingulate gyruses, left and right precuneus and posterior cingulate gyrus, left and right parietal lobes, left and right lateral temporal lobes, left and right occipital lobes, left and right sensorimotor cortices, left and right mesial temporal lobes, cerebellar gray matter, pons, as well as a composite ROI (13).

A K-means clustering ML algorithm programmed in Matlab was used to classify amyloid brain PET scans for the purpose of comparison with clinical interpretation (Figures 1 and 2). The K-means clustering algorithm used K=2 (i.e. assumption of 2 clusters, positive and negative) and the Z-score data (M=19 features) to determine if the data clusters were concordant with the clinical interpretation. The "exact" Clopper-Pearson 95% confidence interval (CI) was calculated with Medcalc. Since K-means clustering does not provide insight into whether a cluster is negative or positive, the clusters were assigned as 'negative' or 'positive' in a manner that maximized classification accuracy using the clinical amyloid PET scan interpretation as the ground truth.

Results:

Based on clinical interpretation, 46 of the [18F]Flutemetamol brain PET studies were negative and 20 were positive. Of the 19 CN participants, 1 (5%) had a positive [18F]Flutemetamol brain PET study. This participant was one of the oldest subjects included in the study at 80 years of age, with clinically derived normal Mini-Mental Status Examination (score 30) but borderline Montreal Cognitive Assessment (score 24). Of the 25 participants with MCI, 9 (36%) had a positive [18F]Flutemetamol brain PET study. Of the 22 participants with dementia, 10 (45%) had a positive [18F]Flutemetamol brain PET study; 7 of the 11 participants with AD (64%), 1 of the 3 participants with VCI (33%) and 2 of the 8 participants with PD/DLB (25%) had a positive [18F]Flutemetamol brain PET study (Table 1).

K-means clustering (K=2) of the data gave: sensitivity 95% (95% CI: 75% to 100%), specificity 98% (95% CI: 89% to 100%), and classification accuracy=97% (95% CI: 90% to 100%). Table 2 shows the confusion matrix. Table 3 shows the mean Z-score of the 19 ROIs for each cluster. The cluster that was interpreted as negative (A) had several ROIs with low amyloid deposition (low SUVrs), while the

cluster interpreted as positive (B) had several ROIs with high amyloid deposition (high SUVrs). The ROIs with the greatest difference in amyloid deposition were the precuneus/ posterior cingulate gyrus, the prefrontal and parietal lobes, as well as the lateral temporal lobe on both the left and right sides. Table 4 summarizes the algorithm parameters according to the reporting guidelines for ML algorithms first suggested earlier this year (14).

Discussion:

Clinically, the diagnosis of dementia is based on cognitive impairment that interferes with the activities of daily living (15). Pathologically, the etiology of dementia is a multi-factorial process and includes an array of pure and mixed pathology at presentation. AD is thought to result primarily from a complex interplay of Aβ protein deposition and aggregation of hyperphosphorylated tau protein (16). VCI is likely related to vascular factors that may also include deposition of beta-amyloid within vessels. For example, dementia occurring after a stroke, is estimated to have a 7% incidence at 1 year with 30% prevalence in stroke survivors (17). While there are several variables associated with post-stroke dementia, vascular lesions are thought to play a part. PD/DLB spectrum disorder is characterized by the aggregation of mis-folded α-synuclein protein in neuronal cell bodies (Lewy bodies) and cell processes (Lewy neurites). Often more than one dementia subtype co-exist (18,19). Indeed, evidence suggests Aβ, tau protein and α-synuclein interact promoting dementia pathology (20). Ultimately, it is thought that while amyloid deposition may be seen in cognitively normal individuals, it is more common in the setting of MCI and dementia. Further, the risk for conversion from MCI to AD is higher when cerebral amyloid deposition increases more rapidly (21,22). Also, while amyloid deposition is much more common in certain pathologies such as AD, it may be seen across the dementia spectrum including VCI and PD/DLB spectrum disorder (23). The diagnosis of dementia is made clinically; however, imaging plays a key role. For example, if amyloid deposition in the brain is detected on imaging it supports the clinical findings might be due to amyloid pathology.

Today, three PET radiopharmaceuticals are clinically approved by the U.S. Food and Drug Administration to image brain amyloid deposition: [18F]Florbetapir (Amyvid^{TM}), [18F]Flutemetamol (Vizamyl^{TM}) and [18F]Florbetaben (NeuraCeq^{TM}). When amyloid is present in the cerebral cortex, the radiotracer binds to it. In clinical practice, a amyloid PET scan is typically interpreted in a subjective and binary manner as positive or negative, where a positive scan correlates with moderate to frequent amyloid plaque, while a negative amyloid PET suggests no or sparse amyloid plaque. In research, a quantitative approach is often used that includes region specific amyloid deposition. Since cerebellar gray matter tends not to accumulate amyloid, this is often used as an internal control or reference region for quantitation. The ratio of radiotracer uptake in the brain ROI to that of the reference region is the SUVr. Quantitative

PET data may then be compared with a normal database such that Z-scores for individual and combinations of brain ROIs can be derived.

The IDEAS study was launched in 2016, in part to assess if imaging amyloid with PET could help clinicians diagnose the cause of cognitive impairment. Of 11,409 participants with MCI or dementia of uncertain etiology, a positive amyloid PET was reported in 55% with MCI and 70% with dementia. While individuals with AD often had a positive amyloid PET, individuals with other types of dementia more commonly had a negative amyloid PET (24,25). A study by Sevigny et al. reported amyloid PET as positive in 61% of 278 participants with clinical criteria for prodromal or mild AD (26). We found 1/19 (5%) of CN, 9/25 (36%) with MCI and 10/22 (45%) of participants with dementia had a positive amyloid PET. This is consistent with the literature in that most CN individuals have negative amyloid PET, while the incidence of positive amyloid PET increases in individuals with MCI and dementia. Further, and also consistent with the literature, we found that while the majority 7/11 (64%) of participants with clinically suspected AD had a positive amyloid PET, participants with clinically suspected non-AD dementia more commonly had negative amyloid PET. Our results also confirm the cerebral cortical regions in which amyloid deposition is preferentially found, specifically the precuneus and posterior cingulate gyrus, prefrontal lobes, parietal lobes, and lateral temporal lobes (27,28).

Recently, there has been an explosion in the use of ML algorithms for the analysis of imaging data. One of the goals of ML algorithm is to assist clinicians in suggesting imaging interpretation. In certain cases, tagged data or data that has been processed by a human being may not be available to train a ML algorithm. K-means clustering is an unsupervised ML algorithm that attempts to identify commonalities between data points and assigns nearby data points to a common representative "cluster." The idea is to iteratively assign each data point to a cluster (*K* the number of clusters) such that data points that are most alike are grouped together (assigned to the same cluster). There are a few papers that use a K-means algorithm to evaluate brain pathology. For example, Abualhaj et al. (29) and Blanc-Durand et al. (30) used a K-means algorithm with dynamic [^{18}F]FET brain PET scans in participants with brain tumors. Matias-Guiu JA et al. clustered participants with primary progressive aphasia (PPA) into different subtypes according to regional brain metabolism using [^{18}F]FDG brain PET scans (31). Escudero et al. used a K-means algorithm to help divide participants according to a bio-profile and postulated this might help identify those individuals with MCI most likely to convert to AD (32). To our knowledge, our results provide the first data suggesting K-means clustering may be a helpful tool to assist clinicians with the classification of [18F]Flutemetamol brain PET studies as positive or negative for amyloid deposition. Given the low computational complexity and high classification accuracy (97%), K-means clustering might be a desirable choice for the task of assisting clinical interpretation of amyloid brain PET. Further, this high fidelity is achieved even in the presence of a pathologically heterogeneous

sample with participants at all stages of disease. This implies that a ML-driven K-means algorithm may be helpful to suggest a PET interpretation that could potentially increase the confidence of the physician interpreting the amyloid brain PET. However, further investigation would be needed to confirm this.

Conclusions:

K-means clustering is a powerful unsupervised ML algorithm that may be a helpful tool for identifying amyloid brain PET positive individuals even in the face of a heterogeneous cohort, as would be encountered in a memory clinic.

References

[1] Dementia. World Health Organization. http://www.who.int/mediacentre/factsheets/fs362/en/. Updated May 14, 2019. Accessed 9 Jul 2019.

[2] Ismail Z, Black SE, Camicioli R, et al. Recommendations of the 5th Canadian Concensus Conference on the diagnosis and treatment of dementia. *Alzheimer's Dement*. 2020;16:1182-1195.

[3] Zukotynski K, Kuo PH, Mikulis D, et al. PET/CT of Dementia. *AJR Am J Roentgenol*. 2018;211(2):246-259.

[4] Silink K. The possibility of designing machines which learn diagnostic. The zero-systems of types and pathotypes in endocrinology. *Act Nerv Super (Praha)*. 1961;3:148-153.

[5] Steinhaus H. Sur la division des corps matériels en parties. *Bull. Acad. Polon. Sci*. 1957;4:801-804.

[6] MacQueen JB. Some methods for classification and analysis of multivariate observations. *Proc Berkeley Symp Mathematical Statistics and Probability*. 1967;281-297.

[7] Uribe CF, Mathotaarachchi S, Gaudet V, et al. Machine learning in nuclear medicine: Part 1-Introduction. *J Nucl Med*. 2019;60:451-458.

[8] ClinicalTrials.gov. BEAM: Brain-Eye Amyloid Memory Study (BEAM). http://clinicaltrials.gov/ct2/show/NCT02524405. Updated Oct 22, 2020. Accessed 25 Feb 2021.

[9] Petersen RC. Mild cognitive impairment as a diagnostic entity. *Journal of Internal Medicine.* 2004;256:183-194.

[10] Albert MS, DeKosky ST, Dickson D, et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations form the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):270-279.

[11] 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(3):263-269.

[12] Thurfjell L, Lilja J, Lundqvist R, et al. Automated Quantification of 18F-Flutemetamol PET Activity for Gategorizing Scans as Negative or Positive for Brain Amyloid: Concordance with Visual Image Reads. *J Nucl Med*. 2014;55(10):1-6.

[13] Jack CR Jr., Lowe VJ, Senjem ML, et al. ^{11}C PiB and structural MRI provide complementary information in imaging of Alzheimer's disease and amnestic mild cognitive impairment. *Brain*. 2008;131:665-680.

[14] Zukotynski K, Gaudet V, Uribe C, et al. Machine Learning in Nuclear Medicine: Part 2-Neural Networks and Clinical Aspects. *J Nucl Med*. 2021;62(1):22-29.

[15] American Psychiatric Association. *Diagnostic and statistical manual of mental disorders, fifth edition (DSM-5)*. Washington, DC: American Psychiatric Press; 2013.

[16] Braak H, Thal DR, Ghebremedhin E,Tredici KD. Stages of the pathologic process in Alzheimer Disease: age categories from 1 to 100 years. *J Neuropathol Exp Neurol*. 2011;70(11):960-969.

[17] Leys D, Henon H, Mackowiak-Cordoliani MA, Pasquier F. Poststroke dementia. *Lancet Neurol.* 2005;4:752-759.

[18] Mrak RE, Griffin WST. Dementia with Lewy bodies: Definition, diagnosis, and pathogenic relationship to Alzheimer's disease. *Neuropsychiatr Dis Treat.* 2007;3(5):619-625.

[19] Boller F, Mizutani T, Roessmann U, Gembetti P. Parkinson disease, dementia, and Alzheimer disease: Clinicopathological correlations. *Ann Neurology.* 1980;7(4):330-335.

[20] Mandal PK, Pettegre JW, Maliah E, et al. Interaction between Aβ peptide and α synuclein: molecular mechanisms in overlapping pathology of Alzheimer's and Parkinson's in dementia with Lewy body disease. *Neurochem Res.* 2006;31:1153-1162.

[21] Hatashita S, Wakebe D: Amyloid-β Deposition and Long-Term Progression in Mild Cognitive Impairment due to Alzheimer's Disease Defined with Amyloid PET Imaging. *J Alzheimers Dis.* 2017;57(3):765–773.

[22] Nordberg A, Carter SF, Rinne J, et al. A European multicentre PET study of fibrillar amyloid in Alzheimer's disease. *Eur J Nucl Med Mol Imaging.* 2013;40:104–114.

[23] Rowe CC, Ng S, Ackermann U, et al. Imaging beta-amyloid burden in aging and dementia. *Neurology.* 2007;68:1718–1725.

[24] Villemagne VL, Ong K, Mulligan RS, et al. Amyloid imaging with (18)F-florbetaben in Alzheimer disease and other dementias. *J Nucl Med.* 2011;52:1210-1217.

[25] Maetzler W, Liepelt I, Reimold M, et al. Cortical PIB binding in Lewy body disease is associated with Alzheimer-like characteristics. *Neurobiol Dis.* 2009;34:107-112.

[26] Sevigny J, Suhy J, Chiao P, et al. Amyloid PET screening for enrichment of early-stage Alzheimer disease clinical trials - Experience in a phase 1b clinical trial. *Alzheimer Dis Assoc Disord*. 2016;30(1):1-7.

[27] Braak H and Braak E. Neuropathological stageing of Alzheimer-related changes. *Acta Neuropathol*. 1991;82:239-259.

[28] Clark CM, Schneider JA, Bedell BJ, et al. Use of florbetapir-PET for imaging beta-amyloid pathology. *JAMA*. 2011;305:275–83.

[29] Abualhaj B, Weng G, Ong M, et al. Comparison of five cluster validity indices performance in brain [^{18}F]FET-PET image segmentation using k-means. *Med Phys*. 2017;44(1):209-220.

[30] Blanc-Durand P, Van Der Gucht A, Verger A, et al. Voxel-based 18F-FET PET segmentation and automatic clustering of tumor voxels: A significant association with IDH1 mutation status and survival in patients with gliomas. *PLoS ONE*. 2018;13(6):e0199379.

[31] Matias-Guiu JA, Díaz-Álvarez J, Ayala JL, et al. Clustering analysis of FDG-PET imaging in primary progressive aphasia. *Front Aging Neurosci*. 2018;10:230.

[32] Escudero J, Zajicek JP, Ifeachor E. Early detection and characterization of Alzheimer's disease in clinical scenarios using Bioprofile concepts and K-means. *Conf Proc IEEE Eng Med Biol Soc*. 2011;6470-6473.

Figure 1: K-means algorithm using an iterative process.

Figure 2: Illustration of K-means clustering using an example dataset with 20 training cases and n=2 features (x and y). A. Training cases plotted on the x-y plane. B. Initial random assignment of K=2 cluster centers (Cluster A: white diamond, and Cluster B: black diamond). C. Assignment of cases to nearest cluster center (white circles to Cluster A, and black circles to Cluster B; dashed line is at equal distance to the two centers). D. Cluster centers move to the center of the points. E. Three cases (arrows) are reassigned from Cluster B to Cluster A. F. Cluster centers move again to new centers of mass. G. Two cases (arrows) move from Cluster B to Cluster A. H. Cluster centers move again. Note that at this point, no cases change clusters so the algorithm has converged.

Input
Training set consisting of N data points (cases), each a vector of M features

Initialization
1. Randomly choose the centers of the K clusters by assigning a random value for each feature

Loop
2. Assign each data point in the training set to the cluster whose center is closest to it (lowest square distance calculated over all M features)
3. Adjust each cluster center to be at the center of the points (ie. average value for each feature) in the cluster
4. Repeat steps 2 and 3 until the solution converges (no cases switch from one cluster to another during step 2)

Output
a. Set of K centroids, each a vector of M features and corresponding to a cluster of data points
b. Classification of each training case, ie. which cluster does it belong to

Figure 2

Click here to access/download;Figure;Figure 2.tiff

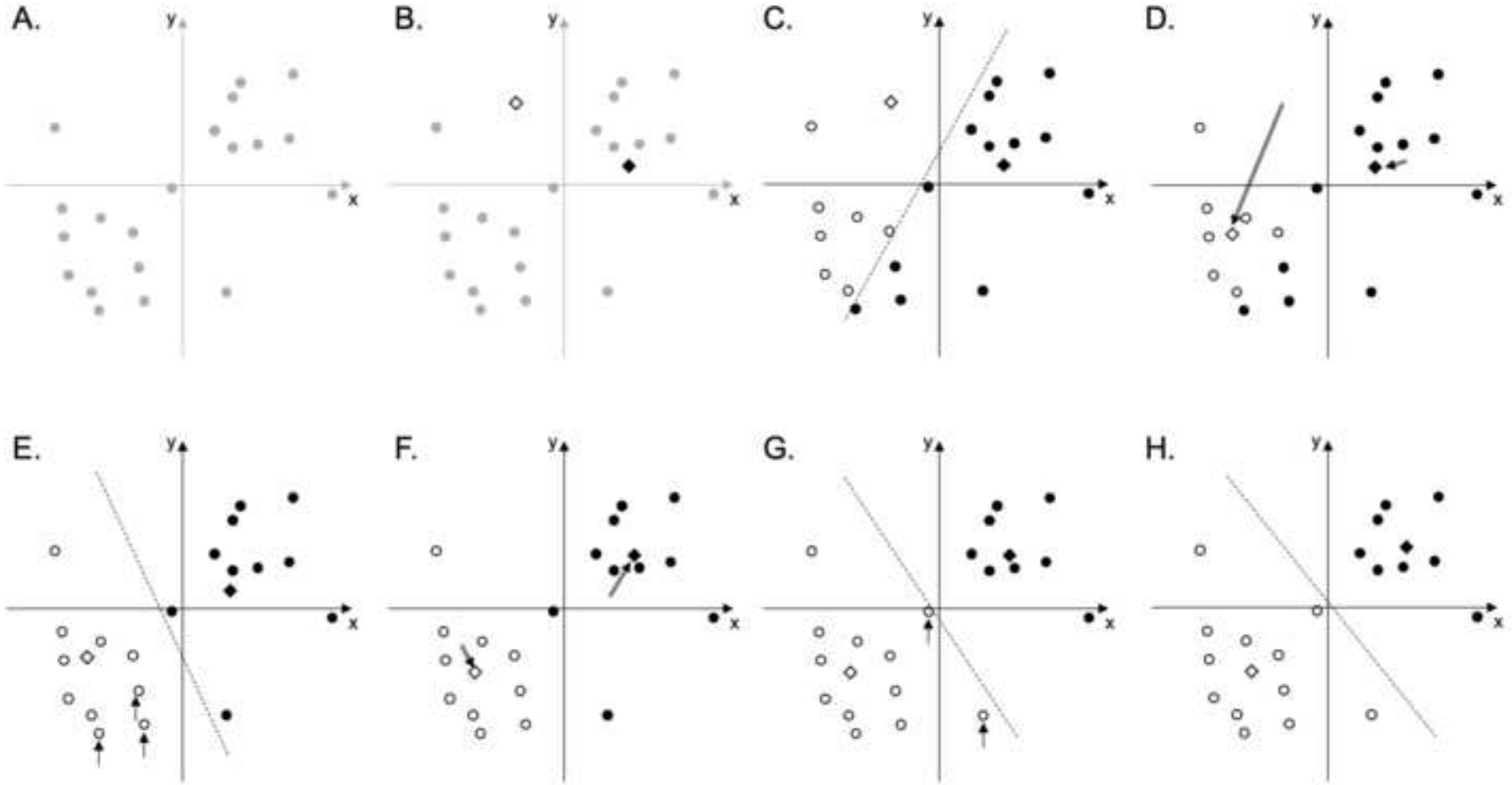

Table 1: Summary of study participants. F indicates female and M indicates male. N indicates negative and P indicates positive amyloid PET.

Clinical Cohort		Number	Age average (standard deviation)	Sex		Clinical PET interpretation	
				F	M	N	P
Clinically normal		19	67.9 (7.4)	11	8	18	1
Mild cognitive impairment		25	68.6 (7.2)	14	11	16	9
Dementia		22	70.2 (9.0)	10	12	12	10
Breakdown	AD	11	69.3 (10.6)	6	5	4	7
	VCI	3	73.7 (7.0)	2	1	2	1
	PD/LBD	8	70.1 (7.7)	2	6	6	2

Table 2: Confusion matrices for K=2. N indicates negative and P indicates positive amyloid PET.

		K-means	
		N	P
Clinical read	N	45	1
	P	1	19

Table 3: Centroids after clustering N=66 cases with M=19 features into K=2 clusters. Feature values are average Z-scores. N indicates negative and P indicates positive amyloid PET. Shaded boxes indicate ROIs where centroids have the largest differences in Z-score (≥ 8).

	N	P		N	P
Composite	0.8	10.4			
Cerebellar grey matter	-0.2	-0.3			
Pons	1.7	1.2			
Left prefrontal cortex	0.5	8.8	Right prefrontal cortex	0.6	9.2
Left anterior cingulate gyrus	0.3	6.5	Right anterior cingulate gyrus	0.5	6.9
Left precuneus/ posterior cingulate gyrus	0.8	9.2	Right precuneus/ posterior cingulate gyrus	0.5	8.8
Left parietal lobe	0.7	9.1	Right parietal lobe	0.9	9.8
Left lateral temporal lobe	1.0	9.5	Right lateral temporal lobe	1.0	9.5
Left occipital lobe	0.8	7.7	Right occipital lobe	1.0	8.0
Left sensorimotor cortex	0.5	5.7	Right sensorimotor cortex	0.7	6.0
Left mesial temporal lobe	0.3	2.3	Right mesial temporal lobe	0.8	3.0

Table 4: Summary of ML algorithms performance following reporting guidelines in (12).

ML algorithm	K-means clustering (unsupervised)
Architecture details	K = 2 clusters
Computational cost	Low: 0.26 s runtime (matlab.mathworks.com cloud server)
Data set	66 cases, 19 features
Figure of merit	Classification accuracy 97%