Automatic classification of MR scans in Alzheimer's Disease *(Klöppel et al., 2008)*

CMIC Journal Club, 29 Oct 2014

Contents

- Introduction
- Research hypotheses
- Methodology
- Results
- Contributions
- Limitations

Introduction

- Up to 2008, Alzheimer's disease (AD) was only "manually" diagnosed by clinicians using clinical exam, history, neuropsychological testing
- Pattern recognition methods applied to structural MRI were reported for the separation of mild cognitive impairment (MCI) from cognitively normal individuals (Davatzikos et al., 2006; Teipel et al., 2007)
- The machine learning technique used was Support Vector Machines
- However, not all patients with amnestic MCI will eventually develop AD

Research hypotheses

- Structural MRI can automatically distinguish AD from normal aging *robustly* using datasets from multiple scanners and different centres
- Classification based on the *grey matter* segment of pathologically confirmed cases can distinguish this using Support Vector Machines (SVMs)
- These methods can be used to differentiate scans between two *different types* of dementia, AD and frontotemporal lobar degeneration (FTLD)

Methodology

- Three datasets were used to distinguish AD from controls (Groups I, II, III) and a fourth one (Group IV) to distinguish between AD and FTLD
- Results on grey matter (GM) from the whole brain and part of it (antero-medial lobe) were obtained and contrasted
- Three main pre-processing steps:
  - Subjects were excluded from analysis if their scans revealed gross structural abnormalities other than atrophy
  - Images were fully segmented into GM, white matter and celebro-spinal fluid
  - Diffeomorphic registration algorithm (Ashburner,2007) was applied to GM segments

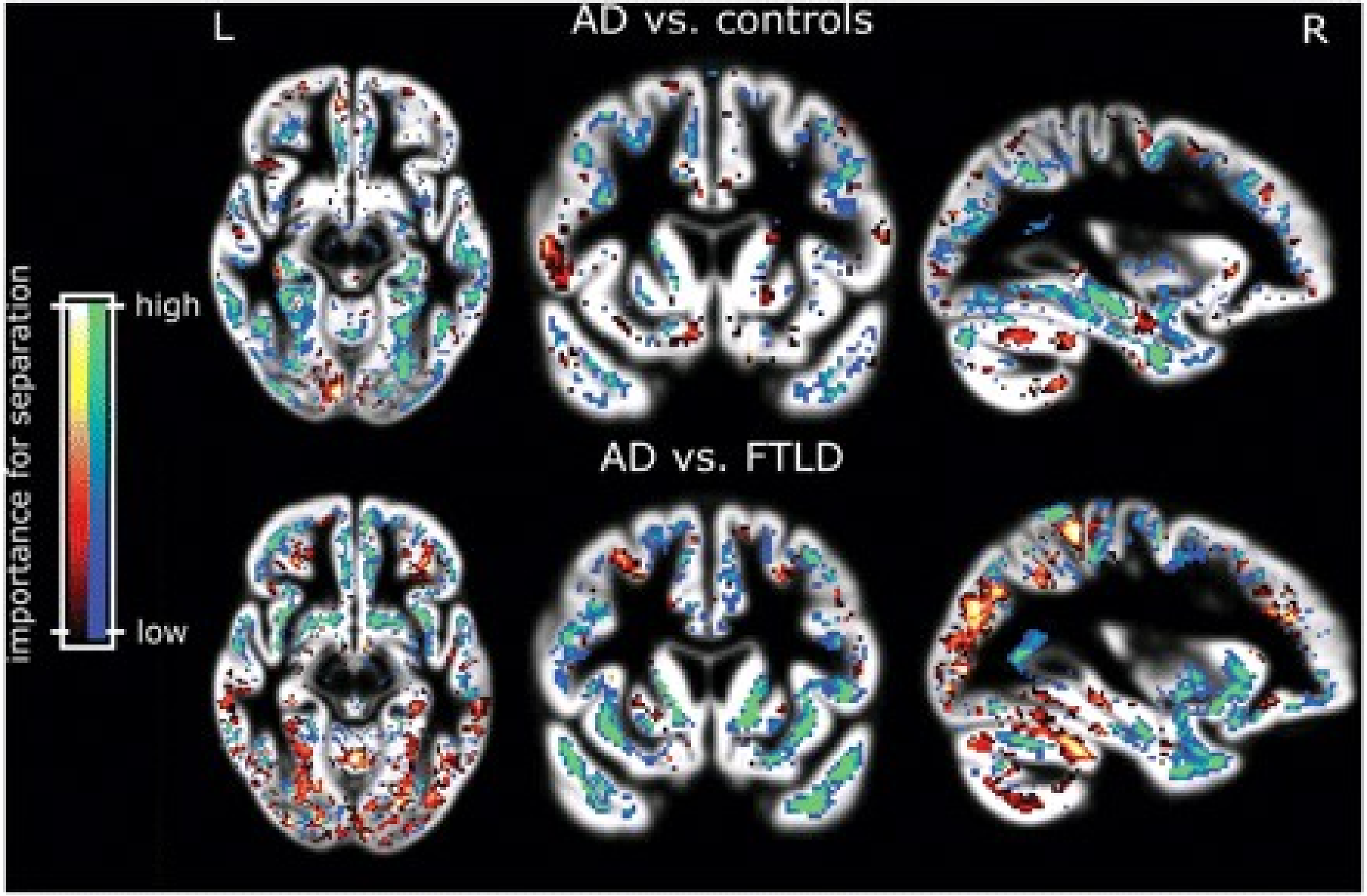
	Group I		Group II		Group III		Group IV	
Group (n)	AD (20)	controls (20)	AD (14)	controls (14)	AD (33)	Controls (57)	AD (18)	FTLD (19)
Sex (F/M)	11/9	10/10	5/9	5/9	10/23	16/41	6/12	8/11
Age at MRI-scan (mean, range)	81.0 (51–102)	79.5 (55–91)	65.0 (53–85)	63.0 (51–81)	73.1 (61–80)	71.9 (61–80)	66.0* (53–85)	61.7* (46–73)
MMSE – score (mean, range)	16.7 (7–29)	29.0 (27–30)	16.1 (10–20)	29.2 (28–30)	23.5 (20–28)	29.1 (27–30)	16.2* (5–29)	18.0 (0–26)
Years from MRI-scan to death (mean, range)	1.7 (0.2–3.4)	NA	3.6 (0.3–7.2)	NA	NA	NA	3.5 (0.3–7.2)	5.8 (1.3–11.0)

Results

- Up to 96% of pathologically verified AD patients versus control using whole brain images and up to 89% for AD versus FTLD
- Data from different centres were successfully combined achieving comparable results from the separate analyses
- Cases from Group III were handpicked as ‘mild’ AD and were classified against controls with correct separation of 89% of cases
- Qualitatively, SVM assigns a specific weight vector to each scan, whose components reflect the importance on the classification

Group	Correctly classified (%)	Sensitivity (%)*	Specificity (%)*
AD and controls Group I	95.0	95.0	95.0
AD and controls Group II	92.9	100	85.7
AD and controls Group III	81.1	60.6	93.0
Dataset I for training, set II for testing	96.4	100	92.9
Dataset II for training, set I for testing	87.5	95.0	80.0
Group I + II	95.6	97.1	94.1
AD from Dataset II and FTLD Group IV	89.2	83.3	94.7

Group	Correctly classified (%)	Sensitivity (%)*	Specificity (%)*
AD and controls Group I	90.0	85.0	95.0
AD and controls Group II	92.9	92.9	92.9
AD and controls Group III	85.6	75.8	91.2
Dataset I for training, set II for testing	71.4	50	92.9
Dataset II for training, set I for testing	70.0	95.0	45.0
Group I + II	94.1	97.1	91.2


L
AD vs. controls
R
high
low
importance for separation
AD vs. FTLD

Appendix 1 – Support Vector Machines

An SVM training algorithm is:

- a Supervised Learning algorithm used for either classification or regression
- a non-probabilistic, binary, linear classifier building a discriminative model
- able to create nonlinear classifiers by applying the kernel method*

* So kernel methods can be: Linear (dot product), Gaussian RBF, or even user-defined

Appendix 1 – Support Vector Machines

How SVM works (for linearly separable data):

1.Find the support vectors

2.Aim for the maximum optimal margin hyperplane (OMH) that optimally divides the data

3.Define a weight vector w, which is perpendicular to the OMH

4.Define an intercept term b, which is the offset/bias of the OMH

5.Solve the resulting Quadratic Optimisation problem

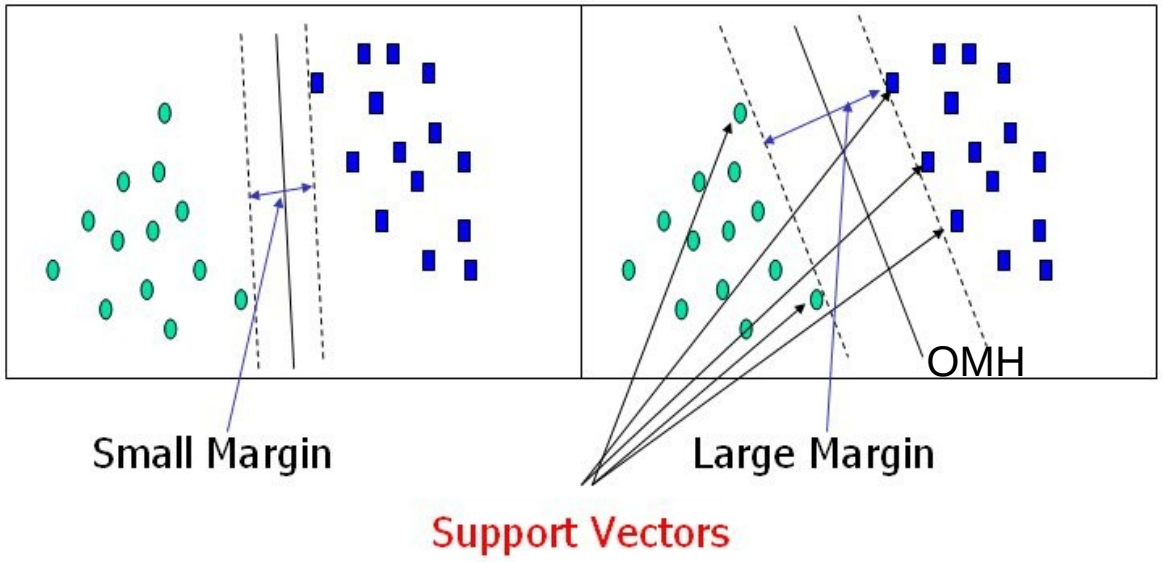
For nonlinearly separable data: Apply the soft margin, i.e. Add a regularisation hyperparameter C

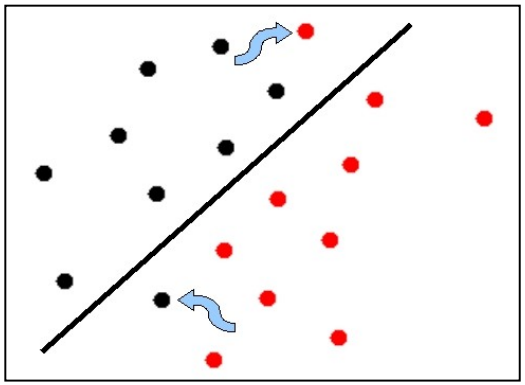
That is why this specific type of SVM is called C-SVM (or Type 1)

Appendix 1 – Support Vector Machines

Linearly separable case

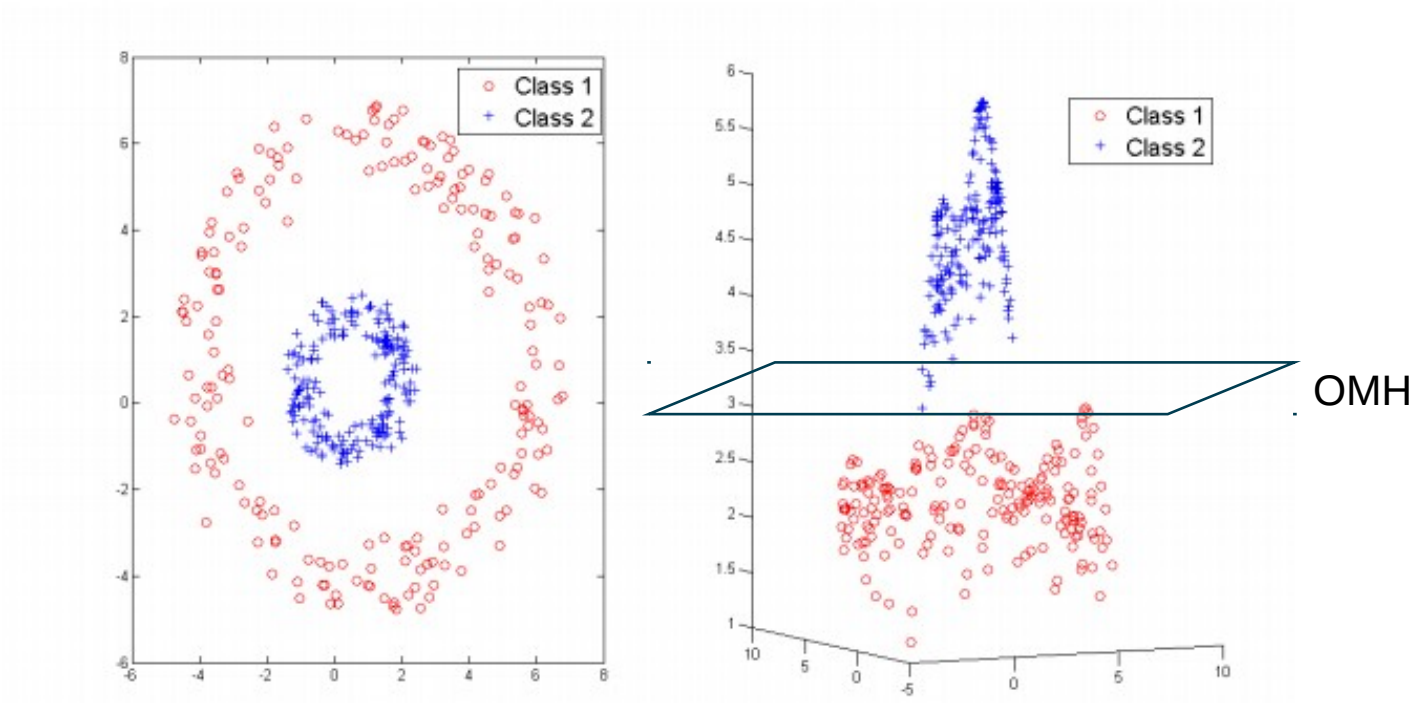
Nonlinearly separable case





Appendix 1 – Support Vector Machines

Kerned method can transform the data from one domain to an other:



Appendix 2 – Statistical measures

		Actual case	
		True	*False*
Test case	*True*	True Positive	False Positive
	False	False Negative	True Negative

$$\text{specificity} = \frac{\text{number of true negatives}}{\text{number of true negatives} + \text{number of false positives}}$$

$$= \frac{\text{number of true negatives}}{\text{total number of well individuals in population}}$$

$$= \text{probability of a negative test given that the patient is well}$$

$$\text{sensitivity} = \frac{\text{number of true positives}}{\text{number of true positives} + \text{number of false negatives}}$$

$$= \frac{\text{number of true positives}}{\text{total number of sick individuals in population}}$$

$$= \text{probability of a positive test, given that the patient is ill}$$

‘Limitations’

Konstantinos Georgiadis

Against

- No table/figure to compare with existing approaches
- Manual inspection of images for artefacts
- Insufficient explanation of methodology, Fig. 2 – aimed at people without math/machine learning BG

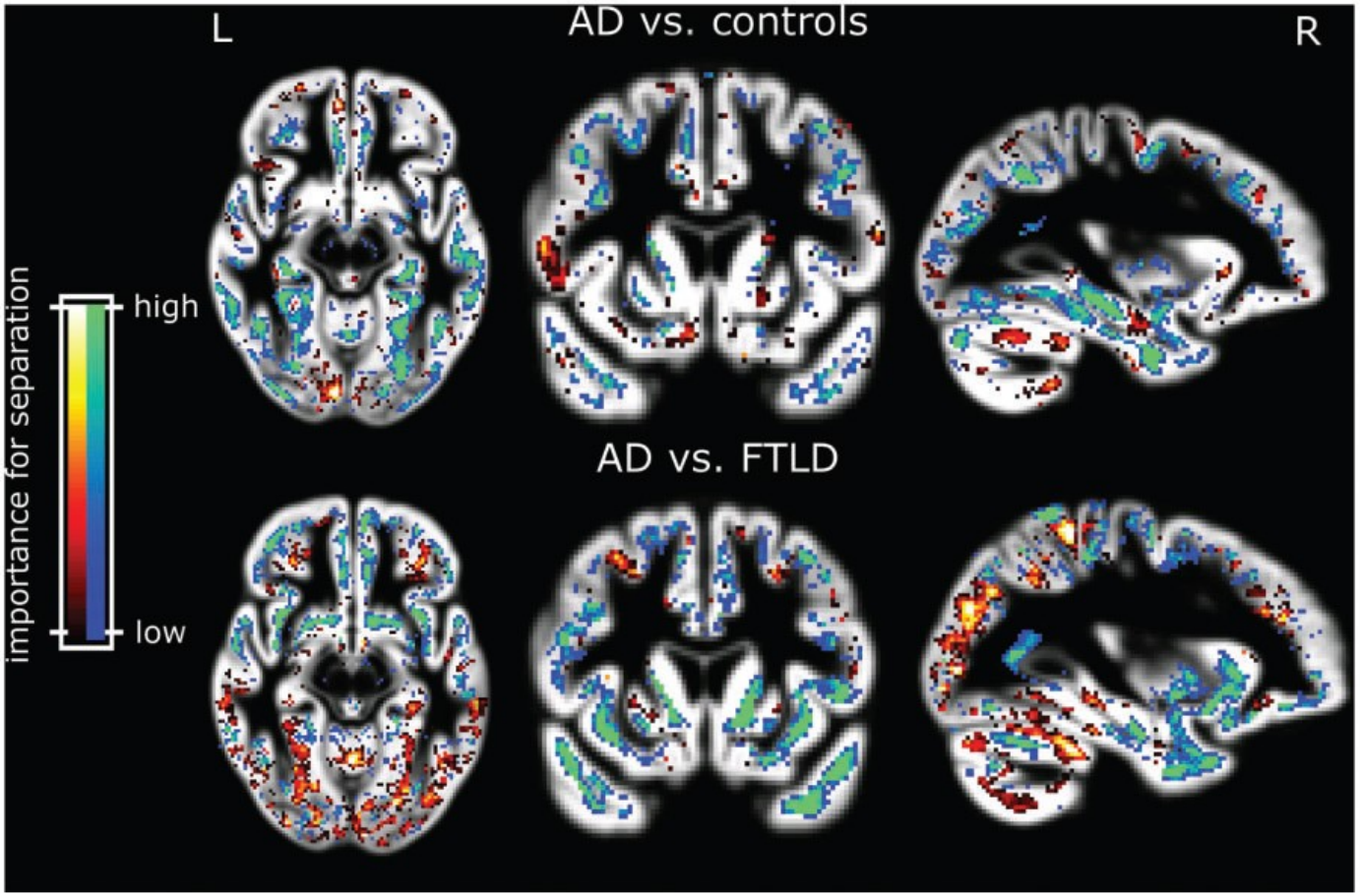


Fig. 2 Voxels most relevant for classification of patients from group I after SVM training with the data from group I (upper panel). The blue and green areas indicate higher grey matter volume increasing the likelihood of classification into normal. Red and yellow show regions where higher grey matter volume indicates the opposite. The lower panel depicts relevant areas for the separation from AD from FTLD. Blue and green indicate areas where lower grey matter volume indicates FTLD. Results are overlaid on the mean grey matter compartment image from all subjects.

Limitations

- For clinical use:
  - Automation
  - Recognize more diseases rather than just AD
  - Use many more scans of people with a variety of diseases (or none) for the learning process
- Dimensionality Reduction
- Probabilistic framework
- Use a space other than pixel intensity?
- Why only grey matter? White matter, CSF?

‘Contributions’

Thore Bücking

Classification Using the Whole Brain /ROI

Group	Correctly classified (%)		Sensitivity (%)*		Specificity (%)*	
AD and controls Group I	**95.0**	90.0	**95.0**	85.0	**95.0**	95.0
AD and controls Group II	**92.9**	92.9	**100**	92.9	**85.7**	92.9
AD and controls Group III	**81.1**	85.6	**60.6**	75.8	**93.0**	91.2
Dataset I for training, set II for testing	**96.4**	71.4	**100**	50	**92.9**	92.9
Dataset II for training, set I for testing	**87.5**	70.0	**95.0**	95.0	**80.0**	45.0
Group I + II	**95.6**	94.1	**97.1**	97.1	**94.1**	91.2
AD from Dataset II and FTLD Group IV	**89.2**		**83.3**		**94.7**	

SVMs Separate AD from Healthy Subjects

- Better than other classification methods
- Takes the whole brain into account

  *p.686 last paragraph
- First to consider every voxel directly for diagnosis

  *Cuingnet et. al., 2011
- Better than clinicians performance (?)
- NINCDS-ADRDA or DSM-IIIR criteria have an average sensitivity of 81% (range 49–100%) and specificity of 70% (range 47–100%).
- Formal comparison required

  *p.682 1st paragraph

Of Potential Use for Clinicians

- (fully) automated
  - No observer bias
  - Reproducible
  - Time saving

    *p.687 2nd paragraph
- Translatable to more general setting
  - Comparable results across different centres

    *p.687 3rd paragraph
  - Possible to incorporate other degenerative diseases?

    *p.687 1st paragraph

Overall…

- Demonstrated success of SVMs as diagnostic tool
- Demonstrated generalizability of approach

- Showed important role of computer based diagnostic image analysis for clinical practice