Classification of fMRI Data Acquired in Individuals With and Without Dyslexia: Comparison of Decision Tree Method and Linear Discriminant Analysis

José M. Maisog[1], Lynn Flowers[2], Guinevere F. Eden[1]
[1] Center for the Study of Learning, Georgetown Univ. Medical Center, Washington, DC 20007, USA
[2] Wake Forest Univ. School of Medicine, 1 Medical Center Boulevard, Winston-Salem, NC 27157, USA

INTRODUCTION

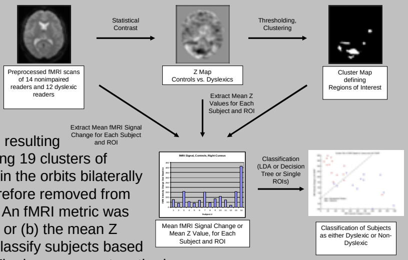
Previous work has shown that task-related signal changes in response to perception of visual motion are weaker in dyslexics compared to typical readers in area V5/MT [5,7]. This observation is consistent with psychophysical data indicative of a magnocellular deficit in developmental dyslexia. Measures of perceptual processes have the advantage that they can be assessed in pre-readers, whereas pure reading ability is not directly assessed prior to formal schooling. This raises the potential of using the neural responses elicited during visual motion perception as an early marker for reading disability.
We explored the possibility of classifying subjects as dyslexic or non-dyslexic based only on the functional magnetic resonance imaging (fMRI) profile elicited during the performance a visual motion task. We compared the discriminatory ability of linear discriminant analysis (LDA) [10,13], decision tree classification (“dtree”) [3,4], and single regions of interest (ROIs). We also compared two metrics: the average fMRI signal change versus the average Z value within a region of interest (ROI).

METHODS

Subjects: 12 adult dyslexic subjects (ages 18-44 years) and 14 matched nonimpaired readers (ages 20-55 years). Difference in mean reading skill was significant; not so for age (Table 1).
Scanning parameters: Multi-slice transverse EPI scans acquired on a 1.5 Tesla Siemens Vision scanner, TR/TE = 40 ms/3500 ms, 64x64 acquisition matrix, 230 mm FOV, 40 non-interleaved 3.0 mm slices, 0.6 mm gap, 3.6 mm^3 voxel size.
Tasks: Subjects viewed (1) moving dots and indicated the direction of the motion; (2) stationary dots and indicated which visual field had a higher dot density; and (3) a fixation cross [7].
Data analysis: fMRI scans were motion corrected [14], high-pass temporal filtered, and spatially smoothed with an 8mm Gaussian filter. A statistical contrast between tasks #1 and #2 was computed, generating a Z map for each subject. The two groups were then contrasted using "Stouffer's Method" [11]. The resulting Z map was thresholded (False Discovery Rate = 0.05 [9]), defining 19 clusters of contiguous suprathreshold voxels [8]. Four of the clusters were in the orbits bilaterally or in the left and right lateral ventricles posteriorly, and were therefore removed from further analysis. The remaining 15 clusters were used as ROIs. An fMRI metric was generated by computing either (a) the mean fMRI signal change or (b) the mean Z value, for each subject and ROI . LDA and dtree were used to classify subjects based on the fMRI metrics. Single ROIs were also tried as classifiers. The leave-one-out method was used for cross-validation, to estimate sensitivity, specificity, and % correctly classified.

Table 1.

		Controls	Dyslexics	t-test	p-value
Age	Minimum	20.00	18.20	0.76	0.46
	Average	40.74	43.68		
	Maximum	55.10	52		
WRAT single word reading standardized score	Minimum	100	75.00	-8.20	5.5E-08
	Average	111.21	90.08		
	Maximum	118	98		



RESULTS

Table 2 shows sensitivity/specificity/number correctly classified, for LDA and dtree.

Table 3 shows the same information, but this time for the single ROI classifiers.

Table 2.

Metric	LDA	dtree
Mean fMRI Signal Change	83% / 29% / 53%	41% / 21% / 12%
Mean Z value	58% / 79% / 69%	17% / 7% / 31%

Table 3.

Cluster #		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Side		Bilateral	Left	Bilat	Left	Right	Left	Right	Right	Right	Left	Bilateral	Right	Left	Left	Left
Brain Region		Cuneus	Lingual Gyrus / Cuneus / Middle Occipital Gyrus	Precuneus	Middle Frontal Gyrus	Superior / Middle Frontal Gyrus	Superior Frontal Gyrus	Middle Frontal Gyrus	Superior / Middle Frontal Gyrus	Precentral Gyrus / Superior Temporal Gyrus	Superior Frontal Gyrus	Culmen of Cerebellum	Cuneus / Lingual Gyrus	Medial Frontal Gyrus	Cuneus	Pons
Brodmann Area		18/17	18/17/19	7	8/6/9	10/11	11/10	9/8	8/6	22/6	10		17/18	11/10	19/7	
Dyslexic - Control		Neg	Pos	Pos	Pos	Pos	Pos	Pos	Pos	Neg	Pos	Pos	Pos	Neg	Pos	Pos
Size (voxels)		1613	1292	767	471	391	348	319	260	190	161	158	153	153	149	146
Mean fMRI Signal Change	Sensitivity	75%	75%	83%	92%	83%	83%	83%	92%	100%	83%	75%	67%	83%	58%	75%
	Specificity	93%	79%	57%	86%	79%	71%	79%	57%	79%	71%	64%	71%	64%	57%	86%
	% Correct	85%	77%	69%	89%	81%	77%	81%	73%	89%	77%	69%	69%	73%	58%	81%
Mean Z Value	Sensitivity	75%	75%	83%	83%	75%	83%	83%	83%	100%	92%	75%	58%	75%	58%	92%
	Specificity	86%	71%	64%	79%	71%	71%	86%	71%	79%	79%	71%	71%	64%	64%	86%
	% Correct	81%	73%	73%	81%	73%	77%	85%	77%	89%	85%	73%	65%	69%	62%	89%

SUMMARY

- LDA had superior performance to dtree, perhaps due to the relatively low prediction accuracy of decision trees in general [2].
- The simple single ROI classifiers did surprisingly well compared to LDA and dtree.
- Neither of the two metrics (mean fMRI signal change vs. mean Z value) seemed to have a decided advantage, although mean fMRI signal change resulted in slightly better sensitivity overall (not statistically significant).
- The accuracy of these results is comparable to classification studies using neuroanatomical data in dyslexia [12] and functional brain imaging data in neurological disorders [1,6].

Further refinement and extension to pediatric populations could lead to an fMRI-based tool for detecting children at risk for dyslexia, and aid in the identification of candidates suitable for intervention at an early age. Future directions: (1) use receiver operator characteristic curve methods to compare classification methods; (2) use larger sample sizes; (3) test other methods of classification, e.g. support vector machines; (4) test other metrics (e.g., maximum value within ROI, exceedance mass); (5) use ensemble methods [for review, see 2] to improve the performance of dtree.

We thank Christian Borgelt (Otto-von-Guericke-University of Magdeburg) and Habtom Ressom (Georgetown University Medical Center) for their kind assistance. This work was funded by NICHD (HD37890, HD40095), NIMH (MH6500), and the Latham Trust. This poster is available online at http://csl.georgetown.edu/publications/DyslexiaClassificationPoster-HBM2006.pdf

REFERENCES

[1] Azari et al. (1993) Journal of Cerebral Blood Flow and Metabolism 13(3):438-47.
[2] Blower and Cross (2006) Current Topics in Medicinal Chemistry 6(1):31-9.
[3] Borgelt (1998a) A Decision Tree Plug-In for DataEngine. 2nd Data Analysis Symposium.
[4] Borgelt (1998b) A Decision Tree Plug-In for DataEngine. 6th European Congress on Intelligent Techniques and Soft Computing 2, 1299-1303.
[5] Demb et al. (1997) Proceedings of the National Academy of Sciences U S A 94(24):13363-6.
[6] Eckert et al. (2004) Neuroimage 21(1):229-35.
[7] Eden et al. (1996) Nature 382(6586):66-9.
[8] Friston et al. (1994) Human Brain Mapping 1(3):214-20.
[9] Genovese et al. (2002) Neuroimage 15(4):870-8.
[10] Klecka (1980) Discriminant analysis. Sage Publications, Beverly Hills, Calif.
[11] Lazar et al. (2002) Neuroimage 16(2):538-50.
[12] Leonard et al. (1999) Biological Psychiatry 46(3):374-82.
[13] Rencher (1995) Methods of multivariate analysis. Wiley, New York.
[14] Woods et al. (1998) Journal of Computer Assisted Tomography 22(1):139-152.