Texture and Regression Tree Analysis in the Characterisation of Ovarian Lesions

Peter Gibbs[1], Martine Dujardin[1], and Lindsay Turnbull[1]
[1]*MRI Centre, HYMS at University of Hull, Hull, East Yorkshire, United Kingdom*

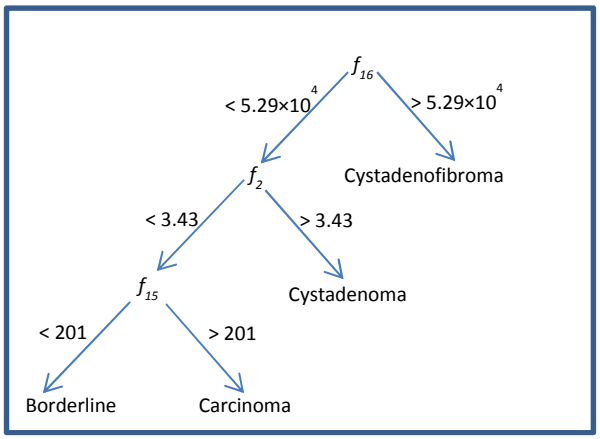
Target Audience Ovarian MR researchers – Physicists and Clinicians

Purpose The characterisation of complex ovarian lesions is an on-going challenge. Because of its greater soft tissue contrast resolution MRI is the preferred technique for characterising complex adnexal masses(1). However, the presence of solid components in benign lesions such as cystadenofibroma, a feature that overlaps with malignant lesions, causes diagnostic difficulty(2). Correct diagnosis is very important since benign ovarian lesions may be treated by simple cystectomy or oophorectomy, whilst malignant lesions require hysterectomy, bilateral oophorectomy, omentectomy and possibly appendectomy. Textural analysis has previously been utilised in contrast enhanced MRI of the breast both as a diagnostic tool(3) and as a predictor of chemotherapeutic response(4). This study aims to explore the utility of texture analysis and subsequent regression tree analysis in the diagnosis of ovarian malignancy.

Methods Data from 96 women with histopathologically proven ovarian cancer (*n*=67), borderline ovarian tumour (*n*=28), cystadenoma (*n*=14) or cystadenofibroma (*n*=19) who underwent pre-operative pelvic MRI using a 32 channel phase array coil on a 3 Tesla scanner was retrospectively analysed. Texture analysis was performed on T_2 weighted images acquired with the following parameters: TE 111 ms, TR 3431 ms, FOV 24×24 cm, matrix size 512×416, slice thickness/gap 4/1 mm, acquisition time 5 mins for ~40 slices. ROIs were manually drawn on a single slice by an expert radiologist, detailing the most complex portion of the lesion. ROI data was then reduced to 16 grey levels using histogram equalisation to ensure adequate counting statistics. Co-occurrence matrices, which represent the probability of finding 2 adjacent pixels of intensities *i* and *j* were then computed for four directions (0°, 45°, 90°, and 135°) to enable subsequent calculation of texture parameters f_1 to f_{16} as described by Haralick *et al*(5) and Conners *et al*(6). Since no intrinsic directionality is anticipated average texture parameters were utilised in statistical analysis. Due to the uneven group sizes, which can result in over emphasis on trying to correctly predict the largest group, repeated testing of equal sample sizes (*n*=14 for each group) via random sampling was employed. The Kruskal-Wallis test was utilised to determine significant differences between groups. Pearson's correlation coefficient was then used to remove redundant parameters prior to regression tree analysis.

Results Significant differences between the four groups were consistently noted for 8 parameters (f_2, f_4, f_5, f_7, f_{10}, f_{11}, f_{15} and f_{16}) and are detailed in the accompanying table for one set of random data samples. Once correlation analysis was performed 5 parameters were retained and thus inputted into regression tree analysis using the CART algorithm. After tree pruning to prevent over fitting a final classification tree (see diagram) contained 3 parameters (f_2 – contrast, f_{15} - cluster shade and f_{16} – cluster prominence). Using this model correctly categorises 11.4/14 (81%) of cystadenofibromas, 9.8/14 (70%) of cystadenomas, 8.7/14 (62%) of borderline ovarian tumours and 9.4/14 (67%) of ovarian cancers.

	Median value				
Parameter	Cystadenofibroma	Cystadenoma	Borderline	Cancer	*p* value
f_2	9.99	4.10	3.12	3.06	<0.0001
f_4	49.1	25.1	18.3	21.6	0.0002
f_5	0.496	0.539	0.556	0.557	0.0711
f_7	214	105	77	89	0.0002
f_{10}	5.34	2.11	1.70	1.44	<0.0001
f_{11}	2.52	2.18	2.04	2.04	0.0002
f_{15}	-830	-678	70	335	0.0042
f_{16}	10.1×10^4	2.9×10^4	1.6×10^4	2.4×10^4	<0.0001



Conclusions Texture analysis has been successfully applied in the diagnosis of ovarian malignancy. After performing repeated testing, via random sampling, a robust diagnostic model has been developed, with an overall accuracy of 70%. By appropriate use of correlation analysis and tree pruning only 3 parameters were retained, thus avoiding over parameterisation in the final model. Future work may include incorporation of other MR quantitative features such as shape.

(1) YZ Tang *et al* (2013) *European Radiology* 23:48-56. (2) SAA Sohaib *et al* (2003) *American Journal of Roentgenology* 180:1297-1304. (3) P Gibbs and LW Turnbull (2003) *Magnetic Resonance in Medicine* 50:92-98. (4) A Ahmed *et al* (2013) *Journal of Magnetic Resonance Imaging* 38:89-101. (5) RM Haralick *et al* (1973) *IEEE Transactions on Systems, Man and Cybernetics* 3:610-621. (6) RW Conners and CA Harlow (1980) *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2:204-222.