Table 6. Performance comparison among three classifiers with GMM features.
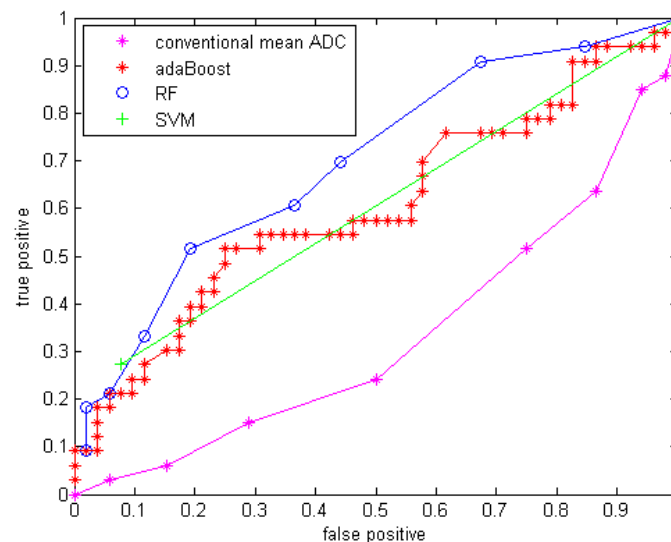
Classifier	Sensitivity	Specificity	Accuracy	Az
AdaBoost	39.39%	80.77%	64.7%[*]	0.60
Random forest	51.52%	80.77%	69.41%[*]	0.70
SVM	27.27%	92.3%	67.06%[*]	0.60

(*: All p-values <0.0001 comparing with accuracy of Table 4)

Figure 5. ROC curve for three classifiers with GMM features.

6. Discussion

Compared to using only the mean ADC value, the quantitative statistical histogram features and the proposed classification system tremendously improved the accuracy from 29.4% to 69.41% (Az increased from 0.33 to 0.70). The statistical analysis indicates that all three classifiers are significantly different from the conventional mean ADC method with our dataset. Compared to general statistical histogram features, the classification with GMM features using random forest technique slightly improved the accuracy from 65.88% to 69.41%, while adaBoost and RF classifiers generated the same accuracy no matter whether GMM features were included. There is no significant difference between the three machine-learned classifiers.

The conventional mean ADC method performs worse than a random classifier (Az < 0.5). The reason is that conventionally researchers hypothesized that mean ADC increases because the tumor cell density decrease after an effective treatment. This assumption may not be valid for our dataset, because it involves in an anti-angiogenesis drug, which suppresses the cancer cell growth without necessary killing tumor cells (decreasing their density) at an early stage (5-7 weeks). Another possible reason is that in our dataset many of the GBM tumors are recurrent GBM tumors that are usually necrotic. The treatment tends to reduce necrosis and edema, which will diminish ADC. Essentially there are two competing processes at work: cell density, edema and necrosis [25].

Another state-of-art study included features that capture spatial information in tumor heterogeneity features. Functional diffusion map (fDM) [22,23] is a popular technique studying the ADC value increase or decrease voxel-by-voxel. Moffat et al. applied fDM to 20 patients, classified patients into

the three categories: PR, SD and PD, and reported 100% accuracy [22]. However, the threshold they used for classification was determined from a single dataset of 20 patients used for both training and testing, while in our experiments, a cross validation analysis was performed. In Moffat et al's study, they explored the assessment of fractionated radiation therapy for different types of brain tumors with 20 patients scanned on the same scanner [22]. However, in our study, we focused on the GBM brain tumors treated by anti-angiogenesis drugs, which suppress the blood supply for the tumor cells and may not directly decrease the tumor cellularity. The difference in accuracy may come from the different mechanism of treatment. Additionally, our dataset is from GBM drug trials across multiple sites, thus our preliminary study is an important contribution for exploring DWI as an early imaging biomarker in a real pharmaceutical drug trial. In future work, we will extract texture feature to include spatial information, and shape features will be extracted as well. By introducing a new richer feature set indicating more useful tumor information, we aim to include more information about tumors and further improve the performance of the classification system.

One limitation of this study is that we classified CR, PR and SD as responders for the ground truth to achieve a binary classification. Since SD and PR may have different patterns in terms of their ADC histogram change, a multi-category classification system will be explored in future work. Another limitation of the study is that we used the Macdonald criteria at the eighth or tenth week after treatment for determining treatment response. In future work, time-to-progression and survival time will be a better endpoint to classify treatment response. Another limitation comes from the 3D ROI mapping tool. This tool is more computationally efficient compared to the co-registration techniques, but it cannot correct for patient motion. Therefore, in our study, a board-certified radiologist's visually checked and edited all segmentation results as needed. In the future, a more sophisticated registration method with an image similarity measure may improve the accuracy of the tumor contours on ADC maps, and consequently improve the accuracy of the extracted features and the classifier performance.

ADC values obtained on pre-operative MRI scans are reported to be of prognostic value in patients with glioblastoma [25,42]. The term "prognosis" refers to predicting the likely outcome of treatment. ADC, reported to be inversely proportional to tumor cellularity, is gaining interest in predicting GBM tumor prognosis. Our proposed framework now uses changes in DW-MRI for early prediction of treatment response; however, the framework with feature extraction and machine learning technique could be generalized to pre-treatment DW-MRI for prognosis prediction.

In this study, we developed a CADrx framework with machine learning techniques to automatically predict tumor treatment response before the size change using DW-MRI. In our preliminary study, our major contributions are extracting statistical ADC histogram features, applying GMM to model the ADC histogram to interpret the competing effects of cellular density and edema, and applying machine learning techniques using all the extracted features. Cell density and edema may be reflected in ADC values before size changes are apparent on standard MRI sequences. Therefore, ADC holds promise as a biomarker, in determining both which tumors are more likely to respond to treatment and which tumors are actually responding.

In conclusion, this work shows that a CADrx system using quantitative ADC histogram features and a machine-learned classifier has better performance in treatment response assessment over conventional analysis using only a mean ADC value. This will have major implications for clinical trials. This work has potential clinical significance for early treatment response assessment in GBM.

References

1. Giger, M.L. *Computer-aided diagnosis in medical imaging — A new era in image interpretation*; Technical Report; World Markets Research Centre: London, UK, 2000; pp. 75–78.

2. Ross, B.D.; Moffat, B.A.; Lawrence, T.S.; Mukherji, S.K.; Gebarski, S.S.; Quint, D.J.; Johnson, T.D.; Junck, L.; Robertson, P.L.; Muraszko, K.M.; Dong, Q.; Meyer, C.R.; Bland, P.H.; McConville, P.; Geng, H.; Rehemtulla, A.; Chenevert, T.L. Evaluation of cancer therapy using diffusion magnetic resonance imaging. *Mol. Cancer Ther.* **2003**, *2*, 581–587.

3. Padhani, A.R.; Liu, G.; Mu-Koh, D.; Chenevert, T.L.; Thoeny, H.C.; Takahara, T.; Dzik-Jurasz, A.; Ross, B.D.; Cauteren, M.V.; Collins, D.; Hammoud, D.A.; Rustin, G.J.S.; Taouli, B.; Choyke, P.L. Diffusion-weighted magnetic imaging as a cancer biomarker: consensus and recommendations. *Neoplasia* **2009**, *11*, 102–125.

4. Phillips, W.E.; Velthuizen, R.P.; Phupanich, S.; Hall, L.O.; Clarke, L.P.; Silbiger, M.L. Applications of fuzzy C-means segmentation technique for tissue differentiation in MR images of a hemorrhagic glioblastoma multiforme. *J. Magn. Reson. Imaging* **1995**, *13*, 277–290.

5. Clark, M.C.; Hall, L.O.; Goldgof, D.B.; Velthuizen, R.; Murtagh, R.; Silbiger, M.S. Automatic tumor segmentation using knowledge-based techniques. *IEEE Trans. Med. Imaging* **1998**, *17*, 187–201.

6. Fletcher-Heath, L.M.; Hall, L.O.; Goldgof, D.B.; Murtagh, R.F. Automatic segmentation of non-enhancing brain tumors in magnetic resonance images. *Artif. Intell. Med.* **2001**, *21*, 43–63.

7. Prastawa, M.; Bullitt, E.; Moon, N.; Leemput, K.V.; Gerig, G. Automatic brain tumor segmentation by subject specific modification of atlas priors. *Acad. Radiol.* **2003**, *10*, 1341–1348.

8. Kaus, M.; Warfield, S.; Nabavi, A.; Black, P.M.; Jolesz, F.A.; Kikinis, R. Automated segmentation of mr images of brain tumors. *Radiology* **2001**, *218*, 586–591.

9. Ho, S.; Bullitt, E.; Gerig, G. Level set evolution with region competition: Automatic 3-d segmentation of brain tumors. In *Proceedings of International Conference on Pattern Recognition*, Quebec, Canada, August, 2002; pp. 532–535.

10. Vinitski, S.; Gonzalez, C.F.; Knobler, R.; Andrews, D.; Iwanaga, T.; Curtis, M. Fast tissue segmentation based on a 4D feature map in characterization of intracranial lesions fast tissue segmentation based on a 4D feature map in characterization of intracranial lesions. *J. Magn. Reson. Imaging* **1999**, *9*, 768–776.

11. Lee, C.H.; Schmidt, M.; Murtha, A.; Bistritz, A.; Sander, J.; Greiner, R. Segmenting brain tumor with conditional random fields and support vector machines. In *Proceedings of Workshop on Computer Vision for Biomedical Image Applications at International Conference on Computer Vision*, Beijing, China, October, 2005; Vol. 3765, pp. 469–478.

12. Zhu, Y.; Yan, H. Computerized tumor boundary detection using a hopfield neural network. *LEEE Trans. Med. Imaging* **1997**, *16*, 55–67.

13. Zhang, J.; Ma, K.; Er, M.H.; Chong, V. Tumor segmentation from magnetic resonance imaging by learning via one-class support vector machine. In *Proceedings of International Workshop on Advanced Image Technology*, Singapore, January, 2004; pp. 207–211.

14. Niea, J.; Xue,.; Liu, T.; Young, G.S.; Setayesh, K.; Guo, L.; Wong, S.T.C. Automated brain tumor segmentation using spatial accuracy-weighted hidden Markov Random Field. *Comput. Med. Imaging Graph.* **2009**, *33*, 431–441.

15. Corso, J.J.; Sharon, E.; Dube, S.; El-Saden, S.; Sinha, U.; Yuille, A. Efficient Multilevel Brain Tumor Segmentation with Integrated Bayesian Model Classification. *IEEE Trans. Med. Imaging* **2008**, *27*, 629–640.

16. Liu, J.; Udupa, J.; Odhner, D.; Hackney, D.; Moonis, G. A system for brain tumor volume estimation via mr imaging and fuzzy connectedness. *Comput. Med. Imaging Graph.* **2005**, *29*, 21–34.

17. Dube, S.; Corso, J.J.; Yuille, A.; Cloughesy, T.F.; El-Saden, S.; Sinha, U. Hierarchical Segmentation of Malignant Gliomas via Integrated Contextual Filter Response. *Proc. SPIE* **2008**, *6914*, 69143Y.

18. Dube, S.; Corso, J.J.; Cloughesy, T.F. ; El-Saden, S.; Yuille, A.; Sinha, U. Automated MR image processing and analysis of malignant brain tumors: enabling technology for data mining. In *Data Mining Systems Analysis and Optimization in Biomedicine*; American Institute of Physics Proceedings: New York, NY, USA, 2007; Vol. 953, pp. 64–84.

19. US Food and Drug Administration, Guidance for industry: clinical trial endpoints for the approval of cancer drugs and biologics. *Federal Register* **2007**, *72*, No. 94.

20. Chenevert, T.L.; Stegman, L.D.; Taylor, J.M.; Robertson, P.L.; Greenberg, H.S.; Rehemtulla, A.; Ross, B.D. Diffusion magnetic resonance imaging: an early surrogate marker of therapeutic efficacy in brain tumors. *J. Natl. Cancer Inst.* **2000**, *92*, 2029–2036.

21. Mardor, Y.; Pfeffer, R.; Spiegelmann, R.; Roth, Y.; Maier, S.E.; Nissim, O.; Berger, R.; Glicksman, A.; Baram, J.; Orenstein, A.; Cohen, J.S.; Tichler, T. Early detection of response to radiation therapy in patients with brain malignancies using conventional and high b-value diffusion-weighted magnetic resonance imaging. *J. Clin. Oncol.* **2003**, *21*, 1094–1100.

22. Moffat, B.A.; Chenevert, T.L.; Meyer, C.R.; Mckeever, P.E.; Hall, D.E.; Hoff, B.A.; Johnson, T.D.; Rehemtulla, A.; Ross, B.D. The functional diffusion map: a noninvasive MRI biomarker for early stratification of clinical brain tumor response. *PANS* **2005**, *102*, 5524–5529.

23. Hamstra, D.A.; Chenevert, T.L.; Moffat, B.A.; Johnson, T.D.; Meyer, C.R.; Mukherji, S.K.; Quint, D.J.; Gebarski, S.S.; Fan, X.; Tsien, C.I.; Lawrence, T.S.; Junck, L.; Rehemtulla, A.; Ross, B.D. Evaluation of the functional diffusion map as an early biomarker of time-to-progression and overall survival in high-grade glioma. *PNAS* **2005**, *102*, 16759–16764.

24. Huo, J.; Kim, H.J.; Pope, W.B.; Okada, K.; Alger, J.R.; Wang, Y.; Goldin, J.G.; Brown, W.S. Histogram-based classification with Gaussian mixture modeling for GBM tumor treatment response using ADC map. *Proc. SPIE* **2009**, *7260*, 72601Y.

25. Pope, W.B.; Kim, H.J.; Huo, J.; Alger, J.R.; Brown, W.S.; Gjertson, D.; Sai, V.; Young, J.R.; Tekchandani, L.; Cloughesy, T.; Mischel, P.S.; Lai, A.; Nghiemphu, P.; Rahmanuddin, S.; Goldin, J.G. Recurrent glioblastoma multiforme: ADC histogram analysis predicts response to bevacizumab treatment. *Radiology* **2009**, *252*, 1–8.

26. Sajda, P. Machine learning for detection and diagnosis of disease. *Annu. Rev. Biomed. Eng.* **2006**, *8*, 537–65.

27. Freund, Y.; Schapire, R.E. A short introduction to boosting. *J. Jpn. Soc. For. Artif. Intell.* **1999**, *14*, 771–780

28. Duda, R.O.; Hart, P.E.; Stork, D.H. *Pattern classification*; Wiley Interscience: Malden, MA, USA, 2000.

29. Ho, T.K. Random decision forest. In *Proceedings of the 3rd International Conference on Document Analysis and Recognition* Montreal, Canada, August, 1995; pp. 278–282.

30. Breiman, L.: Random decision forest. *Mach. Learn.* **2001**, *45*, 5–32.

31. Vapnik, V. *Estimation of Dependencies Based on Empirical Data*; Nauka: Moscow, Russia, 1979.

32. Bishop, C. *Neural Networks for Pattern Recognition*; Clarendon Press: Oxford, UK, 1995.

33. http://en.wikipedia.org/wiki/Support_vector_machine (accessed November 10, 2009).

34. Burbidge, R.; Trotter, M.; Buxton, B.; Holden, S. Drug design by machine learning: support vector machines for pharmaceutical data analysis. *Comput. And. Chem* **2001**, *26*, 5–14.

35. Huo, J.; Alger, J.R.; Kim, H.J.; Pope, W.B.; Okada, K.; Goldin, J.G.; Brown, M.S. Between-scanner variation in normal white matter ADC in the setting of a multi-center clinical trial. *Ismrm* 2009, (in press).

36. Otsu N. A threshold selection method from gray level histograms. *IEEE Trans. Syst. Man. Cybern.* **1979**, *9*, 62–66.

37. Adams, R.; Bischof, L. Seeded region growing. *IEEE Trans. Syst. Man. Cybern. Int.* **1994**, *16*, 641–647.

38. Rubner, Y.; Tomasi, C.; Guibas, L.J. A metric for distributions with applications to image databases. In *Proceedings of ICCV*, Bombay, India, January, 1998; pp. 59–66.

39. Ling, H.; Okada, K. An efficient Earth mover's distance algorithm for robust histogram comparison. *IEEE Trans. Patt. Anal. Mach. Intell.* **2007**, *29*, 840–853.

41. Witten, I.H.; Frank, E. *Data Mining: Practical Machine Learning Tools and Techniques*; Morgan Kaufmann: San Francisco, CA, USA, 2005.

42. Yamasaki, F.; Sugiyama, K.; Ohtaki, M.; Takeshima, Y.; Abed, N.; Akiyamad, Y.; Takabad, J.; Amatyac, V.J.; Saitoa, T.; Kajiwaraa, Y.; Hanayaa, R.; Kurisua, K. Glioblastoma treated with postoperative radio-chemotherapy: Prognostic value of apparent diffusion coefficient at MR imaging. *Eur.J. Aiol*. 2009, (in press).

43. Marzban, C. The ROC curve and the area under it as a performance measure. *Weather Forecast.* **2004**, *19*, 1106–1114.

44. Huhn, S.L.; Mohapatra, G.; Bollen, A.; Lamborn, K.; Prados, M.D.; Feuerstein, B.G. Chromosomal abnormalities in glioblastoma multiforme by comparative genomic hybridization: correlation with radiation treatment outcome. *Clin. Cancer Res.* **1999**, *5*, 1435–1443.