Accurate Classification of Parotid Tumors Based on Apparent Diffusion Coefficient

Anahita Fathi Kazerooni1,2, Sanam Assili2, Mohammad Reza Alviri1, Mahnaz Nabil2, Jalil Pirayesh Islamian4, Hamidreza Saligheh Rad1,2, Leila Agha-Ghazvini1,6*

1- Quantitative MR Imaging and Spectroscopy Group, Research Center for Cellular and Molecular Imaging, Tehran University of Medical Sciences, Tehran, Iran.
2- Medical Physics and Biomedical Engineering Department, Tehran University of Medical Sciences, Tehran, Iran.
3- Department of Statistics, Faculty of Mathematical Science, University of Guilan, Rasht, Iran.
4- Medical Physics Department, Faculty of Medicine, Tabriz University of Medical Sciences, Tabriz, Iran.
5- Department of Radiology, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran.
6- Department of Radiology, Amir Alam Hospital, Tehran University of Medical Sciences, Tehran, Iran.

ABSTRACT

Purpose- In this work, we aimed to propose an automatic classification scheme based on the parameters derived from Apparent Diffusion Coefficient (ADC)-maps for discriminating benign and malignant parotid tumors.

Methods- MRI was carried out prospectively on 41 patients presented with parotid tumors who underwent surgery and post-surgical histopathological assessment was provided for them (32 benign, 9 malignant). Based on anatomical images, Regions Of Interest (ROIs) were selected on the most solid parts of tumors on ADC-maps. Three quantitative parameters, namely ADC-Mean, ADC-Max and ADC-Min were calculated. An automatic classification of parotid tumors using ADC parameters was performed and assessed employing two different classifiers, namely, Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA).

Results- Based on statistical analysis, it was indicated that the ADC values in benign tumors are significantly higher than malignant tumors. ADC-Mean, and -Max presented statistically significant differences among benign and malignant parotid tumors (p<0.05). Among the extracted parameters, ADC-Max is the most relevant quantitative parameter for tumor classification with 82.9% accuracy, 84.4% specificity, 77.8% sensitivity, and 83.3% area under the ROC curve (AUC) by exploiting each of the automatic classifiers. This implies that this parameter is inherently accurate and adding further classification complexity does not improve the results. A linear classifier using LDA classification based on ADC-max is proposed, which indicates that ADC-Max under $1.48 \times 10^{-3}$ mm$^2$/s is highly suggestive of malignancy (with 83% accuracy).

Conclusion- ADC-Max is a potential biomarker for discriminating benign and malignant parotid tumors. Using ADC-Max and LDA, a simple and clinically-feasible classifier is proposed.

1. Introduction

Early 2-3% of head and neck tumors occur in parotid regions [1, 2]. Parotid is the largest salivary gland and comprises the majority of salivary gland tumors [3]. In terms of morphology, parotid tumors represent a variety of histologic types of neoplasms, for which the majority of lesions (about 80%) are benign.
Pleomorphic adenomas are the most frequent forms of epithelial benign salivary gland tumors which could evolve into malignant form. Hence, an early detection and accurate management of parotid tumors is critical for disease management and therapy planning [4].

Generally, local parotidectomy is performed for benign lesions and total parotidectomy is applied for malignant tumors [5]. Due to sample size limitation and existing histopathological inhomogeneity within these tumors, the results of Fine-Needle Aspiration (FNAC) Cytology, as the current standard of parotid tumor diagnosis, are not conclusive [6]. Furthermore, by total parotidectomy of malignant lesions, the risk of damaging the facial nerves increases [7]. Hence, the proximity of parotid glands to facial nerves and their inherent heterogeneity reveal the important role of imaging in differentiation and therapy/surgical planning of these tumors. Nowadays, multi-planar imaging is the most common method for the evaluation of the parotid tumors [4]. In this context, Magnetic Resonance Imaging (MRI) is a useful tool in the detection and classification of salivary gland tumors. Conventional MRI provides morphological information about size, location, and margin of the tumor, while Diffusion-Weighted Imaging (DWI), as a functional MRI modality, and the extracted Apparent Diffusion Coefficient (ADC) map could complement the morphological information by capturing physiological properties of the tumorous region [8, 9]. There is a close relationship between the tumor cellularity and the amount of restricted water molecules [10] which can be detected by DWI and ADC in such a way that areas with decreased ADC values represent high cellularity [11, 12].

DWI predicts the tumor cellularity and physiological abnormalities in initial stages [13]. In recent years, Diffusion Weighted Imaging (DWI) and Apparent Diffusion Coefficient (ADC) have emerged as effective tools for providing biological and physiological information about a variety of tumors including parotid lesions [14, 15]. The proliferating tumor cells are enriched in water, and DWI depends on microscopic movement of water molecules, the so-called Brownian motion. Due to the high density of tumor cells, the water mobility is restricted and water molecules cannot move freely in the extracellular space. Therefore, tumorous lesions appear as regions with higher signal intensities on DWI [16]. The high sensitivity of DWI to the changes within the cellular environment of the tissue and the histopathological characteristics of the tissue, makes this imaging modality and the quantitative ADC values potent biomarkers for differentiating the tumor types [4, 17].

Some of large benign parotid tumors manifest with heterogeneous appearance on T1-weighted images which can arise a high degree of uncertainty in diagnosis. Therefore, incorporating information provided by DWI is beneficial for an accurate diagnosis of such tumors [14, 18, 19]. Although DWI and ADC are now indispensable tools for an accurate diagnosis of a large variety of tumors, such as glioma brain tumors, there have been few attempts in parotid tumor diagnosis context for evaluating the role of quantitative ADC measurements for differentiating benign and malignant tumors [4, 20].

Eida et al. evaluated 31 patients with salivary gland tumors employing ADC-maps. They found that the region with high ADC values could be predictive of benignity whereas malignant tumors show low or extremely low ADC values. Also based on their results, high ADC in 5% or smaller area of tumor region, can be predictive of malignancy with 97% accuracy [18]. Wang et al. showed that ADC values in malignant head and neck tumors are significantly lower than benign lesions. They also identified malignant tumors with sensitivity and specificity of 84% and 91%, respectively [21]. Yerli et al. carried out a research on the accuracy of T2-weighted MRI in combination with DWI and fine-needle aspiration cytology (FNAC) for the detection of 22 parotid tumors. They indicated that T2-weighted images could be helpful for diagnosing malignant tumors and adding DWI would not improve the accuracy but it provides some advantages for parotid tumor characterization [22].

In another study on 40 patients with parotid masses by Balcik et al., pleomorphic adenoma and other types of parotid tumors were differentiated with 94.7% sensitivity by mean-ADC cutoff value of 1.60×10⁻³ mm²/s [23].
Karaman et al. showed that mean-ADC was higher in pleomorphic adenomas in comparison with warthin tumors. For each of the three chosen b-values, there was a significant difference in ADC-value between pleomorphic adenomas and malignant tumors [24].

In the mentioned works [18, 21-24], ADC-mean has been explored for its discriminative capability majorly in salivary gland tumors containing parotid masses. Nonetheless, maximum or minimum ADC may outperform mean-ADC in the differentiation of parotid tumors. In this study, we aimed to explore the significance of mean, maximum, and minimum ADC parameters for diagnosing parotid tumors through classification methods. Furthermore, a simple and clinically-implementable computer-aided decision-tree classifier is proposed based on the prominent ADC measure for objective discrimination of parotid tumor types with a high accuracy.

2. Materials and Methods

2.1. Subjects

Institutional Review Board approval (IRB) was obtained for this prospective study and the patients were included only if they provided their informed consent for the study. Between September 2012 to May 2013 (12 months), 41 patients (19 males, and 22 females, age range, 13-77 years, mean age, 43.1) presented with parotid tumors, who were scheduled for surgery and post-operative histopathological assessment, were enrolled in this study for MR imaging.

2.2. MR Imaging Protocol

The patients underwent MR imaging on a 3T MR scanner (Siemens MAGNETOM Tim TRIO, Erlangen, Germany) using a head coil, two weeks before the surgery. The anatomical sequences included axial T1-weighted (T1-w) imaging with, TE/TR = 11/700ms, FOV = 200×200 mm², matrix size = 205×256, slice thickness = 4 mm, number of slices = 25, and axial T2-weighted (T2-w) imaging with TE/TR = 75/5000ms, FOV = 200×200 mm², matrix size = 307×384, slice thickness = 4 mm, number of slices = 25.

Diffusion-Weighted MR Imaging (DWI) was performed by applying 2D spin echo, single-shot echo-planar imaging with TE/TR = 93/7500ms, slice thickness = 3.6 mm, FOV = 170×200 mm², matrix size = 102×160, number of slices = 25, b-value = 50, 1000 mm²/s. ADC-maps were automatically generated from DW images using the two b-values on the scanner workstation.

2.3. Image Analysis

The whole solid parts of the tumors were delineated on the ADC-maps by an expert radiologist (L.A) with over a decade of experience in head and neck oncology imaging, based on pre- and post-contrast T1-w and T2-w morphological images. Cystic components do not enhance on post-contrast T1-w images and usually have low signal intensity on T1-w and high signal intensity on T2-w images. In both benign and malignant tumors, solid components may present with low to intermediate T1-w signal intensity and high signal on T2-w.

Regions-Of-Interest (ROIs) were selected in regions containing the most solid components of the tumor and also normal contralateral tissue on ADC-maps using ImageJ software (http://imagej.nih.gov/ij/). ADC-values within the tumorous ROI were normalized by the average of the ADC-values within normal contralateral ROI. This step is essential to account for variabilities among different scanners and for various imaging parameters; it is performed by normalizing ADC-value of each pixel within the tumor ROI to an average ADC calculated from contra-lateral normal tissue. Then, the three quantitative parameters, namely mean, maximum, and minimum parameters were derived from the tumor ROIs. Within a ROI, the average of ADC-values represents ADC-mean; the maximum of ADC-values represents ADC-max, and ADC-min is the lowest ADC-value.

2.4. Statistical Analysis

Statistical analysis was performed on the three calculated parameters using two-tailed Student’s t-test (MATLAB). To evaluate the significance of each parameter for differentiating benign and malignant tumors, P-values were calculated. The P-value of less than 0.05 was considered statistically significant.
2.5. Automatic Classification

Classification methods are useful techniques for modeling variations in a dataset by defining rules for assigning an observation to its most related category. Classifiers can be real-valued functions of features [25] and be constructed with measured features and serve as a classification rule [26]. The mathematical forms of these functions may be linear, quadratic or any other forms of kernel functions [19].

Here, we implemented two real-valued classification functions, namely Linear Discriminant Analysis (LDA), and Quadratic Discriminant Analysis (QDA). LDA method seeks for a linear transformation that returns the maximum ratio of between-class to within-class variance, to ensure the best possible discrimination. QDA is a generalized model of LDA, which looks for a quadratic transformation for discriminating the data classes. Both of these methods try to model the pattern of data by implementing simple mathematical functions, such as linear or quadratic.

The simpler classifier, i.e., LDA, was used to gain a general outlook about the dataset and the discriminative capability of each of the extracted parameters. To model the data pattern with a more complicated function possibly to attain a higher accuracy, a quadratic function was designed. By starting from a simple classifier and adding further complexity, we aimed to find the optimum classifier in terms of complexity and accuracy; i.e. the classifier with less complexity which maintains high accuracy.

Following the implementation of the classifiers, the parameters were investigated for their capability in differentiating benign and malignant parotid tumors, in terms of sensitivity, specificity, accuracy, and the area under the Receiver Operating Characteristic (ROC) curve. Classification techniques were implemented using MATLAB software Statistics and Machine Learning Toolbox.

3. Results

3.1. Histopathological Assessment

All the patients were scheduled for surgery, two weeks after MR imaging and post-operative histopathological assessments were performed for all the resected tumors. The pathological findings demonstrated that 32 patients had benign and 9 had malignant parotid tumors. The histology information of the parotid tumors is summarized in Table 1.

<table>
<thead>
<tr>
<th>subtype</th>
<th>Tumor type</th>
<th>No. of patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Warthin Tumor</td>
<td>Benign</td>
<td>6</td>
</tr>
<tr>
<td>Pleomorphic Adenoma</td>
<td>Benign</td>
<td>21</td>
</tr>
<tr>
<td>Lymphoepithelial Sialadenitis</td>
<td>Benign</td>
<td>1</td>
</tr>
<tr>
<td>Neurofibroma</td>
<td>Benign</td>
<td>1</td>
</tr>
<tr>
<td>Myoepithelioma</td>
<td>Benign</td>
<td>1</td>
</tr>
<tr>
<td>Epidermoid</td>
<td>Benign</td>
<td>1</td>
</tr>
<tr>
<td>Lymphangioma</td>
<td>Benign</td>
<td>1</td>
</tr>
<tr>
<td>Mucoepidermoid Carcinoma</td>
<td>Malignant</td>
<td>1</td>
</tr>
<tr>
<td>Ductal Adenocarcinoma</td>
<td>Malignant</td>
<td>1</td>
</tr>
<tr>
<td>Adenoid Cystic Carcinoma</td>
<td>Malignant</td>
<td>4</td>
</tr>
<tr>
<td>Acinic Cell Carcinoma</td>
<td>Malignant</td>
<td>1</td>
</tr>
<tr>
<td>Neuroendocrine</td>
<td>Malignant</td>
<td>1</td>
</tr>
<tr>
<td>Adenoca</td>
<td>Malignant</td>
<td>1</td>
</tr>
</tbody>
</table>
3.2. Image Analysis and Quantification

The pre-operative MR images of 41 patients were assessed by an expert radiologist, who identified ROI on the most solid portions of the parotid lesions in at least two slices of the ADC-maps. Figures 1-2 demonstrate the tumorous regions on T1-w and T2-w images of two patients with histopathologically-proven benign and malignant parotid tumors along with the corresponding ADC-maps overlaid by the tumor borders. Mean, maximum and minimum ADC values were computed within the ROIs. To examine the significance of each of mentioned parameters in discriminating benign from malignant parotid tumors, P-values were calculated. Prior to statistical analysis, a normality test was performed to evaluate the distribution of the ADC parameters. All of these parameters had a normal distribution with 0.01 meaningful value. As it can be observed from Table 2, statistically significant differences can be observed for ADC-Mean and ADC-Max parameters in discriminating benign from malignant parotid tumors (P<0.05). ADC-Max is significantly higher for benign (1.95±0.52×10^{-3} mm^2/s) than malignant tumors (1.31±0.36×10^{-3} mm^2/s). Furthermore, ADC values for benign tumors range from intermediate to high levels (Figure 1-B), while malignant tumors appear with intermediate to low ADC levels (Figure 2-B).

Table 2. The calculation of P-value and mean of each ADC-map derived parameters for benign and malignant parotid tumors.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean (Benign) (×10^{-3} mm^2/s)</th>
<th>Mean (Malignant) (×10^{-3} mm^2/s)</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADC-Mean</td>
<td>1.50±0.53</td>
<td>0.93±0.40</td>
<td>0.005</td>
</tr>
<tr>
<td>ADC-Max</td>
<td>1.95±0.52</td>
<td>1.31±0.36</td>
<td>0.001</td>
</tr>
<tr>
<td>ADC-Min</td>
<td>0.98±0.60</td>
<td>0.64±0.49</td>
<td>0.130</td>
</tr>
</tbody>
</table>
For further investigation, box-and-whisker plots of the ADC-derived parameters are provided in Figure 3 to show the distribution of datasets, mean values and the amount of overlap between groups. As it is apparent in Figure 3-C, for ADC-Min parameter, benign and malignant groups have overlaps across the boxes, but for the other two parameters, i.e. ADC-Mean and ADC-Max, benign and malignant groups are well differentiated.

Figure 3. Box-and-whisker plots of ADC-Mean (A), ADC-Max (B) and ADC-Min (C) showing the data distribution and mean value of each parameters for benign and malignant tumors.

The sensitivity, specificity, accuracy, and area under the ROC curve (AUC) of each ADC-derived parameter, using LDA and QDA classifiers are indicated in Table 3. According to the statistical analysis (Table 3), ADC-Max and ADC-Mean have the same sensitivity (78%) and specificity (84%). Therefore, ADC-Max and ADC-Mean have the same accuracy (83%) which is more than accuracy of ADC-Min (78%). However, the ADC-Max can provide a higher AUC than ADC-Mean (83% versus 78%). As mentioned earlier, according to the statistical analysis, ADC-Min was not a significant parameter (P=0.13) for the differentiation of parotid tumors. Also, regarding the classification outcome, it can be observed that ADC-Min with 0% sensitivity is not capable to separate any of the malignant tumors correctly (Table 3).

Table 3. The calculation of sensitivity, specificity, accuracy and Receive Operating Characteristic (ROC) Analysis of each ADC-derived parameters in discriminating benign and malignant parotid tumors for LDA and QDA.

<table>
<thead>
<tr>
<th>ADC Parameter</th>
<th>LDA</th>
<th>QDA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SEN</td>
<td>SPE</td>
</tr>
<tr>
<td>ADC Mean</td>
<td>78</td>
<td>84</td>
</tr>
<tr>
<td>ADC Max</td>
<td>78</td>
<td>84</td>
</tr>
<tr>
<td>ADC Min</td>
<td>0</td>
<td>100</td>
</tr>
</tbody>
</table>

Abbreviations: SEN: Sensitivity, SPE: Specificity, Acc: Accuracy, AUC: Area under the ROC curve.
It is apparent from the results of Table 3 that by using more complicated classifiers, the performance of the parameters cannot be improved. For instance, ADC-Max using LDA classifier indicates the same sensitivity (78%), specificity (84%), accuracy (83%), and AUC (83%) as ADC-Max using QDA. Similar observations hold for the other ADC-derived parameters. Therefore, the best performance of ADC parameters could be achieved by a simpler classifier like LDA. LDA classification employing ADC-Max appears to be the optimum classifier, in terms of less complexity and high classification accuracy, for discriminating benign and malignant parotid tumors.

ROC curves for LDA methods are illustrated in Figure 4 for all of the three ADC-derived features. According to the ROC curves for LDA, ADC-Max<1.48×10^{-3} mm^2/s had a sensitivity of 78% and a specificity of 84% for predicting malignant tumors. Based on ROC curve analysis of ADC-Max, the two misclassified malignant tumors were adenoid cystic carcinoma lesions. ADC-Min has the lowest sensitivity among the three parameters.

**Figure 4.** Receiver Operating Characteristic curves (ROC) for LDA related to ADC-Mean, ADC-Max and ADC-Min and showing the highest sensitivity for ADC-Max and ADC-Mean.

### 3.3. Parotid Tumor Classifier

Using the mentioned classification scheme, we propose a simple and clinically-implementable prediction formula for differentiation of the benign and malignant parotid lesions:

\[
\text{output} = \text{ADC}_{\text{max}} - 1.48
\]

- \( \text{output} > 0 \): benign
- \( \text{output} < 0 \): malignant

Equation (1)

This prediction formula with an accuracy of 83%, is purely based on ADC-maps and can be exploited prior to the surgery for decision making. Based on the proposed formula, the positive output value indicates a benign tumor and the negative output value presents a malignant tumor. It should be noted that the ADC-Max values must be scaled by 10^{-3}. The examples of such classification scheme for the two patients with benign and malignant tumors are presented in part (D) of Figures 1-2.

### 4. Discussion

In this work, we aimed to investigate the significance of ADC-derived measures in the differentiation of benign and malignant parotid tumors prior to the surgery, and to propose a decision-tree classifier for aiding the diagnosis of radiologists. In the presented experiment, it is indicated that among the three assessed parameters, i.e. ADC-mean, ADC-max, and ADC-min, ADC-max parameter is helpful in distinguishing benign from malignant parotid tumors.

Due to a restriction in the diffusion of water molecules in malignant tumors, these lesions exhibit higher signals on DWI and lower ADC values. On this basis, the mean and maximum ADC values are higher in benign than malignant tumors and demonstrated statistically significant difference among the two groups (P<0.05).
further evaluated the performance of each of the mentioned parameters for distinguishing the tumor types using two different classifiers, namely Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA). Employing these classifiers, the maximum ADC value was shown to have the most AUC among the other parameters (83%), while ADC-mean presents a similar sensitivity (84%) and accuracy (83%) to ADC-max. A specificity of 100% was achieved by ADC-min but this parameter returned a very low sensitivity. Therefore, ADC-max is the most relevant quantitative parameter for separating benign from malignant parotid tumors.

In our work, the investigation of the two implemented classifiers, i.e. LDA and QDA, revealed that the high sensitivity and specificity of the calculated measures could compensate for the complexity of the pattern classification technique. Using LDA classifier, the inherent potential of the parameters became more evident, as we have not tried to fit any complicated hyperplanes or non-linear equations to discriminate the two tumor groups. In many applications, using more complicated classification model can remarkably improve the classification outcome because instead of fitting a line, a non-liner model passes between the two data groups and therefore, it can encompass several data points being incorrectly classified using a linear model [27]. However, estimating more potent features or parameters can improve the classification, as we have indicated in the current study. This can be observed from the lack of improvement of the accuracy of ADC-derived measures following QDA classification. Therefore, the classification scheme exploiting LDA model based on ADC-Max parameter can optimally distinguish the benign from malignant parotid tumors. On this basis, we have proposed a linear classifier that can be reliably and easily applied in clinical diagnosis for predicting parotid tumor malignancy with an accuracy of 83%.

An early detection of benign tumors can be helpful in an effective management and accurate decision making about the optimum treatment strategy for these patients, which could prevent complications arising from the surgical procedures and progression of the tumor into a malignant form, tumor recurrence or metastasis. Nonetheless, an accurate classification of parotid tumors employing quantitative DW-MRI and exploring the role of this imaging technique has not yet been profoundly investigated. More specifically, no data are available on the relevance of ADC-Max for characterizing parotid tumors. This study shows the contribution of quantitative ADC-Max measure in this setting. Specifically, this parameter is beneficial as large benign parotid tumors exhibit heterogeneity across their area. In our work, an ADC-Max under $1.48 \times 10^{-3} \text{mm}^2/\text{s}$ is highly suggestive of malignancy. Accordingly, due to the heterogeneity of malignant lesions, ADC-Min showed no false positives and could detect all benign lesions correctly. However, this parameter had no true positive detections resulting in 0 sensitivity. This shows that due to the heterogeneity of malignant tumors, regions with high and low ADC-values coexist across the tumorous region. As benign tumors also contain low ADC values, unlike ADC-Max, ADC-Min cannot be discriminative of benign and malignant parotid masses.

As DWI acquisition outside of the brain is usually susceptible to a lower quality than conventional MRI due to susceptibility artifact and has lower SNR, the authors have carefully checked primary considerations to achieve the required DW image quality by accurate shimming, in order to increase field homogeneity and by choosing proper EPI parameters to avoid susceptibility and motion artifacts [17]. Nonetheless, some limitations of this study must be addressed. Firstly, the small patient population for each histologic subtype may have biased the results of classification, which may hinder the generalization of the proposed approach to other protocols, scanners, and centers. For instance, the two misclassified malignant lesions were adenoid cystic carcinoma, while the remaining adenoid cystic carcinoma tumors were correctly categorized as malignant tumors. Therefore, by increasing the number of patients, more generalized parameter values would be calculated and decisive conclusions about the diagnosis of the presented tumor can be achieved. Secondly, the proposed ADC measurement is operator-dependent for ROI selection and relies on the expertise of the reader for identifying the most solid part of the tumor. The measurement may be inversely affected if noisy or cystic pixels...
are also incorporated in the analysis. In future works, automatic ROI selection techniques with less dependency on the subjective opinion will be implemented. Finally, the sensitivity of the ADC-derived measures could not exceed 78% which is suggestive of the incapability of using DWI and ADC exclusively for diagnosis of parotid tumors, which could be overcome by integrating the information provided by other MRI modalities.

5. Conclusions

In conclusion, these results are supportive of employing a quantitative ADC for characterizing parotid tumors. ADC-Max was the most specific and accurate ADC-derived quantitative feature for distinguishing benign from malignant parotid tumors. Here, a simple and accurate classifier was introduced that could be feasibly applied in clinical practice.

References


