Differentiation between benign and malignant thyroid nodules using diffusion-weighted imaging, a 3-T MRI study

Leila Aghaghazvini, Hashem Sharifian¹, Nasrin Yazdani², Melina Hosseiny, Saina Kooraki, Pirouz Pirouzi, Afsoon Ghadiri, Madjid Shakiba³, Soheil Kooraki

Department of Radiology, Shariati Hospital, ¹Department of Radiology, Amir Alam Hospital, Tehran University of Medical Sciences, ²Otorhinolaryngology Research Center, Tehran University of Medical Sciences, ³Advanced Diagnostic and Interventional Radiologic Research Center (ADIR), Imam Khomeini Hospital, Tehran University of Medical Sciences, Tehran, Iran

Correspondence: Dr. Soheil Kooraki, Department of Radiology, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran. E-mail: Soheil.kooraki@gmail.com

Abstract

Background: Preoperative differentiation of benign from malignant thyroid nodules remains a challenge. Aims: This study assessed the accuracy of diffusion-weighted imaging (DWI) for differentiation between benign and malignant thyroid nodules.

Materials and Methods: Preoperative DWI was performed in patients with thyroid nodule by means of a 3-T scanner magnetic resonance imaging (MRI). Images were obtained at b value of 50, 500, and 1000 mm²/s to draw an ADC (apparent diffusion coefficient) map. Findings were compared with postoperative histopathologic results. Receiver operating characteristic curve was used to assess the accuracy of different cutoff points. Results: Forty-one thyroid nodules (26 benign and 15 malignant) were included in this study. None of static MRI parameters such as signal intensity, heterogeneity, and nodule border was useful to discriminate between benign and malignant lesions. Mean ADC value was $(1.94 \pm 0.54) \times 10^{-3} \text{mm}^2/\text{s}$ and $(0.89 \pm 0.29) \times 10^{-3} \text{mm}^2/\text{s}$ in benign and malignant nodules, respectively ($P$-value < 0.005). ADC value cutoff of $1 \times 10^{-3} \text{mm}^2/\text{s}$ yielded an accuracy, sensitivity, and specificity of 93%, 87%, and 96% to discriminate benign and malignant nodules. Conclusion: DWI is highly accurate for discrimination between benign and malignant thyroid nodules.

Key words: Diffusion; nodule; thyroid

Introduction

With 42% frequency in ultrasound examination, thyroid nodules play a big challenge for both radiologists and clinicians. Most of these nodules are benign; however, up to 7% will turn out to be malignant, requiring proper management.[1] According to American Thyroid Association guidelines, ultrasound is the imaging modality of choice for assessment and risk stratification of thyroid nodules.[2] Various sonographic features are used for characterization
of thyroid nodules and for deciding if fine-needle aspiration (FNA) is required or not. For further evaluation of suspicious nodules, FNA is considered the diagnostic method of choice.\textsuperscript{[3]} FNA is an invasive procedure and highly inaccurate in nodules larger than 4 cm.\textsuperscript{[4]}

Thyroid nodules are frequently discovered in computed tomography or magnetic resonance imaging (MRI); however, these modalities are not reliable for characterization of thyroid nodules.\textsuperscript{[5]} Ultrasound is usually indicated in these incidentally found nodules. Various studies have assessed the power of diffusion-weighted imaging (DWI) and its quantitative counterpart apparent diffusion correction (ADC) for differentiation between benign and malignant thyroid nodules. The results of these studies have been promising but wide variability has been encountered.\textsuperscript{[6-11]}

This study was designed to assess the accuracy of DWI for differentiation between benign and malignant thyroid nodules using a 3-T MRI.

**Materials and Methods**

This cross-sectional study was conducted from April 2012 to March 2014. Patients with thyroid nodules in which surgery plan was already established were referred from an academic otolaryngology clinic. Surgery plan was based on the findings of clinical symptoms and thyroid ultrasound or the results of FNA. MRI was carried out in all patients with focus on DWI. Individuals who could not tolerate the exam (because of claustrophobia, etc.) were excluded. Patients with contraindicated conditions for MRI (e.g. pacemaker) were not included as well. Overall, 30 patients were examined by MRI; among them, 4 patients were excluded because of unsatisfactory quality of DWI. Surgery was carried out in all participants within 7–14 days after MRI acquisition as total thyroidectomy, lobectomy, and lymph node excision. Then, a single expert pathologist evaluated histopathologic specimens.

The institutional review board and local ethics committee at the university approved this study. After explaining the study process in detail to the patients, written informed consent was obtained from all participants.

Magnetic resonance imaging technique and interpretation

MRI was performed with a 3-T superconductive magnet (Siemens, MAGNETOM Trio, Germany) using a dedicated neck coil. Axial turbo spin-echo T2-weighted (TR/TE: 2500/75, matrix size of 307 × 384, FOV: 220 mm, and slice thickness = 4 mm with intersection gap = 1 mm), coronal turbo spin-echo T2-weighted (TR/TE: 3000/69, matrix size of 307 × 384, FOV: 240 mm, and slice thickness = 5 mm with intersection gap = 1 mm), and axial T1-weighted (TR/TE: 746/11, matrix size of 307 × 384, FOV: 220 mm, and slice thickness = 4 mm with intersection gap = 1 mm) images were obtained.

DWI was obtained in axial plane by single shot spin-echo echo-planar imaging on b values of 50, 500, and 1000 (TE: 70, TR: 5400, slice thickness: 4 mm with no intersection gap, and the number of signal acquisition: 4). The total scan time was 4 min and 54 s. After acquiring images, regions of interest were drawn in every suspected nodule on ADC maps, consisting of an area of 5–10 mm\textsuperscript{2}, not consisting areas with necrosis, calcification, and cystic components.

Two expert radiologists with more than 5 years of practice in head and neck imaging interpreted the acquired images in consensus. Interpreting radiologists were completely blind to the clinical data of the patients.

Statistical analysis was performed using SPSS software (version 16.0, SPSS Inc. Chicago, IL, USA). Receiver operating characteristic (ROC) curve was carried out to measure sensitivity, specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV) in various cutoff values. All data are expressed as mean ± SD and an alpha of 0.05 was used as the cutoff for significance.

**Results**

Forty-one thyroid nodules in 26 patients (19 females and 7 males, mean age of 38 ± 12) were included in this study. Mean nodule size was 34 ± 11.22 mm ranging from 16 to 63 mm. According to surgery specimen, 26 nodules were proved to be benign, whereas 15 nodules were malignant. Table 1 provides frequency of each entity in detail. There was not any significant difference in age between benign and malignant nodules (37 ± 12 vs. 40 ± 12, P value: 0.16).

None of static MRI parameters such as signal intensity, heterogeneity, and nodule border was useful to discriminate between benign and malignant lesions (P-value > 0.05).

Mean ADC value was (1.94 ± 0.54) × 10\textsuperscript{-3} mm\textsuperscript{2}/s and (0.89 ± 0.29) × 10\textsuperscript{-3} mm\textsuperscript{2}/s in benign and malignant nodules, respectively (P-value < 0.005). ROC curve [Figure 1] showed ADC cutoff of 1 × 10\textsuperscript{-3} mm\textsuperscript{2}/s to result in accuracy

<table>
<thead>
<tr>
<th>Pathology</th>
<th>Frequency</th>
<th>Percent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multinodular goiter</td>
<td>11</td>
<td>26.8</td>
</tr>
<tr>
<td>Nodular goiter</td>
<td>5</td>
<td>12.2</td>
</tr>
<tr>
<td>MNG with lymphocytic thyroiditis</td>
<td>1</td>
<td>2.4</td>
</tr>
<tr>
<td>Nodular hyperplasia with cystic degeneration</td>
<td>1</td>
<td>2.4</td>
</tr>
<tr>
<td>Colloid nodule</td>
<td>5</td>
<td>12.2</td>
</tr>
<tr>
<td>Colloid cyst</td>
<td>3</td>
<td>7.3</td>
</tr>
<tr>
<td>Papillary thyroid carcinoma</td>
<td>15</td>
<td>36.6</td>
</tr>
</tbody>
</table>

MNG = Multinodular goiter
of 93% for discrimination between benign and malignant nodules (sensitivity: 87%, specificity: 96%, and PPV and NPV: 93%) [Table 2, Figures 2 and 3]. With ADC cutoff of $1.8 \times 10^{-3} \text{mm}^2/\text{s}$, sensitivity of 100% was achieved, whereas in cutoff of $0.72 \times 10^{-3} \text{mm}^2/\text{s}$, specificity of 100% was seen.

**Discussion**

DWI has emerged as a noninvasive and complimentary tool in the assessment of thyroid nodules. DWI is a measure of random Brownian motion of free water molecules in a given environment. Most commonly, DWI is performed by adding two motion-sensitive gradient pulses to either side of a 180 refocusing pulse in a T2-spin-echo sequence. The first gradient creates a phase shift, which is controlled by the gradient strength at the position of each spin. Then during the second gradient pulse, nonmobile (restricted) spins will experience the same gradient strength and will rephase to their stationary positions, but mobile spins will experience a different gradient strength and will go on a total phase shift, not returning to their initial locations. This will cause signal loss for mobile spins. To obtain highly sensitive diffusion-weighted images, the amplitude and duration of two gradients should be increased. These gradient features are generally known as $b$ value. After obtaining diffusion images, ADC map is drawn by deducting signal of diffusion images with two different $b$ values. The main problem with obtaining DWI in head and neck tissues is the motion artifact, caused by nearby moving organs. Breathing, swallowing, coughing, and jaw movements can greatly affect the quality of the images. Moreover, diffusion images are prone to susceptibility artifact in airborne interfaces. To reduce these artifacts and increase signal-to-noise ratio, patients were asked not to swallow or cough during examination, and in some patients, head was immobilized to prevent involuntary movements.

The present study showed that DWI is highly valuable in differentiation between benign and malignant thyroid nodules. DWI can lower the burden of unnecessary surgery in cases with inconclusive FNA. A number of studies have evaluated the role of DWI for assessment of thyroid nodules. Erdem et al. found significantly increased ADC value in benign and significantly decreased ADC value in malignant thyroid nodules compared to the normal

---

**Table 2: Sensitivity, specificity, efficacy, predictive values, and likelihood ratios of ADC with a cutoff of 1 mm²/s**

<table>
<thead>
<tr>
<th>Cutoff: ADC≤1 mm²/s</th>
<th>TP=14/FN=1/TN=22/FP=4</th>
<th>Sensitivity=0.87 (95% CI: 0.60–0.98)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specificity=0.96 (95% CI: 0.80–0.99)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficiency (correct classification rate) = 0.93 (95% CI: 0.80–0.98)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predictive value of positive test = 0.93 (95% CI: 0.66–0.99)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predictive value of negative test = 0.93 (95% CI: 0.76–0.89)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likelihood ratio of positive test = 22.53 (95% CI: 3.26–155.57)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Likelihood ratio of negative test = 7.21 (95% CI: 1.98–26.26)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CI: Confidence interval, FN: false negativity, TN: True negativity, FP: False positivity, TP: True positivity

---

**Figure 1:** With an ADC cutoff of $1 \times 10^{-3} \text{mm}^2/\text{s}$, area under the curve of 93% was achieved for differentiation between benign and malignant nodules.

**Figure 2 (A-C):** A 25-year-old female was presented with a palpable firm nodule in the right lobe of thyroid. The nodule was indeterminate on ultrasound exam. T2-weighted sequence (A) shows a moderately hyperintense nodule in the right lobe of thyroid. Diffusion weighted imaging at $b$ value of 1000 (B) and corresponding ADC (C) show a mean ADC value of $0.84 \pm 0.09 \times 10^{-3} \text{mm}^2/\text{s}$. Later, pathology revealed the diagnosis of papillary thyroid carcinoma.

**Figure 3 (A-C):** A 50-year-old male was referred with enlarged thyroid. Ultrasound showed several thyroid nodules. T2-weighted sequence (A) shows a hyperintense well-defined nodule. Diffusion weighted imaging at $b$ value of 1000 (B) and corresponding ADC (C) show mean ADC of $2.21 \pm 0.09 \times 10^{-3} \text{mm}^2/\text{s}$. Pathology showed multinodular goiter without evidence of malignancy.
The study was subject to several limitations, which should be considered when interpreting the results. First, the number of malignant nodules was low and the results of this study might be confirmed in those with higher number of malignant nodules. Moreover, other than papillary thyroid carcinoma, we did not encounter any other types of malignant nodule in this study. Second, nodules measuring below 10 mm could not be evaluated with DWI. Third, the ultrasound and FNA data were not available in all patients, and therefore, we were not able to carry out any correlation between findings of ultrasound and DWI. Fourth, we did not evaluate the whole lesion ADC value and only evaluated ADC value of the lesion excluding cystic and necrotic components. Finally, the presence of motion artifact made the interpretation difficult in some individuals.

In conclusion, the present study confirms the results of previous studies by showing that DWI is highly valuable in differentiation of benign and malignant thyroid nodules. Using $b$ factor of 1000, a cutoff of $1 \times 10^{-3} \text{mm}^2/\text{s}$ for ADC value resulted in accuracy, sensitivity, and specificity of 93%, 87%, and 96%, respectively. Studies with larger number of thyroid nodules or meta-analysis are recommended for validation of the results.

Financial support and sponsorship
Nil.

Conflicts of interest
There are no conflicts of interest.

References

11. Mutlu H, Sivrioglu AK, Sonmez G, Velioglu M, Sildiroglu HO,


