Can Quantitative Diffusion-Weighted MR Imaging Differentiate Benign and Malignant Cold Thyroid Nodules? Initial Results in 25 Patients


AJNR Am J Neuroradiol 2009, 30 (2) 417-422
doi: https://doi.org/10.3174/ajnr.A1338
http://www.ajnr.org/content/30/2/417
Can Quantitative Diffusion-Weighted MR Imaging Differentiate Benign and Malignant Cold Thyroid Nodules? Initial Results in 25 Patients

BACKGROUND AND PURPOSE: The characterization of cold nodules of the thyroid gland is mandatory because approximately 20% of these nodules are of malignant origin. The purpose of this study was to evaluate the distinction of cold thyroid nodules by using quantitative diffusion-weighted MR imaging (DWI).

MATERIALS AND METHODS: In 25 patients with cold nodules on scintigraphy and suggestive findings at fine-needle aspiration, thyroid carcinoma was suggested. In these patients, cold nodules and the normal parenchyma of the contralateral thyroid lobe were prospectively investigated with quantitative DWI (echo-planar imaging sequence; maximum b-value, 800 s/mm²) before surgery. The differences in the mean apparent diffusion coefficient (ADC) values in benign and malignant nodules were tested by using a Mann-Whitney U test.

RESULTS: Histologically, there were 20 carcinomas with a minimum size of 8 mm and 5 adenomas. The mean ADC values (in 10⁻³ mm²/s) differed significantly among carcinoma, adenoma, and normal parenchyma (P < .05). The ranges (95% confidence interval) of the ADC values for carcinoma (2.43–3.037), adenoma (1.626–2.233), and normal parenchyma (1.253–1.602) showed no overlap. When an ADC value of 2.25 or higher was used for predicting malignancy, the highest accuracy of 88%, with 85% sensitivity and 100% specificity, was obtained.

CONCLUSIONS: Quantitative DWI seems to be a feasible tool with which to differentiate thyroid carcinomas from adenomas; however, further studies are required including larger numbers of patients to confirm our results.

The diagnostic approach to the evaluation of the thyroid is primarily based on clinical evaluation, palpation, and determination of the levels of thyroid hormones.¹

Diagnostic imaging methods, such as radionuclide scintigraphy or sonography, are commonly used to facilitate the differentiation of benign and malignant changes in the thyroid parenchyma. Some indications for surgical intervention are the presence of a hypofunctional, or so-called “cold” nodule; a diagnosis of cancer on fine-needle aspiration (FNAC); or the presence of a large thyroid lesion that causes symptoms such as hoarseness or dysphagia.²,³ Cold nodules are characterized by a reduced uptake of the radionuclide tracer. However, further tissue characterization of these nodules is required because approximately 20% are of malignant origin.³ Sonography-guided FNAC (USgFNAC) is a worldwide established technique that provides further evaluation of the nodules on a cytologic level.⁴,⁵ However, depending on many factors, the accuracy of USgFNAC can be low. The expertise of the cytopathologist plays a crucial role; and in regions with lack of iodine, where multiple nodules in the thyroid lobe can be present, it might be difficult to obtain the cytology of the suggested nodule. Indeterminate cytologic findings demand additional diagnostic methods and, if necessary, surgery.⁶ In addition, USgFNAC is an invasive procedure and uncomfortable for the patient. Thus, there is a demand for both noninvasive and more accurate diagnostic methods for further evaluation of cold thyroid nodules.

To date, little information is available about the use of MR imaging in the diagnosis of thyroid cancer.⁷⁻¹⁰ Diffusion-weighted MR imaging (DWI) is a noninvasive technique with no radiation exposure, which has the potential to differentiate benign from malignant tissues.¹¹ During the last 2 decades, DWI has evolved as a helpful diagnostic tool for assessing in vivo tumor characterization, not only in neural lesions but also in extraneural tissue, such as bone marrow pathologies, lymph nodes, and liver tumors.¹²⁻¹⁷ Structural changes characteristic of malignancies or benign tissue may result in different signals on DWI, which may be quantified by calculating the apparent diffusion coefficient (ADC). The ADC is an objective parameter that reflects the tissue-specific diffusion capacity and is already being used for tissue characterization and follow-up measurements in therapeutic monitoring.¹⁸⁻²⁰ To our knowledge, this study is the first to evaluate cold thyroid nodules that include a high number of malignant cold nodules.

The purpose of this prospectively designed study was to evaluate whether quantitative DWI could distinguish benign from malignant cold thyroid nodules.

Materials and Methods

Patients

During 18 months, 31 patients (20 women and 11 men; median age, 57.7 years; range, 25–82 years; median body mass index, 28.22; range, 18.04–41.80) were included in this prospectively designed study. The thyroid glands of the 31 patients scheduled for total thyroidectomy were investigated with MR imaging, including DWI, the day before
surgery. Patients were scheduled for thyroid surgery because of cold thyroid nodules diagnosed by radionuclide scintigraphy by using technetium Tc99m sodium pertechnetate; suggested thyroid carcinoma or follicular neoplasia, diagnosed by USgFNAC; and suggestive sonography of the thyroid nodules. No instrumentation or biopsy was performed within 3 weeks before MR imaging examination to avoid distortion of DWI measurements due to hemorrhage and susceptibility artifacts. Only cold thyroid nodules with a size >8 mm on sonography and MR imaging (minimum size, 90 mm² on sagittal DWI images) and with suggestive or inconclusive findings on USgFNAC were included in this study.

Six patients were excluded from the study due to motion artifacts and poor image quality in 3 patients and a solitary nodule with a size <8 mm in 3 patients. The nodules with a size <8 mm were excluded from further analysis because these could not be evaluated with DWI due to the low spatial resolution of 1T MR imaging. Finally, 25 patients were selected for further evaluation. The study was performed with institutional review board approval. All patients included in the study gave written informed consent before imaging after a complete explanation of MR imaging.

At surgery, 5 patients had adenoma, 10 had papillary thyroid carcinoma (PTC), 6 had medullary thyroid carcinoma (MTC), and 4 had follicular thyroid carcinoma (FTC).

MR Imaging
Both conventional MR imaging and DWI were performed on a 1T MR imaging unit (Gyroscan T10-NT; Philips Medical Systems, Best, the Netherlands) with a dedicated head and neck coil (Philips Medical Systems). The conventional MR imaging protocol for this study consisted of the following sequences: T1-weighted spin-echo imaging (TR/TE, 550/20 ms) in an axial orientation with a matrix of 512 × 512 pixels, T2-weighted fast spin-echo imaging (TR/TE, 3500/100 ms) in a sagittal orientation with a matrix of 512 × 512 pixels, and short tau inversion recovery imaging (TR/TE/T1, 2100/60/100 ms) in the coronal orientation with a matrix of 256 × 256 pixels.

DWI was performed in exactly the same sagittal plane and orientation that were used in the routine sequences showing the suggestive lesion in the thyroid gland. The exact localization of the sagittal DWI plane was correlated with the midline of the lesion detected on the axial images. In case of multiple nodules in the thyroid gland, the most suggestive and largest lesion was depicted with a combination of sonography, USgFNAC, and scintigraphic findings regarding the location of the nodule. Additional DWI sequences, by using the same parameters, were obtained of the contralateral thyroid lobe, which did not present any evidence of pathology, either on MR imaging or scintigraphy. For DWI, a navigated interleaved echo-planar imaging (EPI) sequence with fat saturation (TR/TE, 2500/102 ms; 13 gradient echoes per shot; 2 signals acquired; FOV, 250 × 250 mm; section thickness, 6 mm; intersection gap, 1.0 mm; and matrix, 128 × 128 pixels, which was zero-filled and reconstructed to 256 × 256 pixels) was performed.21

Diffusion-weighted gradients were applied in 3 orthogonal directions. The signal-intensity-to-noise ratio (SNR) was increased by obtaining isotropic DWIs from the geometric average derived from these images. The b-values used in this study were determined by the imager gradient system (maximum, 10 mT/m), which resulted in a maximal b-value of 800 s/mm² ($b_1$).

For quantitative analysis of tissue-specific diffusion capacities, the ADC was calculated according to the following equation: \[ \text{ADC} = \log(\text{SI}_1 / \text{SI}_2) / (b_2 - b_1), \] where $\text{SI}_1$ and $\text{SI}_2$ are the signal intensities (SI) of DWIs obtained with 2 different b-values ($b_1 = 0$ s/mm² and $b_2$, respectively) in the user-defined regions of interest.20 The region of interest was placed in the pathologic lesion on DWI by an experienced head and neck radiologist, who matched the SI changes on T1- and T2-weighted turbo spin-echo images. Cystic parts of the nodule were excluded from the region of interest to avoid a falsely elevated ADC value. Cystic areas of the lesion were defined by characteristic MR imaging SIs on T1- and T2-weighted images and were correlated with sonographic findings. In the contralateral nonpathologic thyroid lobe, regions of interest corresponding in size to the lesion were drawn and the ADC values were calculated. The difference between the ADC values of the pathologic lesion and the normal thyroid parenchyma and the difference between the various ADC values of the lesion in each patient were evaluated. Multiple measurements of the ADC values were performed to confirm the reproducibility of the values.

For objective image quality, the SNR of the thyroid lesion was calculated according to the following ratio: $\frac{\text{SI}}{\text{SD}_{\text{AIR}}}$ where $\text{SI}_{\text{ThyLes}}$ is the SI of the thyroid lesion of the DWI with a $b_2$-value, and $\text{SD}_{\text{AIR}}$ is the SD of the SI of the air representing the background noise. The SNR of the normal parenchyma in the contralateral lobe was evaluated according to the SNR of the thyroid lesion. ADC values were corrected for the background noise. Only those pixels that could be clearly differentiated from background noise were considered for calculation. The selection threshold was set above 2.5 times the SD of the background noise from the DWI with a $b_2$-value. In 6 patients, pixels within the lesion with a lower SI than 2.5 times the $\text{SD}_{\text{AIR}}$ were excluded.22

Tomographic examination was performed with a 21-gauge (0.8-mm) or 22-gauge (0.7 mm) needle in 21 patients. Part of the aspirated material was air-dried, methanol-fixed, and stained according to the May-Grunwald-Giemsa method. Another part of the aspirated material was fixed with a dedicated fixation spray to assure the preservation of the characteristic cell nuclei by a layer of polyglycol and Papanicolaou stain (Merckofix; Merck, Darmstadt, Germany).

The findings of USgFNAC were classified into 6 categories: inconclusive; benign; suggestive (follicular neoplasia); and papillary, medullary, and follicular thyroid carcinoma.

Thyroid Scintigraphic Examination
Radionuclide scintigraphy was performed by using technetium Tc99m sodium pertechnetate in all patients. After intravenous injection of 185 MBq of the tracer, emission thyroid scans were obtained by using a gamma camera and data analysis system (Nuclide TH33; Mediso Medical Imaging Systems, Budapest, Hungary) and a low-energy high-sensitivity parallel-hole collimator. Approximately 20 minutes after injection, imaging was started and the tracer uptake and the length of each thyroid lobe were measured. Scintigraphic findings were classified according to regular, reduced, or increased tracer uptake.

FNAC
USgFNAC of suggestive cold nodules was performed with a 21-gauge (0.8-mm) or 22-gauge (0.7 mm) needle in 21 patients. Part of the aspirated material was air-dried, methanol-fixed, and stained according to the May-Grunwald-Giemsa method. Another part of the aspirated material was fixed with a dedicated fixation spray to assure the preservation of the characteristic cell nuclei by a layer of polyglycol and Papanicolaou stain (Merckofix; Merck, Darmstadt, Germany).

The findings of USgFNAC were classified into 6 categories: inconclusive; benign; suggestive (follicular neoplasia); and papillary, medullary, and follicular thyroid carcinoma.
### Table: Mean ADC values and 95% CI of carcinoma, adenoma, and normal parenchyma

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>ADC of Lesion</th>
<th>ADC of Normal Parenchyma</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carcinoma (n = 20)</td>
<td>2.73 ± 0.65</td>
<td>1.439 ± 0.648</td>
</tr>
<tr>
<td>Adenoma (n = 5)</td>
<td>1.93 ± 0.25</td>
<td>1.341 ± 0.245</td>
</tr>
</tbody>
</table>

Note: ADC indicates apparent diffusion coefficient.

* ADC values are expressed as the [mean ± SD] × 10⁻³ mm²/s (P < .05).

** Sonography **

Two radiologists with at least 7 years’ experience in thyroid sonography examined all thyroid glands with gray-scale and power Doppler sonography. A sonography scanner (HDI 5000; Philips Medical Systems) was used with a 12-MHz linear transducer for gray-scale and power Doppler imaging. Thyroid lesions were classified as benign or suggestive of further evaluation according to the criteria reported in the literature. Histology that was obtained from all nodules served as the gold standard for evaluating the diagnostic value of MR imaging. Intraproactively, thyroid tissue was evaluated with an hematoxylin-eosin stain, and further histologic analysis of the tissue structure and cells was performed by a dedicated pathologist experienced in the thyroid gland. The histopathologic results were correlated with the ADC values.

** Statistical Methods **

Statistical analysis was performed by using a PC spreadsheet software package (Statistical Package for the Social Sciences, Version 14.0; SPSS, Chicago, Ill). For statistical analysis of differences in the ADC values in patients with benign and malignant thyroid nodules, a Mann-Whitney U test was used because of skewed data and heterogeneous variances. Differences in the ADC values between thyroid lesions and the normal thyroid gland parenchyma were tested by using paired t tests for each group separately. Confidence intervals (CI) of 95% were evaluated for ADC values. The threshold value of the mean ADC value for discrimination between benign and malignant thyroid nodules was determined; and sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy, and the 95% CIs were calculated. A P value < .05 was considered to indicate a statistically significant difference.

** Results **

The average size of the region of interest measured in the thyroid gland was 590.6 mm² (minimum, 76 mm²; maximum, 1142 mm²). The size of the investigated lesion measured on T1-weighted images varied from 90 to 1200 mm². The measured SNR of the DWI with a b value of 800 s/mm² was >12.

The mean ADC values calculated from the thyroid DWIs are summarized in the Table. The ADC values for thyroid cancer differed significantly from the ADC values of adenomas (P = .004) (Figs 1 and 2). There were no significant differences between the ADC values for the 3 types of carcinoma (P > .05). Comparisons between pathologic thyroid tissue, including adenoma and carcinoma, and nonpathologic thyroid tissue demonstrated that the ADC values of pathologic thyroid tissue were significantly higher than those of the reference tissue (normal thyroid gland parenchyma) (P < .009). The range of ADC values in thyroid carcinoma (95% CI, 2.43–3.037 × 10⁻³ mm²/s), adenoma (95% CI, 1.626–2.233 × 10⁻³ mm²/s), and normal thyroid parenchyma (95% CI, 1.253–1.602 × 10⁻³ mm²/s) showed no overlap (Fig 3).

An ADC value of 2.25 × 10⁻³ mm²/s or higher was predictive of carcinoma with an accuracy of 88%. Using this threshold, we correctly diagnosed carcinoma in 17 of 20 patients, with a sensitivity of 85% and a specificity of 100% (PPV, 100%; NPV, 65%).

** Discussion **

DWI is based on the random translational motion of water protons, which reflects the tissue-specific diffusion capacity. The diffusion capacity is indirectly proportional to diffusion barriers.24 Therefore, the motion of diffusion capacity is restricted and results in increased SIs and low ADC values.

The initial results of this pilot study indicate that DWI has the potential to enable the differentiation of thyroid carcinoma, adenoma, and normal parenchyma. The surprising finding is that in contrast to neural and extraneural malignancies, thyroid tumors present with low SI on DWI and high ADC, which indicate unhindered diffusion of water protons. Our results show rather low ADC values for the normal thyroid parenchyma (1.439 × 10⁻³ mm²/s), whereas an intermediate ADC value was observed for adenomas (1.929 × 10⁻³ mm²/s), and high ADC values were noted for thyroid carcinomas (2.733 × 10⁻³ mm²/s) (Figs 1–3, Table). Our data suggest that the differentiation of the different types of carcinoma, such as PTC, MTC, and FTC, is not possible with DWI. However, blood tests enable accurate diagnosis of MTC on the basis of elevated basal and pentagastrin-stimulated calcitonin levels, and follow-up of calcitonin levels in postoperative patients is state-of-the-art management.

Adenoma could be differentiated from carcinoma with an accuracy of 88% by using the ADC threshold value of 2.25 × 10⁻³ mm²/s.

Razek et al10 evaluated sonographically detected solitary thyroid nodules by DWI to discriminate between benign and malignant nodules. In contrast to our findings, Razek et al reported that malignant thyroid nodules showed low ADC values compared with benign nodules. Their possible explanation for high ADC values in benign nodules was the relative abundance of colloid follicles, microcystic necrosis, hemorrhage, fibrous tissue, and calcification. Calcified psammoma bodies in PTC and hyperplastic nuclei are responsible for low ADC values in malignant nodules. In our opinion, hemorrhage and microcystic necrosis occur in both adenoma and carcinoma. In addition, calcifications do occur in adenoma and, therefore, might cause restricted diffusion in adenoma as well as in carcinoma.27 Evaluating the ADC values of the normal parenchyma would have been of great value in the study of Razek et al10 to compare these ADC values with pathologic thyroid tissue and to provide an internal standard of reference. In addition, the number of malignant thyroid nodules in their study was low (n = 7), because every solitary nodule that was detected by sonography was included in their study. As performed in our study, the combined diagnostic approach of sonography and scintigraphy, which detects cold nodules, in-
creases the likelihood of discovering more malignancies in solitary thyroid nodules.

The interpretation of our findings is based on the following: If thyroglobulin is increasingly produced in the thyroid gland, an increase in the number and size of macrofollicles filled with colloid is observed. In thyroid tissue, the diffusion capacity is mainly influenced by the presence or absence and the size of macrofollicles. In the abnormal thyroid gland, the number and size of macrofollicles filled with colloid are elevated; thus, the motion of water protons is increased. This leads to an increased diffusion capacity and high ADC values.20 Our results are in accordance with the findings of Wang et al.28 In contrast to other malignant head and neck lesions, they found high ADC values in papillary thyroid carcinomas. Their possible explanation for the high ADC values in thyroid carcinomas included abundant extracellular fluid in the follicular portions.

In the normal thyroid parenchyma, there are only small cells filled with colloid, depending on the functional activity of the gland. The diffusion capacity in the normal thyroid parenchyma is, therefore, influenced by the high cell attenuation, which results in a restricted motion of water protons and low ADC values.

It is not possible to differentiate the various thyroid gland carcinomas by their specific cell attenuation. Therefore, it was no surprise that we could not differentiate the different types of carcinoma on the basis of their specific ADC values. The
major histopathologic differentiation between adenoma and carcinoma is determined by the invasion of the vessels. There is no completely satisfying explanation for why the ADC values in adenoma are lower than those in carcinoma, because both pathologies are characterized by an elevated number of macrofollicles. It might be possible that areas of bleeding, which are often found in carcinoma, result in an increased extracellular space and, therefore, in an increase in diffusion capacity and high ADC values for carcinoma. However, it remains to be explored whether these areas of bleeding cause susceptibility artifacts on MR imaging and might change the SI on DWI. Eggshell and coarse calcifications in benign thyroid nodules might be more responsible for low ADC values than microcalcifications in malignant nodules. However, the extent of calcifications can vary widely in benign and malignant nodules. Histopathologic correlation with DWI showed that the abundance of follicular portions in carcinoma (PTC, FTC) resulted in an increase of extracellular fluid and, therefore, in high ADC values in carcinoma.

In our study, a b-value of 800 s/mm² was used with good image quality, as is shown by the high SNR on DWIs. This higher level for the b-value was chosen to reduce the influence of perfusion on DWI.²⁹ Therefore, the ADC values, in fact, indicate the diffusion capacity in normal and abnormal thyroid tissue without being influenced by perfusion. Razek et al.¹⁰ used an EPI sequence of DWI with b-values of 250 and 500 s/mm².

The navigated EPI sequence we used in this study is superior to other DWI sequences because of its shorter acquisition time and, consequently, fewer artifacts.²¹ EPI methods have the disadvantage of sensitivity to resonance offsets, such as tissue susceptibility differences, magnetic field inhomogeneities, and chemical shift artifacts, as well as susceptibility to motion artifacts. The application of navigator echoes and the adaptation of the acquisition parameters can overcome these disadvantages, and sufficient image quality can be achieved.²¹,³⁰ A fat-saturation prepulse was used for the DWI sequences in this study to avoid severe chemical shift artifacts.²⁶

There were several limitations of this study. First, the small number of adenomas and the correlation of the ADCs of lesions with histopathologic findings were limited. The high prevalence of malignant lesions among cold thyroid nodules was attributable to the preselection of suggested cold nodules by USgFNAC findings, as described in the “Materials and Methods” section. Further investigations in a larger series that would include a larger number of benign lesions are necessary to clarify the correlation between the ADC values and histopathologic findings. Second, DWI is limited by the low diagnostic value for microcarcinomas due to the low spatial resolution. Histopathologically verified microcarcinomas, <8 mm, in adenoma or normal thyroid gland tissue cannot be detected by the DWI technique. Although the sagittal plane for DWI might be unusual, the lesions were completely shown on sagittal images and motion artifacts occurred in only 3 patients who were excluded from further evaluation. Third, the exact threshold value for predicting malignancy in cold nodules should be determined for each MR imaging unit because there are some variations in MR imaging systems, pulse sequences, and coils. The threshold value in our study was retrospectively evaluated and should be tested for cold thyroid nodules in a further study. However, our data might be helpful in the characterization of cold thyroid nodules, considering the large number of patients with cold nodules.

EPI DWI in the head and neck is still limited by technical problems with regard to susceptibility artifacts, spatial resolution, and motion artifacts due to swallowing, respiration, or blood flow.¹⁰ Therefore, we believe that in contrast to ≥3T MR imaging, the use of a 1T or 1.5T MR imaging machine might be beneficial with regard to artifacts.²⁶ However, technical developments in DWI sequences and dedicated coils in the field of ≥3T MR imaging could overcome this disadvantage.²¹ Studies performed with ≥3T MR imaging machines are required to evaluate the benefit of high spatial resolution with a small voxel size (<0.5 mm³) for the detection of even small carcinomas.

Conclusions
The results of this pilot study show that quantitative DWI seems to be a reliable noninvasive diagnostic technique by which to differentiate benign from malignant cold thyroid nodules of >8 mm. DWI seems to have the potential to be of great value as a diagnostic tool in addition to thyroid sonography and USgFNAC for tissue characterization of thyroid nodules; however, further studies are required, including a larger number of patients, to confirm our results.

Acknowledgment
We thank Mary McAllister, Johns Hopkins University Hospital, Baltimore, Md, for manuscript assistance.

References
7. Stark DD, Clark OH, Moss AA. Magnetic resonance imaging of the thyroid, thymus, and parathyroid glands. Surgery 1984;96:1083–91
15. Kim YJ, Chang KH, Song IC, et al. Brain abscess and necrotic or cystic brain