Are signal intensity and homogeneity useful parameters for distinguishing between benign and malignant soft tissue masses on MR images? Objective evaluation by means of texture analysis

Marius E. Mayerhoefer[a,b,*], Martin Breitenseher[a,c], Gabriele Amann[d], Martin Dominkus[e]

[a]Department of Radiology, Medical University of Vienna, Austria
[b]Institute of Radiology, LKH Oberwart, Oberwart, Austria
[c]Department of Radiology, Waldviertelklinikum Horn, Horn, Austria
[d]Department of Clinical Pathology, Medical University of Vienna, Vienna, Austria
[e]Department of Orthopedics, Medical University of Vienna, Vienna, Austria

Abstract

Objectives: To objectively identify possible differences in the signal characteristics of benign and malignant soft tissue masses (STM) on magnetic resonance (MR) images by means of texture analysis and to determine the value of these differences for computer-assisted lesion classification.

Method: Fifty-eight patients with histologically proven STM (benign, $n=30$; malignant, $n=28$) were included. STM texture was analyzed on routine T1-weighted, T2-weighted and short tau inversion recovery (STIR) images obtained with heterogeneous acquisition protocols. Fisher coefficients (F) and the probability of classification error and average correlation coefficients (POE+ACC) were calculated to identify the most discriminative texture features for separation of benign and malignant STM. $F>1$ indicated adequate discriminative power of texture features. Based on the texture features, computer-assisted classification of the STM by means of k-nearest-neighbor (k-NN) and artificial neural network (ANN) classification was performed, and accuracy, sensitivity and specificity were calculated.

Results: Discriminative power was only adequate for two texture features, derived from the gray-level histogram of the STIR images (first and 10th gray-level percentiles). Accordingly, the best results of STM classification were achieved using texture information from STIR images, with an accuracy of 75.0% (sensitivity, 71.4%; specificity, 78.3%) for the k-NN classifier, and an accuracy of 90.5% (sensitivity, 91.1%; specificity, 90.0%) for the ANN classifier.

Conclusion: Texture analysis revealed only small differences in the signal characteristics of benign and malignant STM on routine MR images. Computer-assisted pattern recognition algorithms may aid in the characterization of STM, but more data is necessary to confirm their clinical value.

1. Introduction

Magnetic resonance imaging (MRI) is generally recognized as the method of choice for the depiction and staging of soft tissue masses (STM) [1–4]. Establishing the correct diagnosis for STMs on magnetic resonance (MR) images, however, is considered difficult and requires a high level of experience in musculoskeletal image interpretation. To improve the results of STM classification, visual and, thus, subjective assessment of different morphologic parameters, including those that describe the signal behavior (e.g., intensity and homogeneity) of the lesions on different pulse sequences, has been recommended in the literature [3–8], but its value is still controversial [2,3,9].

Texture analysis (TA) of medical images is a sophisticated computer-assisted technique that allows detection of mathematical patterns in the gray-level distribution of the pixels

* Corresponding author. Tel.: +43 664 2046495; fax: +43 1 40400 4898.
E-mail address: marius.mayerhoefer@meduniwien.ac.at
(M.E. Mayerhoefer).

of digital images, thus providing an objective description of the signal behavior of anatomic structures or pathological processes. Because different types of tissue are enhanced by different pulse sequences, TA of native and contrast media-enhanced MR images is of particular interest and has been used with success in the detection of breast cancer [10,11] and muscle atrophy [12], characterization of vertebral trabecular bone structure [13], differentiation of brain tumors [14,15] and staging of liver cirrhosis [16].

In the work presented in this article, texture analysis was used to quantify the signal characteristics of STM on non-contrast media-enhanced MR images. The aim was to objectively determine whether, on MR images obtained in routine clinical practice with variations of acquisition parameters, differences in the signal characteristics exist between benign and malignant STM that might be useful for computer-assisted discrimination of the two tumor groups.

2. Material and methods

2.1. Patients and imaging

MR images from 58 patients (22 females and 36 males; mean age, 50 years; age range, 11–82 years) with STM were included in this retrospective study. All patients had undergone diagnostic MR-guided needle biopsy, and the histological diagnosis had been recorded. STM were divided into benign ($n=30$; nonneoplastic "tumor-like," $n=19$; neoplastic, $n=11$) and malignant lesions ($n=28$) (Table 1).

For texture analysis and computer-assisted STM classification, patients were assigned a numeric identifier and their personal data (name, age, gender) was rendered unidentifiable. The MR examinations, acquired at the time of biopsy,

Table 1
Histological diagnoses in the patient population (number of patients)

Benign nonneoplastic ($n=19$)	Benign neoplastic ($n=11$)	Malignant ($n=28$)
Connective tissue/ muscle/fat (5)	Fibromatosis (3)	Metastases (6)
Hematoma (4)	Angioma (2)	Sarcoma NOS [a] (4)
Pigmented villonodular synovitis (2)	Neurofibroma (2)	Carcinoma NOS (3)
Amyloid (1)	Cellular angiofibroma (1)	Liposarcoma (2)
Necrosis (1)	Lipoma (1)	Extraskeletal myxoid chondrosarcoma (2)
New bone formation (1)	Lymphangioma (1)	Plasmocytoma (2)
Myositis ossificans (1)	Chondromatosis (1)	Adenocarcinoma (2)
Cyst (1)		Follicular lymphoma (2)
Abscess (1)		Synovial sarcoma (1)
Gout (1)		Leiomyosarcoma (1)
Periostitis (1)		Rhabdomyosarcoma (1)
		Myxoid liposarcoma (1)
		Myofibroblastic sarcoma (1)

[a] NOS, not otherwise specified.

included T1-weighted (T1w), T2-weighted (T2w) and short tau inversion recovery (STIR) sequences and were obtained using a 1.0-Tesla MR scanner (Philips, Best, The Netherlands). Different receiver coils (body coil or extremity coil), depending on the location of the STM, and MR acquisition parameters were used for imaging.

For the T1w turbo spin echo (TSE) sequence (imaging plane: axial, $n=56$; coronal, $n=1$; sagittal, $n=1$), the acquisition parameters were as follows: repetition time (TR), 310–783 ms; echo time (TE), 9–20 ms; flip angle, 90°; matrix size, 256×256 ($n=25$) or 512×512 ($n=33$) and field of view (FOV), 13.9×13.9–41.8×41.8 cm.

For the T2w TSE sequence (imaging plane: axial, $n=57$; sagittal, $n=1$), the following acquisition parameters were used: TR, 2000–5336 ms; TE, 95–130 ms; flip angle, 90°; matrix size, 256×256 ($n=56$) or 512×512 ($n=2$) and FOV, 13.9×13.9–41.8×41.8 cm.

For the TSE STIR sequence (imaging plane: axial, $n=53$; coronal, $n=5$), the acquisition parameters were the following: TR, 1069–2309 ms; TE, 12–14 ms; inversion time, 130 ms; flip angle, 90°; matrix size, 256×256 ($n=54$) or 512×512 ($n=4$) and FOV, 13.9×13.9–44.8×44.8 cm.

2.2. Texture analysis and tissue classification

While the radiologist performing the texture measurements was aware of the category (benign or malignant) of the STM, he/she was blinded to the patients' clinical history and precise histological diagnosis. Because STM sometimes show different signal patterns, even on neighboring slice positions, two nonadjacent images that clearly depicted the lesion were selected from the image stack for each MR sequence of a patient and exported in DICOM format, yielding a total of 116 study objects. The radiologist responsible for the selection of the images was advised to choose those two images of a STM that were visually most distinguishable from each other, with regard to signal intensity and homogeneity.

Texture analysis was performed using the software package MaZda 3.20 (Institute of Electronics, Technical University of Lodz, Poland), which allows computation of almost 300 texture features. For each MR image, the entire visible STM, excluding, if present, extralesional reactions such as edema, as well as imaging artifacts, was used as the region of interest (ROI) (Fig. 1). In the next step, the signal intensity of each ROI was normalized using the limitation of dynamics to $\mu \pm 3\sigma$ (μ, gray-level mean; and σ, gray-level standard deviation). This was done to minimize the influence of brightness and contrast variation, caused, among other factors, by the differences in MR acquisition parameters. Then, texture features derived from the gray-level histogram, the co-occurrence matrix (information about the gray-level value distribution of pairs of pixels, separated by a defined distance in a given direction), the run-length matrix (information about runs of pixels with the same gray-level values in a given direction), the absolute gradient (information about the spatial variation of gray-level values) and the autoregressive model (description of texture based on the