European Journal of Radiology 95 (2017) 102–110

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Research article

Diagnostic hierarchy of radiological features in soft tissue tumours and proposition of a simple diagnostic algorithm to estimate malignant potential of an unknown mass

Leonhard Gruber*, Hannes Gruber, Anna K. Luger, Bernhard Glodny, Benjamin Henninger, Alexander Loizides

Department of Radiology, Medical University Innsbruck, Anichstraße 35, 6020 Innsbruck, Austria

ARTICLE INFO

Keywords:
Soft tissue tumours
Sarcoma
Features
Prediction
Classification

ABSTRACT

Objective: To quantify the diagnostic utility of imaging features in soft tissue masses (STMs) and to provide a ranked list of predictors for malignancy.

Subjects and methods: Imaging features in 260 cases of STMs with verified histology were assessed. Diagnostic properties including sensitivity, specificity, positive and negative predictive values, likelihood/odds ratios (OR) and normalized variance (NV) via *random forest* analysis were calculated. The diagnostic utility of an 8-item checklist consisting of the highest-ranked features was evaluated through a receiver-operating-characteristics (ROC) curve.

Results: The most predictive features (*NV/OR in parentheses*) were heterogeneous contrast-enhancement in ultrasound (297.9/15.1) and MRI (197.3/11.9), lesion roundness (209.8/5.5), diffusion restriction (175.8/9.3), cystic/necrotic intralesional areas (167.1/8.3), higher patient age (159.0/2.6), surrounding oedema (155.4/6.5) and intralesional Doppler hypervascularity (134.4/5.1).

A simple 8-item checklist was highly predictive of malignancy in cases with at least 75% positive features (0.90 area under the ROC curve, 87.0% sensitivity, 84.5% specificity, 59.5% positive and 96.1% negative predictive value, 36.5 odds ratio) even in cases with only partial feature availability.

Conclusion: Features vary widely in their diagnostic value in STMs; an 8-item checklist based on the eight most decisive features can be a simple tool to assess the likelihood for malignancy in unknown soft tissue masses, even though a stratified approach is certainly still advisable when first confronted with an STM.

1. Introduction

Soft tissue masses (STM) are a very heterogeneous group of tumours and still pose a significant challenge in clinical routine. While STMs have a high incidence, detection of malignant subtypes is complicated by a low relative incidence of soft tissue sarcomas (STS) and other malignant entities of the musculoskeletal system – the incidence rate of STS is considered to be around 1%–30% of overall STMs [1–3] –, and oftentimes nonspecific appearance as well as overlapping imaging features of benign, intermediate and malignant entities [4–9].

The early detection of STS is important as a delay in diagnosis and treatment – usually surgical resection – can lead to greater local complications, the need for a debilitating surgical removal [10] and a higher likelihood of metastases.

To identify malignant entities, a multi-facetted approach is usually employed including radiography (CR) [11,9,12], ultrasound (US) [13–16] and magnetic resonance imaging (MRI) examinations [4,9,16–22]. There are several publications on the diagnostic role of patient characteristics such as age and sex [2,3] and on single or combined lesion features such as size [10,20], border definition [23,24], lesion depth [17,19,20], affected region [15], relation to the investing fascia [25], MR signal intensity [18] or homogeneity [24], diffusion restriction [21,22], central necrosis [18] and tumour perfusion [22,26]. Extensive tissue characterization available from modern

Abbreviations: STM, soft tissue mass; STS, soft tissue sarcoma; CR, radiography; CT, computed tomography; US, ultrasound; CE, contrast-enhanced; CEUS, contrast-enhanced ultrasound; STIR, short tau inversion recovery; RI, resistive index; Gd, gadolinium; T1w, T1-weighted; T2w, T2-weighted; SI, signal intensity; CNB, core needle biopsy; PPV, positive predictive value; NPV, negative predictive value; OR, odds ratio; LR, likelihood ratio; NV, normalized variance; AUC, area under the curve; ROC curve, receiver-operating characteristics curve

* Corresponding author.
E-mail addresses: leonhard.gruber@i-med.ac.at (L. Gruber), hannes.gruber@i-med.ac.at (H. Gruber), anna.luger@i-med.ac.at (A.K. Luger), bernhard.glodny@i-med.ac.at (B. Glodny), benjamin.henninger@i-med.ac.at (B. Henninger), alexander.loizides@i-med.ac.at (A. Loizides).

http://dx.doi.org/10.1016/j.ejrad.2017.07.020

MRI examinations has led to the proposition of semi-quantitative image evaluation [19,21] and multifactorial algorithms [17,18].

Currently the replacement of core needle biopsy (CNB) or other forms of tissue sampling such as open biopsy stills appears elusive, as imaging methods do not provide sufficiently reliable information on a lesion's differentiation in many cases [4,23,24]. Nonetheless a focused set of differential diagnoses is important for pre-biopsy planning [27], interpretation of ambiguous CNB results and surgical therapy.

In this retrospective study we assessed the diagnostic utility of several demographic, CR, US and MRI features in the diagnosis of STMs and propose a hierarchical selection of features derived through a classical statistical analysis as well as a probabilistic *random forest* analysis [28,29].

2. Subjects and methods

After exclusion of 2 cases with inconclusive histology, a total of 260 consecutive cases of patients with STMs who had undergone US-guided biopsy between September 2009 and July 2015 were examined in this retrospective study. For further information on distribution and frequency of entities and subgroups, please refer to Table 2.

Primary inclusion criteria:

1) a histopathologic diagnosis
2) either a diagnostic US or MRI study (preferably both).

Data handling was performed according to the World Medical Association Declaration of Helsinki 2008 [30] and national legal norms. Institutional review board approval was granted by means of a general waiver by the local ethics committee for retrospective studies (February 20th, 2009).

Ultrasound-guided biopsy was performed in a standardized fashion as described [13]. Ultrasound machines used in the study were a Philips iU22 (Philips; Bothwell, WA, USA), a GE Logiq E9 (GE Healthcare; Chalfont St Giles, UK) and a Hitachi Arietta V70 (Hitachi; Tokyo, Japan). The transducer properties were chosen to achieve best tumour visualization.

If both CR and computed tomography (CT) were available, then CT studies were used for the analysis.

Minimum inclusion requirements for any MRI examinations were the acquisition of at least the following sequences: T1- (TE: 9–20 msec, TR: 303–793; T1w) and T2-weighted (TE: 77–116, TR: 3390–6750 msec; T2w) images in coronal, axial or sagittal orientation and a fat-suppressed or STIR sequence in coronal, axial or sagittal orientation (TE: 33–86 msec, TR: 2500–7720 msec, TI 150–160 msec) (sequence timing details depending on examination type).

Analysed features in CR or CT, US and MRI are given in Table 1 (for illustrative cases please refer to Fig. 1). Three readers blinded to the final diagnoses and with 8, 4 and 3 years of experience in the field of musculoskeletal radiology performed the readings. In cases with unclear findings a consensus was sought between readers.

The final histological diagnosis was retrieved from our institutional patient data management system (Powerchart 2012.10.1.5; Cerner, Gmund, Austria). Lesions were categorized according to WHO literature [31]. Histopathological results from open biopsy (n = 12), re-biopsy (n = 3) or full tumour resection overruled any CNB diagnoses.

2.1. Statistical analysis

Statistical evaluation was performed in GraphPad Prism®6.05 (GraphPad Software Inc.; CA, USA) and SPSS version 22.0 (SPSS Inc., Chicago, Illinois, USA). A p-value < 0.05 was considered significant; mean values and standard deviation (SD) values as well as confidence intervals (CI) are provided as indicated.

The relative distribution of benign, intermediate and malignant cases grouped by single features was calculated. According values for sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), likelihood ratios (LR) and odds ratios (OR) for malignancy were calculated for each predictor. Contingency tables were analysed via a two-sided Fisher's exact test or with a chi-square test in case of more than 2×2 cells. Continuous variables were compared via a two-sided ANOVA with Dunnet's correction for multiple testing.

To rank all imaging and other features by their diagnostic value, a *random forest* analysis with 100.000 trees (bootstrap sampling of 8 items per iteration) was carried out within *R* (version 3.2.3, R Core Development Team, [28]) using the *randomForest* package [29]. Missing values were computed through the *rfImpute* function provided by the *randomForest* package. Features were ranked by their respective OR for malignancy (from Fisher's exact or chi-square tests) and their normalized out-of-bag variance for classification (NV).

A naïve Bayes analysis using the *e1071* package [32] was then carried out in *R* to determine the prediction performance of a probabilistic model with increasing numbers of predictors, ordered by their NV.

Finally, the value of a simple additive score calculated from the number of positive features among the available eight top-ranked features was assessed. The score result is given as percent positive of available features. The distribution of benign, intermediate and malignant lesions is given for eight score ranges; corresponding ORs for malignancy and ROC curves were calculated to determine cut-offs and diagnostic power.

3. Results

3.1. Study population

Mean patient age at biopsy was 55.15 ± 17.9 years (range 13–93 years). 51.3% of patients were female. 99.2% of patients had received an US examination and 90.3% an MRI.

58.2% of lesions were benign, 10.6% intermediate and 30.0% malignant. 0.8% of lesions had to be excluded due to an inconclusive histology (see Table 2).

3.2. Age and gender

Higher patient age was correlated with an increasing prevalence of malignancies; the prevalence of malignancies was lowest in patients between 30 and 40 years ($p < 0.0001$). Mean patient age was 53.4 ± 17.1 years for benign lesions, 44.9 ± 18.2 years for intermediate lesions ($p = 0.02$) and 61.9 ± 17.2 years for malignancies ($p = 0.0008$). There was no association of sex and lesion classification ($p = 0.67$).

3.3. Localization, compartment, depth, size and roundness

STMs were most frequently encountered at the thigh (27.0% of overall cases). In general there was a higher rate of malignancy close to or at the torso such as the gluteal/hip region (47.6%), thighs (36.6%) or ventral thorax (36.4%), while the occurrence of malignancy was low in hands (8.3%) and feet (4.0%, $p = 0.21$).

Most lesions were encountered in skeletal muscle (59.6%), where the frequency of malignancies (32.9%) was also highest among the compartments ($p = 0.45$).

Malignant tumours were significantly larger than benign ones ($p = 0.008$), but a substantial overlap between benign, intermediate and malignant lesions was evident (64.5 ± 48.3 mm vs. 78.8 ± 59.7 mm vs. 85.3 ± 50.5 mm, respectively). Malignant tumours exhibited significantly higher values for roundness (i.e. aspect ratio of maximum width to length) than benign masses ($p < 0.0001$) (cut-off 0.5, Fig. 2a). Mean values for roundness were 0.52 ± 0.17 for benign, 0.49 ± 0.23 for intermediate ($p = 0.23$ vs. benign) and