Diffusion-weighted imaging in oral squamous cell carcinoma using 3 Tesla MRI: is there a chance for preoperative discrimination between benign and malignant lymph nodes in daily clinical routine?

Christina M Wendl¹, Steffen Müller², Johannes Eiglsperger³,*, Claudia Fellner¹, Ernst M Jung¹ and Johannes K Meier²

Abstract

Background: Preoperative staging of cervical lymph nodes is important to determine the extent of neck dissection in patients with oral squamous cell carcinoma (OSCC).

Purpose: To evaluate whether a preoperative discrimination of benign and malignant cervical lymph nodes with diffusion-weighted imaging (DWI) (3T) is feasible for clinical application.

Material and Methods: Forty-five patients with histological proven OSCC underwent preoperative 3T-MRI. DWI (b = 0, 500, and 1000 s/mm²) was added to the standard magnetic resonance imaging (MRI) protocol. Mean apparent diffusion coefficients (ADC_{mean}) were measured for lymph nodes with 3 mm or more in short axis by two independent readers. Finally, these results were matched with histology. Results: Mean ADC was significantly higher for malignant than for benign nodes (1.143 ± 0.188 * 10⁻³ mm²/s vs. 0.987 ± 0.215 * 10⁻³ mm²/s). Using an ADC value of 0.994 * 10⁻³ mm²/s as threshold results in a sensitivity of 80%, specificity of 65%, positive predictive value of 31%, and negative predictive value of 93%.

Conclusion: Due to a limited sensitivity and specificity DWI alone is not suitable to reliably discriminate benign from malignant cervical lymph nodes in daily clinical routine. Hence, the preoperative determination of the extent of neck dissection on the basis of ADC measurements is not meaningful.

Keywords

Oral squamous cell carcinoma, lymph nodes, diffusion-weighted imaging (DWI), apparent diffusion coefficient, magnetic resonance imaging (MRI)

Date received: 3 May 2015; accepted: 1 September 2015
same oncologic results with less postoperative morbidity (2–4). Since the number of negative elective neck dissections could still be up to 80%, an even more restricted type of dissection, the so-called superselective neck dissection, was proposed. This technique is expected to further reduce the postoperative morbidity without compromising oncological outcomes (5). An important precondition for this surgical technique is a highly sensitive and specific imaging method which allows a reliable preoperative staging of the neck in order to determine the necessary extent of neck dissection. Clinical examination, B-mode ultrasound, and computed tomography (CT) do not provide the required sensitivity and specificity and are therefore not suited for surgical planning in terms of the extent of lymph node dissection (6–8). PET/CT cannot be used to modify the indication of elective neck dissection as PET/CT is not sufficiently sensitive for micrometastases (9–11).

Diffusion-weighted imaging (DWI), initially introduced for the detection of cerebral ischemia, is the most frequently researched functional magnetic resonance imaging (MRI) tool of the past few years. DWI derives its contrast from the (indirect) measurement of the Brownian motion. Varying grades of mobility of water molecules within different types of tissues lead to signal changes in DWI and changes in the quantitative measure of diffusion, the apparent diffusion coefficient (ADC). ADC has also been shown to correlate to tissue cell density, with highly dense tissues having a low ADC (12). Therefore, DWI has shown to be an appropriate method to depict lymph nodes (13,14). Whether DWI is suitable to reliably differentiate malignant and benign lymph nodes is subject of investigations and controversy. If tumor cells invade the lymph node tissue a decrease in diffusion is conceivable, possibly resulting in signal alteration in ADC. Regarding the efficiency of differentiating malignant and benign lymph nodes the results do diverge. With this histological matched study we aim to evaluate whether a preoperative discrimination of benign and malignant cervical lymph nodes with DWI using a 3T MRI is feasible, in particular with regard to clinical practicability.

**Material and Methods**

**Patients**

Fifty-one patients (31 men, 20 women; median age, 62 years; range, 47–82 years) with histologically proven OSCC were included in this prospective study after they gave their written informed consent. Six patients were excluded due to poor MR image quality. The study was approved by the local ethic committee (vote: 12-101-0198). Inclusion criteria were: (i) histologically proven OSCC; and (ii) treatment consisting of surgery including superselective neck dissection.

**MRI**

Imaging was performed with a 3.0 Tesla MR unit (Skyra; Siemens Healthcare, Erlangen, Germany) using a 24-channel head/neck array coil. For anatomical purpose a T2-weighted (T2W) fast spin-echo (FSE) sequence and a T1-weighted (T1W) turbo spin-echo (TSE) sequence were acquired in axial plane before (T2 and T1) and after contrast medium injection (only T1). Additionally, an axial diffusion-weighted sequence in the same plane was performed for study purpose before contrast medium application. Sequence parameters are listed in Table 1.

**ADC analyses**

ADC maps were automatically generated by the MR system using all b-values (NUMARIS/4, Syngo MR D13, Siemens Healthcare).

The analyses were performed by two radiologists, with 5 and 20 years of experience in head and neck imaging. Radiologists, surgeons, and pathologist were blinded to the findings of the other observers. All lymph nodes located in the levels I–IIa/b on both sides with short axes of 3 mm or more were selected in the DWI ($b = 1000 \text{s/mm}^2$). These neck levels contain helpful anatomical structures as salivary glands and different muscles, simplifying the correct topographical matching between the MR images and the pathological specimen. For the calculation of the mean ADC ($ADC_{mean}$) a polygonal ROI was drawn free-handed around each subcentimeter and solid supracentimeter lymph node on the b1000 images.

In case of obvious necrotic portion within a lymph node, determined as a hyperintense region on the T2W and hypointense region on T1W images, this area was excluded from the ROI as seen in Fig. 1. The correct anatomic placements of the ROIs were verified linking the transverse T2 TSE sequence with the b1000 images. Linking also the ADC map with the b1000 images by slice position, the ROIs could simply be transferred to the referring ADC map by copying the ROIs and inserting them again in the ADC map.

**Histological correlation**

During surgery lymph nodes were dissected systematically level-by-level with the adjacent reference structures as muscles and salivary glands analogue to the system devised by the American Joint Committee on Cancer (15). Each lymph node level specimen was pinned on a
cork board separately, with filaments anatomically marked. The fresh specimens were then transported to the pathology department. Each lymph node was carefully delineated and its short axis was measured. Thereafter, the lymph nodes were separately put into cassettes for further detailed histological workup. Residual fatty tissue was submersed in formol to detect very small lymph nodes that might have been overlooked. Finally, lymph nodes were fixed, sectioned, hematoxylin-eosin stained, and examined microscopically with regard to tumor infiltration.

Recording the maximum short axial diameter, the positional relation to the marking filaments and the exact anatomic location of each lymph node per level (related to adjacent anatomic structures such as muscles, veins, and salivary glands) a topographic correlation between the MR images and the pathological specimen was possible for each lymph node.

Table 1. Sequence parameters of the diffusion-weighted image.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence type</td>
<td>Spin-echo echo-planar imaging</td>
</tr>
<tr>
<td>Field of view (mm)</td>
<td>230</td>
</tr>
<tr>
<td>Matrix</td>
<td>144 × 144</td>
</tr>
<tr>
<td>Voxel size (mm³)</td>
<td>1.6 × 1.6 × 3.0</td>
</tr>
<tr>
<td>Slice thickness (mm)</td>
<td>3</td>
</tr>
<tr>
<td>Intersection gap (mm)</td>
<td>0.3</td>
</tr>
<tr>
<td>Slices (n)</td>
<td>35</td>
</tr>
<tr>
<td>Repetition time (ms)/Echo time (ms)</td>
<td>8200/76</td>
</tr>
<tr>
<td>Averages (n)</td>
<td>4</td>
</tr>
<tr>
<td>Type of fat saturation</td>
<td>SPAIR</td>
</tr>
<tr>
<td>B values (s/mm²)</td>
<td>0, 500, 1000</td>
</tr>
<tr>
<td>Acquisition time (min)</td>
<td>4.16</td>
</tr>
</tbody>
</table>

Fig. 1. Example of ROI placement for the ADC measurement of cervical lymph nodes. Axial (a) T2W and (b) T1W images are used to verify the correct matching between lymph nodes and diffusion-weighted images. (c) The region of interest is drawn around a lymph node metastasis on b1000 image, sparing a region suspect for necrosis (white arrow) and (d) afterwards copied on the corresponding ADC map.

Statistical analyses

Statistical analyses were performed using the statistic computing environment R (version 2.15, R Foundation for Statistical Computing). In order to compare the ADC of benign and malignant nodes a two-sided Student’s t-test was performed. Receiver-operating characteristic (ROC) curve analyses were
performed to evaluate the diagnostic performance for the differentiation of metastatic nodes based on ADC of the nodes. Sensitivity, specificity, positive and negative predictive value, and the area under the ROC curve (AUC) were calculated. Threshold values were determined according to the point nearest to the upper left corner in the ROC curves. Inter-observer agreement for the nodal ADC measurement was analyzed by calculating the intra-class correlation coefficient and by using the method of Bland and Altman.

Results

Histopathology

Of the 45 patients included in the study 26 had a pN0 status, eight a pN1 status, and 11 a pN2 status at histology. A total of 389 lymph nodes were found within the histological specimens. A node-by-node matching of histology and MRI was performed in 149 lymph nodes, which presented with a short axis diameter of 3 mm or more. Of these 149 nodes 25 were malignant and 124 benign. The mean axial diameter (short axis) of the benign lymph nodes was 5.3 mm (range, 3–12 mm; standard deviation, 1.6 mm) while the mean axial diameter of the malignant lymph nodes was 11.0 mm (range, 4–27 mm; standard deviation, 5.9 mm). Table 2 shows the distribution of the benign and malignant lymph nodes with regard to the size of the short axis. In the group of malignant lymph nodes 11 showed necrotic portions, seven in the group of lymph nodes with over 10 mm in short axis, and four with 6–10 mm in short axis.

ADC analyses

The mean ADCs (mean values of the two observers) were $0.987 \pm 0.215 \times 10^{-3}$ mm²/s (range, 0.550–1.712 $\times 10^{-3}$ mm²/s, median 0.943 $\times 10^{-3}$ mm²/s) for the benign lymph nodes and 1.143 ± 0.188 $\times 10^{-3}$ mm²/s (range, 0.855–1.603 $\times 10^{-3}$ mm²/s, median 1.123 $\times 10^{-3}$ mm²/s) for the malignant lymph nodes. The difference between ADC values for malignant lymph nodes and benign lymph nodes was found to be significant ($P = 0.0007$). The area under the curve for the detection of metastatic lymph nodes using the mean ADC values was 0.73 (95% CI, 0.64–0.82). The ROC curve is displayed in Fig. 2. A mean ADC value of 0.994 $\times 10^{-3}$ mm²/s was found to be the optimal threshold, resulting in a sensitivity of 80% (95% CI, 64–92%), specificity of 65% (95% CI, 56–73%), positive predictive value of 31%, and negative predictive value of 93%.

Fig. 3 depicts the distributions of the mean ADC of the two observers for the benign and malignant lymph nodes together with the found optimal threshold value.

Table 2. Distribution of malignant lymph nodes based on short axis diameter (d).

<table>
<thead>
<tr>
<th>Lymph nodes (n)</th>
<th>Metastatic lymph nodes (n)</th>
<th>Prevalence (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>d ≤ 5 mm</td>
<td>80</td>
<td>3</td>
</tr>
<tr>
<td>5 mm &lt; d ≤ 10 mm</td>
<td>55</td>
<td>9</td>
</tr>
<tr>
<td>D &gt; 10 mm</td>
<td>14</td>
<td>13</td>
</tr>
</tbody>
</table>

The Bland-Altman plot for the inter-observer agreement for the nodal ADC mean measurements is given in Fig. 4. The mean absolute difference (bias) in ADC mean values between the two readers was 6.3 $\times 10^{-6}$ mm²/s (95% CI, −84.5 $\times 10^{-6}$ mm²/s to 97.0 $\times 10^{-6}$ mm²/s) and the intra-class correlation coefficient for the nodal ADCmean measurements was 0.97.

Discussion

The aim of our study was to judge whether diffusion-weighted images could be used to differentiate malignant from benign cervical lymph nodes with the intent to preoperatively determine the extent of neck dissection in patients with OSCC. In concordance with previously published studies we observed a statistical significant difference in the mean ADC values of benign and malignant lymph nodes with malignant lymph nodes (unexpectedly) presenting with higher ADC values than benign ones (Figs 5 and 6).

Nonetheless, our results indicate that diffusion-weighted MRI is not accurate enough for surgical planning of neck dissections due to the overlap of ADC value distributions of benign and malignant lymph nodes (Fig. 3).

In 2013 Heijnen et al. came to a similar result after they examined 102 lymph nodes in the small bowel of 21 patients with primary rectal cancer. They found higher ADC values for malignant than for benign lymph nodes ($1.15 \pm 0.24 \times 10^{-3}$ mm²/s vs. $1.04 \pm 0.22 \times 10^{-3}$ mm²/s) resulting in a sensitivity of 67% and a specificity of 60%. Furthermore, they measured the ADC of each lymph node (ADCnode) relative to the main tumor ADC (ADCrel) resulting in a slight increase in sensitivity and specificity (75%/61%). In consideration of these results they reasoned that DWI in case of rectal cancer can demonstrate but not characterize lymph nodes (13).

Another study of Heusch et al. in 2013 dealt with the question whether ultrasound, 18F-FDG-PET/CT, or 18F-FDG-PET-MRI with DWI could improve the accuracy of lymph node detection in the neck. Interestingly, no differences in sensitivity, specificity,
and accuracy between 18F-FDG-PET-MRI with or without DWI were noted (52%, 97%, 92% versus 53%, 97%, 92%, respectively). Ultrasound was superior to 18F-FDG-PET-MRI with DWI concerning nodal staging (9 versus 12 patients). In their opinion, 18F-FDG-PET-MRI plus DWI may not increase nodal detection and preoperative staging in patients with OSCC and therefore the method cannot replace surgical staging of cervical lymph nodes in the near future (16).

In 2009 Vandecaveye et al. presented their study results on diffusion-weighted MRI for nodal staging of the neck including 301 lymph nodes. Using b-values of 0 and 1000 s/mm² they found a mean ADC of $1.19 \times 10^{-3}$ mm²/s for benign and $0.85 \times 10^{-3}$ mm²/s for malignant lymph nodes ($P < 0.0001$). With a cutoff value of $0.94 \times 10^{-3}$ mm²/s they had sensitivity of 84%, specificity of 94%, and accuracy of 91%. The false-negative findings were ascribed to intranodal necrosis which should according to the authors be ruled out in the anatomic sequences before rating the DWI (14). In concordance to these findings necrosis also plays an important role in our study as six of the 11 necrotic malignant lymph nodes presented with macroscopic non-visible necrotic portions in the histological workup.

Looking at the behavior of the ADC values with regard to necrosis the stadium of the necrosis at the time of MRI examination seems to be decisive. In terms of early necrosis the ADC value is often low due to restricted water diffusion caused by intracellular edema in the tumor cells during the transition to complete necrosis with liquefaction (17,18). On the contrary, already liquid necrotic portions with degraded cellular structures lead to high ADC values (13). As long as the anatomic sequences do not allow a reliable selection of necrotic portions within in other respects (e.g. size) benign appearing lymph nodes, the meaningfulness of defining an ADC cutoff value for the differentiation of malignant and benign lymph nodes remains questionable.

Besides necrosis, tissue cellularity, cell membrane permeability, and extracellular matrix composition can affect ADC results (19). Consequently, several
previous studies reported lower ADC values for malignant lymph nodes compared to benign lymph nodes (14,20). These differences could, for example, be explained by densely packed carcinoma cell deposits in the nodal microarchitecture which restrict the diffusivity resulting in low ADCs. However, increased ADC values in metastatic lymph nodes, as found in our study, could mainly be caused by proliferating tumor cells and low cellularity of metastatic tumors (21,22).

In addition, there are strong indications that the partial volume effect (PVE) influences the accuracy of
ADC measurements to a high degree (23). Former studies dealing with brain diffusion tensor imaging came to the conclusion that if the voxel size exceeds 1.5 mm on one side the accuracy of DTI measurements can be influenced significantly (23–25). In case of an EPI-based DWI these observations could be transferred to ADC measurement.

The PVE originates from the acquisition as signal averaging had to be performed if finite-size voxels include more than one structure. In case of neck lymph node measurements surrounding perilymphatic tissue as, for example, muscles or salivary gland tissue as well as hilar vascular structures could be incorporated in the ROI if the lymph node is of small size. The contribution of these falsified voxels to the measurement influences the analysis of ADC values to an unknown variable extent (26).

Another problem could arise from the size of the metastases within the lymph nodes. If the zone of tumor cell infiltration is small or if the infiltration pattern is diffuse a contamination of the ADC value by the inclusion of surrounding benign lymph node portions could occur.

Some of the limitations of the presented study need to be taken into consideration. First, although we tried to minimize the susceptibility artifacts by modifying the vendor supplied sequence, in some cases image distortions along the phase encoding direction could be observed, which potentially affect the ADC values measured in small lesions. This issue was already observed by other study groups (20,27).

Second, the unbalanced distribution of malignant and benign lymph nodes in our dataset results in a high negative predictive value. The predictive values are in general not independent of the setting in which the test is used but are strongly influenced by the prevalence (28). The negative predictive value tends to be high when applied in a population with a low likelihood of disease. Due to early detection of OSCC and the infrequency of lymph node metastases in this special type of tumor high negative predictive values of tests can often be seen in this field of research (14,16).

In conclusion, our findings together with the heterogeneous results of previous studies point out that DWI is of limited use in the discrimination of benign and malignant cervical lymph nodes due to a limited sensitivity and specificity. The preoperative determination of the extent of neck dissection reclusively on the basis of ADC measurements is thereby not recommended.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.
Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

References