Accuracy of MRI in prediction of tumour thickness and nodal stage in oral squamous cell carcinoma

Christine T Lwin\textsuperscript{a}, Rebecca Hanlon\textsuperscript{b}, Derek Lowe\textsuperscript{c}, James S. Brown\textsuperscript{a,d}, Julia A. Woolgar\textsuperscript{d,e}, Asterios Triantafyllou\textsuperscript{e}, Simon N. Rogers\textsuperscript{a,f}, Fazilet Bekiroglu\textsuperscript{e}, Huw Lewis-Jones\textsuperscript{b}, Hulya Wiesmann\textsuperscript{b}, Richard J. Shaw\textsuperscript{a,d,*}

\textsuperscript{a}OMFS, Aintree University Hospitals NHS Foundation Trust, United Kingdom
\textsuperscript{b}Aintree University Hospitals NHS Foundation Trust, United Kingdom
\textsuperscript{c}CStat, Aintree University Hospitals NHS Foundation Trust, United Kingdom
\textsuperscript{d}University of Liverpool, United Kingdom
\textsuperscript{e}OMFS, Aintree University Hospitals NHS Foundation Trust, United Kingdom
\textsuperscript{f}University of Liverpool, United Kingdom

\textbf{SUMMARY}

We aim to compare radiological with histological tumour thickness (RTT with HTT) for oral squamous cell carcinoma (OSCC), and the ability of both to predict cervical metastasis. The MRI images and histopathology reports of 102 consecutive OSCC cases were compared and the relationship between RTT and HTT, calculated as a “shrinkage factor” by the gradient of the best fitting regression line. Most (69%) tumours appeared thicker on MRI than was revealed by histopathology. Shrinkage factor was 0.70 (interquartile range 0.63–0.77, correlation co-efficient 0.63) for all cases, 0.87 (IQ 0.80–0.95, CC 0.88) for tongue and 0.65 (IQ 0.49–0.82, CC 0.45) for floor of mouth sub-sites. RTT did not correlate well with the presence of nodal metastases in any sub-site, i.e. there was no clinically applicable cut-off value of RTT to determine the prescription of elective neck dissection. Although RTT has some predictable relationship with HTT, this varies between sub-sites with tongue the most accurately predicted shrinkage using axial MRI. It is not possible from either the MRI staging of neck or tumour thickness to safely determine the need for neck dissection in OSCC. It is necessary to re-evaluate the benefit of MRI as a staging investigation (particularly for early stage OSCC) and further explore the contribution of molecular biomarkers and ultrasounds.

\begin{tabular}{l}
\textit{Corresponding author. Address:} 5th Floor UCD, Duncan Bldg, Daubly Street, Liverpool L69 3GA, United Kingdom. Tel.: +44 0151 706 4183, +44 0151 529 5290. \\
E-mail addresses: Richard.shaw@liverpool.ac.uk, rjshaw@liv.ac.uk (R.J. Shaw).
\end{tabular}

\textcopyright{} 2011 Elsevier Ltd. All rights reserved.

\textbf{Introduction}

The value of imaging in the staging of oral squamous cell carcinoma (OSCC) is in judging operability, assessment of the prognostic characteristics and dimensions of the primary tumour, the presence of cervical metastasis and diagnosis of synchronous primary tumours.\textsuperscript{1} The prognosis of OSCC is closely related to features of the primary tumour such as tumour thickness\textsuperscript{2,3} and of the cervical metastases such as extracapsular spread (ECS).\textsuperscript{4} Although it is recognised that magnetic resonance imaging (MRI) lacks the sensitivity to replace elective neck dissection,\textsuperscript{5} it remains a widely used technique in clinical practice. Various techniques have been proposed for preoperative evaluation of tumour thickness. Clinical palpation to gauge the thickness of a tumour has not been evaluated, but palpation has been found to be inaccurate in assessment of cervical lymphadenopathy.\textsuperscript{6} The few published studies of MRI in prediction of tumour thickness and have been small (18–50 patients) and limited to oral tongue carcinoma sub-site.\textsuperscript{7,11} There is little evidence in other common\textsuperscript{12} oral cavity sites and it seems likely that imaging protocols are applied to OSCC rather than oral tongue alone. These reports have suggested that radiologically measured tumour thickness (RTT) is predictive of cervical metastases, but the cut off value has varied widely from 3 to 10 mm for tongue SCC. Evidence is emerging that HTT may be predicted with intra-oral ultrasound,\textsuperscript{13,14} although this may be limited by pain, trismus, and the subjective nature of interpretation. Many patients with early stage OSCC receive no proven methods to assess tumour thickness pre-operatively, although it is common practice to attempt evaluation in order to determine the requirement for elective neck dissection in the N0 neck.

Various techniques have been proposed for preoperative evaluation of tumour thickness. Clinical palpation to gauge the thickness of a tumour has not been evaluated, but palpation has been found to be inaccurate in assessment of cervical lymphadenopathy.\textsuperscript{6} The few published studies of MRI in prediction of tumour thickness and have been small (18–50 patients) and limited to oral tongue carcinoma sub-site.\textsuperscript{7,11} There is little evidence in other common\textsuperscript{12} oral cavity sites and it seems likely that imaging protocols are applied to OSCC rather than oral tongue alone. These reports have suggested that radiologically measured tumour thickness (RTT) is predictive of cervical metastases, but the cut off value has varied widely from 3 to 10 mm for tongue SCC. Evidence is emerging that HTT may be predicted with intra-oral ultrasound,\textsuperscript{13,14} although this may be limited by pain, trismus, and the subjective nature of interpretation. Many patients with early stage OSCC receive no proven methods to assess tumour thickness pre-operatively, although it is common practice to attempt evaluation in order to determine the requirement for elective neck dissection in the N0 neck.

Various techniques have been proposed for preoperative evaluation of tumour thickness. Clinical palpation to gauge the thickness of a tumour has not been evaluated, but palpation has been found to be inaccurate in assessment of cervical lymphadenopathy.\textsuperscript{6} The few published studies of MRI in prediction of tumour thickness and have been small (18–50 patients) and limited to oral tongue carcinoma sub-site.\textsuperscript{7,11} There is little evidence in other common\textsuperscript{12} oral cavity sites and it seems likely that imaging protocols are applied to OSCC rather than oral tongue alone. These reports have suggested that radiologically measured tumour thickness (RTT) is predictive of cervical metastases, but the cut off value has varied widely from 3 to 10 mm for tongue SCC. Evidence is emerging that HTT may be predicted with intra-oral ultrasound,\textsuperscript{13,14} although this may be limited by pain, trismus, and the subjective nature of interpretation. Many patients with early stage OSCC receive no proven methods to assess tumour thickness pre-operatively, although it is common practice to attempt evaluation in order to determine the requirement for elective neck dissection in the N0 neck.
un-necessary neck dissections due to the uncertainty of staging accuracy; hence further evidence is needed in options for safe de-escalation of surgery.

The aim of our study is to evaluate the accuracy of MRI measured tumour thickness (RTT) with HTT for all sub-sites in OSCC. We also aim to discover if RTT gives a cut-off value to predict metastasis to cervical lymph nodes.

Patients and methods

Participants

Consecutive patients with previously untreated OSCC were identified from our Head and Neck Oncology treatment records. Patients were excluded if they had missing MRI imaging or pathology reports. The following data were recorded retrospectively from the histopathology reports: sub-site and pTNM stage of primary tumour, according to AJCC/UICC classification, total histological tumour thickness to the nearest mm, presence and type of neck dissection performed. Histological tumour thickness (HTT) was measured to nearest mm with an optical micrometer. Sections were prepared according to Royal College of Pathologists protocol whereby the specimen is sliced, each slice is processed, and the one showing thickest tumour is recorded as HTT (the advancing front included any satellite islands and nerve invasion or vascular emboli). HTT reports were carried out by experienced head and neck pathologists (JAW and AT) with specialist interest in OSCC. Although inter- and intra-observer errors were not investigated, the pathologists were blind to the detail of the MRI report at the time of their measurement of HTT. For those with a neck dissection, the presence of intranodal and extra-capsular spread were also recorded according to our established protocols. In order to ensure uniformity of radiological reporting criteria, all pre-operative MRI images were sequentially re-reported by a single Head and Neck Radiologist (RH) using digitised images. This re-reporting was blinded to any correlating pathological or clinical information other than the original radiology request. Measurements were: radiological tumour thickness (RTT) to the nearest mm, nodal status and the radiological suggestion of ECS. The criteria for defining a node as radiologically positive was based on local modifications of the van den Brekel protocol: lymph node dimension in the short axis of nodes greater than 10 mm (5 mm for retropharyngeal nodes, 15 mm for jugulodigastric node) were used. Standard diagnostic MRI protocols using 5 mm slice thickness were employed throughout. Short-tau inversion recovery (STIR) and T2-weighted images were used to measure RTT in the most appropriate plane perpendicular to the mucosal orientation. Axial views were used to measure RTT for tongue (see Figs. 1 and 2), tonsil, and buccal sub-sites where tumours lie vertically and hence tumour thickness is reflected in a medial to lateral dimension. Reconstructed craniocaudal views were used for the floor of mouth (see Figs. 3 and 4), retromolar, soft palate and alveolus where tumour thickness is reflected in a vertical dimension.

Statistical method

Linear regression methods were used to investigate the relation between MRI reported tumour thickness (RTT) with the pathology reports (HTT). A ‘shrinkage factor’ was defined as the slope of the best fitting line constrained to fit through the origin. Pearson correlation coefficient was also computed to quantify the strength of linear correlation between the two measures of thickness, using the best fitting straight line not constrained through the origin. The Mann–Whitney test was used to compare both HTT and RTT in regard to pN and pECS status.

Results

Cohort characteristics

A total cohort of 102 consecutive patients comprising 34 females and 68 males, with a mean age 59 years (range 23–89) underwent primary surgery for previously untreated OSCC between October 2007 and December 2008. Selective neck dissection was carried out in 99/102 patients. There were 77 (76%) patients with radiologically negative necks (rN0), out of which 74 (96% of rN0 cases) received a neck dissection: 65 (84%) selective ipsilateral neck dissection (SND) and 9 (12%) bilateral SND. The 25 rN+ patients all underwent SND: 16 (64%) ipsilateral and 9 (36%) bilateral SND. For the total cohort, 60 (61%) had SND Levels I–III, 21 (21%)
had SND Levels I–IV, and 19 (19%) had bilateral SND. No radical or Levels I–V neck dissections were performed in this cohort.

Surgical resection of the primary tumour was carried out in all cases. Twenty-four (24%) cases were managed without reconstruction, of which 5 were laser resections and 19 underwent primary closure. Microvascular free flap reconstruction was carried out in 78 (76%) patients.

Clinical T stages of 102 patients were as follows: T1 = 28 (27%), T2 = 24 (24%), T3 = 7 (7%) and T4 = 43 (42%). AJCC pTNM Staging for the total cohort was: Stage I n = 11 (11%), Stage II n = 30 (29%), Stage III n = 16 (16%), Stage IVa n = 45 (44%). Tumour sub-site distribution was: oral tongue 44 (43%), floor of mouth (FOM) 29 (28%), alveolus 8 (8%), soft palate 7 (7%), tonsil 6 (6%), buccal mucosa 4 (4%), retromolar 3 (3%) and hard palate 1 (1%). There were 4 patients for whom the HTT was not stated in the pathology report (3 mandibular bone as deep margin and 1 tonsil sub-site). In 20/102 (20%) cases, the tumour could not be visualised on MRI scan. For 7 of these cases the scan was technically inadequate (5 through movement artifact, 2 where tumour site was not encompassed by imaging) they were excluded from correlation analysis with HTT. For the remaining 13 the scan was technically satisfactory but tumour not seen, and in these cases the RTT recorded as 0 mm. The commonest reason for this was a small anterior floor of mouth tumour missed between adjacent (5 mm) axial images.

Correlation/regression between HTT and RTT

When comparing the MRI reported tumour thickness (RTT) with the pathology reports (HTT), most values fell below the 1:1 line of agreement (dashed line in Fig. 5) i.e. RTT was greater than HTT in 63/91 cases, the same value for 2/91 cases and less than HTT for 26/91. This diminution of thickness between imaging and histology could be calculated as a “shrinkage factor” of 0.7, reflecting graphically by the slope 0.70 of the best fitting regression line constrained to pass through the origin, seen in Fig. 5 as the dotted line. Hence a tumour measured on MRI as 10 mm RTT typically translates to 7 mm HTT on the pathology report.

The analysis between HTT and RTT was also carried out for the two most numerous sub-site groups: oral tongue (Fig. 6) and floor of mouth (FOM) (Fig. 7). For tongue the shrinkage factor was 0.87 and for FOM it was 0.65. It is notable that a number of substantial tumours proved invisible in technically adequate MRI scans (cases aligned vertically on the y axis with RTT 0 and HTT 2–24 mm). Other sub-sites (tonsil, buccal, oropharynx, retro-molar, soft palate and alveolus) were analysed together as one group (“other”), due to small numbers, with a shrinkage factor of 0.59. The shrinkage factor for tongue gave a more accurate prediction of HTT than the FOM or other sites, as demonstrated by the narrower confidence intervals and higher correlation co-efficient, as seen in Table 1.

\[
\text{rN vs pN stage}
\]

Forty-three (43%) were pN1 or 2 (“pN+”) and 56 (56%) were pN0. MRI sensitivity for detecting histologically positive nodes...
was 51% and specificity was 95%. The positive predictive value (PPV) was 88% and the negative predictive value was 72%. Table 2.

A total of 23 patients had pathological extra-capsular spread pECS+ but none of these were detected on MRI (sensitivity 0%).

Correlation between tumour thickness measures (RTT and HTT) and pN status

The results shown in Table 3 replicate the previously reported, and anticipated, correlation that thicker tumours are more likely pN+, more particularly for tongue sub-sites. A similar pattern of results was seen for both HTT and RTT. There was stronger evidence that thicker tumours were more likely pECS+.

Cut-off in RTT to predict pN+

There was no valid RTT cut-off that gave a satisfactory clinical test in predicting pN+ cases. This was also evident when considering individual sub-sites such as tongue or floor of mouth only, or as a clinical test to predict pECS (data not shown). Predictive models using cut-off of 5 mm and 10 mm (previously cited cut-off values for HTT) are illustrated in Table 4, both self-evidently unsatisfactory.

Discussion

The data reported here show that MRI has some predictive value in estimating tumour thickness although the clinical relevance of this to management has not been conclusively demonstrated. A shrinkage factor is reported, giving a prediction of the histological tumour thickness (HTT) from the MRI determined tumour thickness (RTT). This shrinkage factor was 0.70 for all cases, 0.87 for tongue and 0.65 for floor of mouth sub-sites. The accuracy of prediction of HTT for tongue using this shrinkage factor was more accurate for tongue than other sites. RTT did not correlate well with pN+, i.e. there was no convincing cut-off value of RTT to guide the prescription of neck dissection.

Our study is based on a large number of consecutive cases of various oral sub-sites using strict and uniform protocol for both radiological and histological measures, but avoiding complex measures and reconstructions not in everyday practice. In that regard the data may offer superior clinical applicability than other published series. It may appear that the clinical stage in the OSCC cases seen was rather advanced, certainly in comparison with our previous practice, method of reconstruction and rate of ECS. However this appears merely a result of chance in randomly chosen cases rather than by design.

Although the imaging seen in our study was entirely representative of our routine clinical practice, it was evident that the axial MRI images were offered higher resolution that the cranio-caudal reconstructions based on these acquisitions. This was, in part, presumably responsible for the insensitivity in reporting floor of mouth tumours and relatively poor correlation co-efficient seen in the shrinkage factor. In this study we have not addressed the
Table 1
Shrinkage factor and correlation co-efficient between RTT and HTT for OSCC sub-sites.

<table>
<thead>
<tr>
<th>Sub-site</th>
<th>N</th>
<th>Shrinkage factor (slope of regression line constrained to pass through the origin)</th>
<th>95% CI for shrinkage factor</th>
<th>P value of regression</th>
<th>Best fitting regression line</th>
<th>Correlation coefficient</th>
<th>P value of correlation/best fitting regression line</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tongue</td>
<td>43</td>
<td>0.87</td>
<td>0.80–0.95</td>
<td>&lt;0.001</td>
<td>–0.4 + 0.89RTT</td>
<td>0.88</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>FOM</td>
<td>24</td>
<td>0.65</td>
<td>0.49–0.82</td>
<td>&lt;0.001</td>
<td>7.7 + 0.23RTT</td>
<td>0.45</td>
<td>0.03</td>
</tr>
<tr>
<td>Other sites</td>
<td>24</td>
<td>0.59</td>
<td>0.44–0.74</td>
<td>&lt;0.001</td>
<td>5.6 + 0.34RTT</td>
<td>0.50</td>
<td>0.01</td>
</tr>
<tr>
<td>Total</td>
<td>91</td>
<td>0.70</td>
<td>0.63–0.77</td>
<td>&lt;0.001</td>
<td>5.0 + 0.47RTT</td>
<td>0.63</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

Table 2
Comparison of rN and pN status.

<table>
<thead>
<tr>
<th>Radiological N stage</th>
<th>pN Stage</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pN0</td>
<td>pN+</td>
</tr>
<tr>
<td>rN0</td>
<td>53</td>
<td>21</td>
</tr>
<tr>
<td>rN+</td>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>56</td>
<td>43</td>
</tr>
</tbody>
</table>

Table 3
Correlation of HTT/RTT and pN status.

<table>
<thead>
<tr>
<th>HTT</th>
<th>N</th>
<th>Median</th>
<th>IQR</th>
<th>Range</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>All cases</td>
<td>pN0</td>
<td>54</td>
<td>9.3</td>
<td>6.7–13.8</td>
<td>0.8–22</td>
</tr>
<tr>
<td></td>
<td>pN+</td>
<td>41</td>
<td>12.0</td>
<td>7.7–19.5</td>
<td>2–45</td>
</tr>
</tbody>
</table>

**Tongue:**

| pN0 | 25 | 9.0 | 6.0–12.7 | 0.8–20 | 0.13 |
| pN+ | 19 | 13.0| 7.0–20.0 | 2–45   |       |

**FOM:**

| pN0 | 17 | 11.3| 9.1–16.0 | 5–21   | 0.37 |
| pN+ | 9  | 10.7| na       | 2–19   |       |

| All cases | pECSneg | 77 | 9.4 | 6.0–14.0 | 0.5–30 | 0.001 |
|           | pECS+   | 21 | 16.0| 10.2–20.0| 7–45   | 0.37  |

In comparison with other published literature on tongue SCC, Preda et al. showed a direct correlation between measured MRI and histological tumour thickness (correlation coefficient = 0.80, p < 0.0001) in their study of 33 patients. This is similar to our own findings of our tongue sub-site group whose RTT and HTT correlation was found to be 0.88 (p < 0.001). Iwai et al. in a study of 30 T1–T3 tongue SCC patients found a correlation of 0.609 (p = 0.002) between their MRI and histological tumour thickness, which became more significant at 0.839 (p < 0.0001) when they used reconstructive tumour thickness on MRI (defined as the difference in distance between the distance from the surface to septum of the tongue on the unaffected side and the distance from the septum to the deepest extent of tongue carcinoma). Lam et al. studied 18 patients with T1–T3 tongue SCC showed correlation between MRI and histological tumour thickness to be 0.941 (p < 0.0005). Jung et al. in 50 patients with T1–T2 tongue SCC, reported a correlation of 0.813 (p < 0.001). Our study shows similar correlation figure to those in the literature, for our tongue sub-site group, despite being the only study to include T1–T4 tongue tumours. Additionally, our study has shown poor correlation between RTT and HTT for floor of mouth sub-site group (correlation coefficient 0.45, shrinkage coefficient 0.65, p < 0.001). Similar poor correlation was true of the remaining OSCC sub-sites. The omission of data other than tongue is striking, particularly in the

na: not appropriate for small N.

* Mann–Whitney test.

light of our previous observation that floor of mouth SCC is more common that tongue SCC.\(^{12}\)

The clinical message of this work is clear. Firstly, it is not possible, from either the MRI staging of neck or tumour thickness, to safely determine the need for neck dissection in OSCC. It is then necessary to re-evaluate the benefit of MRI as a staging investigation in terms of other valuable data generated. Although synchronous tumours were not studied here, it is likely that imaging may easily miss additional small tumours in the head and neck. Hence endoscopic visualisation of all relevant mucosae is mandatory irrespective of any imaging technique used. In larger tumours, MRI may offer valuable information regarding bone invasion and operability (cT4a versus cT4b), however this will be generally limited to cases deemed T3/4 on inspection. The value of MRI in judging operability is beyond the scope of this data, as surgical excision and resultant pTNM on pathology report were pre-requisites for inclusion. In T1 and T2 stages, it may be difficult to justify the additional costs of MRI in the light of the data presented, particularly in subsites other than tongue. MRI may offer false re-assurance to clinician and patient alike that accurate information about dimensions of the primary tumour and chance of metastasis has been established. Understandably, clinicians are under pressure to clarify the benefit of costly investigations based on careful audit.

In the light of this data, a re-appraisal of MRI compared with other available staging methods is timely. A prospective re-evaluation of clinical assessment (i.e. clinical palpation) seems reasonable. Newly emerging molecular biomarkers offer a possibility of risk stratification in judging the biological aggression and risk of metastasis but so far these have not made a clinical impact. Also ultrasound of the primary tumour, combined with needle aspira-

### References


