Does primary tumour volumetry performed early in the course of definitive concomitant chemoradiotherapy for head and neck squamous cell carcinoma improve prediction of primary site outcome?

¹K S S BHATIA, FRCR, ¹A D KING, FRCR, ²K-H YU, FRCR, ³A C VLANTIS, FCS(SA)ORL, MD, ⁴G MK TSE, FRCP, ²F KF MO, PhD and ¹A T AHUJA, FRCR

¹Department of Imaging and International Radiology, ²Department of Clinical Oncology, ³Department of Otorhinolaryngology, Head and Neck Surgery, and ⁴Department of Anatomical and Cellular Pathology, Faculty of Medicine, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong SAR, China

ABSTRACT. Although previous studies have documented correlations between pre-treatment or post-treatment primary tumour volumes and local outcome following definitive concomitant chemoradiotherapy (CCRT) in head and neck squamous cell carcinoma (HNSCC), no study has included and compared tumour volumes during CCRT. We reviewed the MRIs of 69 HNSCC patients treated with a 6 weeks course of CCRT and who underwent successful MRI pre-treatment ($n=69$), 2 weeks intra-treatment ($n=48$) and 6 weeks post-treatment ($n=61$). Primary tumour volumes on MRI at the three time points were calculated and compared for their predictive value for primary site outcome. Volume thresholds optimised to predict failure with the highest accuracy and positive predictive value (PPV) were calculated. The mean pre-treatment volume was 24.6 cm^3 (range, $1.1\text{--}187.9\text{ cm}^3$) and the mean follow-up interval was 41 months (range, 12–100 months). 23 primary tumours failed treatment (33%). Volumes before, during and after CCRT were positively associated with local failure ($p=0.015$, $p=0.009$, $p<0.0001$). Volume reductions during and after CCRT were negatively associated with local failure ($p=0.021$, $p=0.001$). Pre-treatment and intra-treatment volume thresholds achieved the highest accuracy and produced intermediate PPVs (51–64%) for predicting local failure. Optimised intra-treatment thresholds did not identify any more treatment failures than the pre-treatment thresholds. By comparison, a 6 weeks post-treatment volume reduction ($<35\%$) achieved 100% PPV for failure, albeit with 26% sensitivity. In conclusion, primary tumour volumetry performed early in CCRT provides minimal additional information compared with pre-treatment volumetry, with respect to predicting post-treatment local failures. Therefore, volumetry during CCRT is unlikely to be useful for guiding individual response-based therapeutic modifications.

Received 20 July 2009
Revised 30 September 2009
Accepted 19 October 2009

DOI: 10.1259/bjr/27631720

© 2010 The British Institute of Radiology

Concomitant chemoradiotherapy (CCRT) is an established definitive treatment for both early and advanced head and neck squamous cell carcinoma (HNSCC) at several primary sites. This treatment can achieve high rates of locoregional control and survival, while permitting preservation of organ function [1]. The evaluation of treatment response following definitive CCRT is a major challenge, however, owing to post-CCRT effects that hinder clinical and imaging assessment. In this regard, conventional CT and MRI are limited in the post-CCRT period by a delayed anatomical response in the treated tumour as well as by severe anatomical distortions caused by early mucositis and late fibrosis. Furthermore, functional

anatomical imaging using positron emission tomography (PET) CT has a poor positive predictive value (PPV) for failure shortly after (chemo) radiation owing to the high number of false-positives caused by irradiation-induced acute inflammation [2–4]. Despite these limitations, early prediction of potential treatment failures during or shortly after CCRT is an important goal, as salvage surgery is the only subsequent curative option and should be performed early to prevent residual tumour from becoming surgically unresectable and before irradiation-induced fibrosis has become established.

The primary objective of this study was to determine whether tumour volume indices on MRI performed early in the course of definitive CCRT can predict primary site outcome; to our knowledge, this has not previously been investigated. A secondary objective was to compare the discriminatory performance of volumetric data from MRI performed during CCRT with that obtained from pre-therapeutic and early post-therapeutic MRI.

Address correspondence to: Dr Kunwar S S Bhatia, Department of Diagnostic Radiology and Organ Imaging, Faculty of Medicine, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, 30-32 Ngai Shing Street, New Territories, Hong Kong SAR, China. E-mail: drkbhatia@cuhk.edu.hk

Methods and materials

Patient selection

Between 2001 and 2008, 89 patients with biopsy-confirmed HNSCC were identified at our institution who had received definitive CCRT to the primary site and for whom pre-treatment MRI scans were available. Treatment comprised four weekly doses of cisplatin (40 mg m^{-2}) on days 1, 8, 15 and 22 of a course of accelerated concomitant boost radiotherapy. A dose of 54 Gy/30 fractions/6 weeks was delivered to the gross tumour and sites at risk of microscopic disease. A second daily dose (1.5 Gy/fraction) to the gross tumour was delivered on the final 12 days of treatment. A total of 20 patients were subsequently excluded from the study: 19 had less than 1 year of disease-free follow-up and 1 had primary tumour that was inseparable from nodal metastases. In the remaining 69 patients, pre-treatment MRI ($n=69$) and 6 weeks post-treatment MRI ($n=681$) were evaluated in addition to MRI performed 2 weeks after the start of CCRT ($n=48$). The 2-week intratherapeutic MRI scan was part of another prospective study at our institution. Local ethics committee approval had been granted and written informed patient consent had been obtained for this study.

The final study population comprised 63 men and 6 women with a mean age of 59 years (range 45–75 years). Tumour sites were as follows: 24 oral cavity/oropharyngeal, 24 hypopharyngeal, 18 laryngeal and 3 in the nasal cavity. All patients had Stage III or IV disease with no known distant metastases at presentation [5].

MRI evaluation

Patients underwent MRI on a 1.5 T whole-body system (Intera NT; Philips Medical Systems, Best, the Netherlands) with a 30 mT m^{-1} maximum gradient capability using a standard receive-only head and neck coil. The imaging protocol included axial T_2 weighted turbo spin-echo images with fat suppression (repetition time [TR] ms/echo time [TE] ms 2500/100; 4 mm slice thickness with no interslice gap; two signals acquired), axial T_1 weighted spin-echo images (477/12; 4 mm slice thickness with no interslice gap; two signals acquired) and contrast-enhanced axial T_1 weighted spin-echo images obtained following a bolus injection of 0.1 mmol kg^{-1} gadoteric acid (Dotarem; Guerbet, Aulnay, France) with utilisation of a 512 imaging matrix.

MRI examinations were retrospectively reviewed on the imaging workstation by two observers (KSSB, ADK) in consensus; these observers have 19 and 4 years' specialist experience in head and neck cancer imaging, respectively. Reviewers were blinded to the final outcome during the assessment of MRI examinations. Primary tumour volumes were measured using a summation-of-areas technique in which the tumour was manually contoured on the MRI workstation on successive axial T_1 weighted post-intravenous gadolinium MR images; the summed areas of contours were then multiplied by the slice interval plus thickness (0.4 cm). In this respect, other imaging planes and sequences including T_1 weighted pre-contrast, T_2 weighted and fat saturation sequences were used to assist in anatomical localisation (Figure 1). Areas of

Figure 1. (a) A supraglottic cancer is contoured on an axial T_1 weighted post-gadolinium MRI sequence. (b) An axial T_2 weighted MRI sequence with fat suppression at the same level is used to assist in anatomical localisation (arrows).

obvious peritumoural inflammation and frank tumoural necrosis were specifically excluded from the contoured volume. Absolute values of tumour volume at the three time-points were documented and percentage reductions in tumour volume at 2 weeks into CCRT and 6 weeks after CCRT, compared with the pre-treatment volume, were calculated.

Assessment of local control

On completion of CCRT, patients underwent clinical follow-up every 1–2 months for the first 2 years and every 3–4 months thereafter. Clinical follow-up included physical examination, fiberoptic endoscopy, where relevant, and endoscopy under general anaesthesia with targeted biopsies if there was clinical or radiological suspicion of residual tumour. Patients underwent MRI at 6 months, 1 year and 18 months post CCRT; thereafter, MRI was carried out depending on clinical suspicion of residual tumour. Local failure was assigned if residual tumour was confirmed histologically at the primary site after completion of CCRT or in the absence of histology if there was definite radiological disease progression after CCRT on serial imaging. All other patients who survived at least 1 year without primary site residual disease or recurrence were designated as primary site control.

Data analysis

Statistical analyses were performed using SPSS 17.0 software (SPSS, Chicago, IL). The relationship between the primary tumour volume and clinical outcome in terms of local control or failure was assessed at the various time points using the Mann–Whitney *U*-test. Other statistical tests were used as shown. Statistical tests were two-sided; *p*-values <0.05 were considered statistically significant. The discriminatory performances of volumetric parameters were assessed by receiver operating characteristic (ROC) analysis. We determined optimal thresholds as

those producing the highest accuracy (ratio of true-positive tumours plus true-negative tumours to number of tumours) followed by the highest PPV for local failure if the highest accuracy was tied between more than one threshold. This approach was adopted because it could potentially be used to identify a subgroup of high-risk patients that might benefit from CCRT dose escalations or early salvage surgery, while minimising the number of patients reported with false-positives who would be overtreated.

Results

Local failures occurred in 23 of 69 (33%) patients with a mean interval of 8 months (range 3–29 months) from the start of treatment. There was no evidence of local failure in 46 (67%) patients, with a mean follow-up of 41 months (range 12–100 months). Primary failures consisted of 11 of 24 hypopharyngeal (46%), 8 of 18 laryngeal (44%), 3 of 24 oral cavity/oropharyngeal (13%) and 1 of 3 nasal cavity (33%) tumours. Local failure was confirmed histologically in all cases, except for one patient with a progressively enlarging carcinoma of the tongue identified on serial post-therapeutic MRI studies who declined further biopsies.

For all sites combined, the mean pre-treatment primary tumour volume was 24.6 cm³ (median 12.8 cm³, range 1.1–187.9 cm³). The frequency distribution of all tumours according to volume is shown in Figure 2.

For all primary sites combined, patients with local failure had higher absolute tumour volumes than patients with local remission at diagnosis (*p*=0.015), 2 weeks into CCRT (*p*=0.009) and 6 weeks after CCRT (*p*<0.0001) (Table 1). Patients with local failure had smaller percentage reductions in tumour volume 2 weeks into CCRT and 6 weeks after CCRT than patients with local remission (*p*=0.021 and *p*=0.001, respectively) (Table 1). Out of 48 tumours, 14 (29%) increased in volume by 2 weeks from the start of CCRT (median 29%, range 24–5.6%). An increase in tumour volume by

Figure 2. Tumour frequency according to pre-treatment volume for all tumours combined.

Table 1. Absolute and relative changes in primary tumour volume correlated with primary outcome for all primary sites

Parameter (number of patients)	Local control	Local failure	<i>p</i> -Value ^a
	Median (range)	Median (range)	
Volume pre-treatment (69)	8.1 (1.4–187.9) cm ³	18.9 (3.3–126.8) cm ³	0.015
Volume 2 weeks into CCRT (48)	3.3 (0.6–229.3) cm ³	17.2 (1.4–121.5) cm ³	0.009
Volume reduction 2 weeks vs pre-treatment (48)	44.4 (–24.7 to 89)%	22.1 (–81 to 90)%	0.021
Volume 6 weeks after CCRT (61)	0.9 (0.0–42.0) cm ³	6.0 (0.2–26.9) cm ³	<0.001
Volume reduction 6 weeks vs pre-treatment (61)	89.1 (35–100)%	76.5 (0.1–97)%	0.001

CRT, concomitant chemoradiotherapy.

^aMann–Whitney *U*-test.

2 weeks was not a risk factor for local failure ($p=0.095$, Fisher's exact test). Of the 14 tumours that increased in volume by 2 weeks, 13 were smaller than the pre-treatment volume by 6 weeks after CCRT; of these, 8 (57%) had local control. Out of 61 patients, 17 (28%) had no measurable residual mass on MRI 6 weeks after CCRT and none of these patients had local failure on follow-up. By comparison, 19 out of 44 patients (43%) with a residual mass at 6 weeks had local failure ($p<0.0001$; Fisher's exact test).

Table 2 shows the optimal thresholds for the volume parameters for all primary sites in terms of achieving the highest accuracy for local failure; these values were used in conjunction with the highest PPV if accuracy values were tied for more than one threshold. In descending order of accuracy, the best predictors of local failure were an absolute volume 6 weeks after CCRT >5.7 cm³, a percentage reduction in volume 6 weeks after CCRT of $<35\%$, an absolute volume 2 weeks into CCRT >10.6 cm³, a percentage reduction in volume 6 weeks after CCRT of $<-9.4\%$ and an absolute volume pre-treatment >11.2 cm³. It was not possible to adjust these thresholds to predict local failure with fewer false-positives (higher PPV) without appreciably lowering the sensitivity. For instance, 2 weeks into CCRT, raising the threshold used to predict failure from -9.4% to 50% volume reduction only marginally increased the PPV (from 64% to 75%), but appreciably lowered the sensitivity from 41% to 18%. Six weeks after completion of CCRT an interval reduction of $<35\%$ produced the highest attainable accuracy (75%) as well as the highest PPV (100%), although the sensitivity was only 26%. By comparison, using an arbitrary volume

reduction of $<50\%$ at 6-week after CCRT increased the sensitivity by 11% (to 37%), but lowered the PPV by 30% (to 70%). The overall discriminatory performance of each volume parameter, as indicated by the calculated area under the ROC curve, was highest for the 6-week post-CCRT scan, followed by the 2-week scan and, finally, the pre-treatment scan (Table 2).

To evaluate whether there was any additional benefit in performing the 2-week scan compared with the pre-treatment scan, the subgroup of 48 patients that had MRI performed at both of these time-points were assessed. For this subgroup, the optimal pre-treatment threshold produced 14 true-positives and 3 false-negatives for failure, whereas the absolute and relative thresholds 2 weeks into CCRT produced 13 and 7 true-positives, respectively. Furthermore, the PPVs of these parameters increased marginally from 51% using the pre-treatment threshold to 62% and 64% for the 2-week absolute and relative thresholds, respectively; the sensitivity was lowered more appreciably from 83% to 76% and 41%, respectively. On a case by case basis, the 2-week thresholds did not correctly identify any of three false-negatives for failure using the pre-treatment threshold.

Hypopharyngeal tumours were the only primary site suitable for subgroup analysis owing to the sufficient number of tumours ($n=24$) and an appreciable proportion of local failures (44%) at this site. For this site, patients with local failure had higher tumour volume pre-treatment, 2 weeks into CCRT and 6 weeks after CCRT than patients with local remission ($p=0.002$, $p=0.011$ and $p<0.0001$, respectively) (Table 3). Median

Table 2. Discrimination performances of volume parameters for all primary sites in terms of predicting post-treatment local failure

Parameter	Threshold for local failure ^a	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)	Area under ROC curve ^b
Volume pre-treatment	>11.2 cm ³	83	61	51	88	68	0.638
Volume 2 weeks into CRT	>10.6 cm ³	76	74	62	85	75	0.721
Volume reduction 2 weeks vs pre-treatment	$<9.4\%$	41	87	64	73	71	0.689
Volume 6 weeks after CRT	>5.7 cm ³	58	90	73	83	79	0.808
Volume reduction 6 weeks vs pre-treatment	$<35\%$	26	93	100	76	75	0.798

CRT, chemoradiotherapy; PPV, positive predictive value, NPV, negative predictive value.

^aThresholds are calculated from receiver operating characteristic (ROC) analyses as those achieving highest accuracy (true positives plus true negatives/total number of cases). ^bArea under the receiver operating characteristic (ROC) curve estimates global accuracy of the parameter, whereby a result of 1 indicates perfect discrimination and 0.5 indicates no discrimination.

Table 3. Absolute and relative changes in primary tumour volume correlated with primary outcome for hypopharyngeal tumours

Parameter (number of patients)	Local control	Local failure	<i>p</i> -Value ^a
	Median (range)	Median (range)	
Volume pre treatment (24)	8.1 (1.1–29.6) cm ³	20.1 (8.6–126.8) cm ³	0.002
Volume 2 weeks into CRT (17)	2.5 (0.6–10.7) cm ³	17.9 (1.4–121.5) cm ³	0.011
Volume reduction 2 weeks vs pre treatment (17)	29.1 (13.7–54.8)%	15.9 (–64.2 to 73.7)%	0.321
Volume 6 weeks after CRT (21)	1.5 (0.0–5.7) cm ³	9.3 (1.9–26.9) cm ³	<0.001
Volume reduction 6 weeks vs pre treatment (21)	92.6 (44.8–100)%	73.1 (–32.8 to 91)%	0.036

CRT, chemoradiotherapy.

^aMann–Whitney *U*-test.

percentage volume reduction at 2 weeks into CCRT was lower in patients with local failure compared with local control, although there was no statistically significant difference between these groups ($p=0.321$). Percentage volume reduction 6 weeks after CCRT was lower in patients with failure compared with control, which was statistically significant ($p=0.036$). Table 4 shows the discriminatory performance of the volume parameters for hypopharyngeal tumours using thresholds optimised to achieve maximum accuracy; these values were used in conjunction with the highest PPV if accuracy values were tied for two or more thresholds. In descending order of accuracy (plus descending order of PPV if accuracy values were tied), the best predictors of local failure were an absolute volume 6 weeks after CCRT >5.7 cm³, an absolute volume pre-treatment >11.0 cm³, an absolute volume 2 weeks into CCRT >10.6 cm³, a percentage reduction in volume 6 weeks after CCRT of $<44.8\%$ and a percentage reduction in volume 2 weeks after CCRT of $<47.2\%$. Unlike the situation for all primary sites combined, if volume thresholds for failure were determined as those producing the highest accuracy, these thresholds also achieved high PPVs ranging from 77% to 100%. Furthermore, some thresholds also achieved high sensitivities. In this context, the 6-week absolute threshold produced 100% PPV and 80% sensitivity for failure, the volume 2 weeks into CCRT produced 100% PPV and 78% sensitivity, and the pre-treatment volume threshold produced 77% PPV and 91% sensitivity. The overall discrimination performance of volume parameters, as indicated by the calculated area under the ROC curve, was highest for the pre-treatment threshold, followed by

the 6-week post-CCRT threshold and, finally, the 2-week intratherapeutic threshold.

Discrimination analysis of the subgroup of 17 hypopharyngeal tumours with MRI scans both pre-treatment and 2 weeks into CCRT indicated that the pre-treatment threshold produced seven true-positives, whereas the intratherapeutic threshold produced six true-positives (there were eight failures in total in this subgroup). Conversely, there was one false-positive for failure using the optimal pre-treatment volume threshold compared with no false-positives using the optimal 2-week absolute volume threshold.

Subgroup statistical or ROC analysis of other sites was not performed owing to the smaller number of tumours at these sites or low numbers of tumours that ultimately failed. Nevertheless, we noted that only 3 of 24 oral/oropharyngeal tumours failed locally after CCRT and, within this group, none of the 13 tonsil tumours had local failure. In this context, tonsil tumour size (mean 21.7 cm³, range 4.4–64.3 cm³) was not significantly different from that of hypopharyngeal tumours (mean 12.8 cm³, range 1.1–126.8 cm³; $p=0.594$, Mann–Whitney *U*-test).

Discussion

The present study has shown that patients with local failure after CCRT have higher primary volume measurements at diagnosis, 2 weeks into CCRT and 6 weeks after CCRT than tumours with local control. Similarly, reductions in volume 2 weeks into CCRT and 6 weeks after CCRT are lower in patients with local failure. Correlations

Table 4. Discrimination performances of volume parameters for hypopharyngeal tumours in terms of predicting post-treatment local failure

Parameter	Threshold for local failure ^a	Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)	Area under ROC curve ^b
Volume pre-treatment	>11.0 cm ³	91	77	77	91	83	0.952
Volume 2 weeks into CRT	>10.6 cm ³	78	100	100	80	88	0.873
Volume reduction 2 weeks vs pre-treatment	$<47.2\%$	78	63	70	71	71	0.635
Volume 6 weeks after CRT	>5.7 cm ³	80	100	100	85	90	0.937
Volume reduction 6 weeks vs pre-treatment	$<44.8\%$	40	100	100	65	71	0.841

^aThresholds are calculated from receiver operating characteristic (ROC) analyses as those achieving highest accuracy (true positives plus true negatives/total number of cases). ^bArea under the ROC curve estimates global accuracy of the parameter, whereby a result of 1 indicates perfect discrimination and 0.5 indicates no discrimination. CRT, chemoradiotherapy; PPV, positive predictive value; NPV, negative predictive value.

between pre- or post-therapeutic tumour volumes and primary outcome are well documented [6–17], although to our knowledge no other study has assessed the utility of primary tumour volumes early during CCRT. Volumetric analysis during CCRT has the potential to alter definitive treatment itself; for instance, by identifying a subgroup of patients with a suboptimal early response to treatment who might benefit from timely CCRT dose escalations or early conversion to surgery.

Despite the statistically significant correlations between volume parameters and local outcome in our series, we believe that the discriminatory performances of these parameters are suboptimal for use towards therapeutic intensifications. In this respect, for primary sites combined, volumetric thresholds at diagnosis and 2 weeks into CCRT that achieved the maximum accuracy tended to produce intermediate PPVs. Consequently, we cannot recommend that patients in the unfavourable prognostic group should receive more aggressive therapy, as this would be unnecessary in approximately 40% of these patients and would increase their treatment-related morbidity. Furthermore, although the high-to-intermediate negative predictive value (NPV) and specificity of volumetric data suggest that a subgroup of patients that will achieve local control can be identified before completion of CCRT, this information cannot be used to improve their already excellent local outcomes; however, we acknowledge that there might be other benefits in terms of reducing morbidity through CCRT dose reductions, as well as less aggressive post-therapeutic monitoring of the primary site.

Unsurprisingly, volume parameters 6 weeks after completion of CCRT were more predictive of local outcome than pre- or intratherapeutic parameters. In this respect, optimal thresholds achieved higher specificities, NPVs and PPVs, although sensitivities were relatively low (26% and 58%). Despite the low sensitivities, these thresholds might be useful in clinical practice to identify a subset of patients that is likely to have local failure on the basis of the 6-week scan; for these patients more aggressive monitoring is justified.

For hypopharyngeal tumours, statistically significant correlations between volume parameters at all time-points and local outcome were found, except for the percentage change in volume 2 weeks into CCRT (although this might be due to type II error). Furthermore, volumetric parameters at all three time-points in the study achieved relatively high sensitivity and PPVs. These findings are encouraging, as the hypopharynx is difficult to assess non-invasively; early volumetry might be useful for identifying poor responders who could benefit from therapeutic intensifications as well as from early examinations under anaesthesia and guided biopsies. Of interest, other studies have documented higher optimal cut-offs for pre-treatment hypopharyngeal volumetry to predict failure, between 19 cm³ and 30 cm³, than the present study (11.0 cm³) [13, 14]. Differences in optimal thresholds between studies might reflect differences in treatment efficacy, although they might also be explained by systematic differences in tumour measurement. In this respect, these other studies performed volumetry using CT as opposed to MRI. Because CT has inferior soft-tissue contrast compared with MRI, these other studies might have overestimated

primary tumour volumes owing to the inadvertent inclusion of peritumoural inflammation. Indeed, the authors of one study made no attempt to delineate tumour from surrounding oedema when performing volumetry [13].

Subgroup statistical or ROC analysis was not possible for other primary sites, although none of the 13 tonsillar tumours in our series had local failure compared with a 44% failure rate for hypopharyngeal tumours, and tumours at these sites were not significantly different in size. This is unsurprising as tonsillar tumours are more radiosensitive than several other head and neck subsites [17]; studies have also shown site-related differences in tumour molecular biology that correlate with tumour aggressiveness [18–21]. Consequently, future studies of volumetric data in HNSCC should be site specific.

Several aspects of the study design should be addressed. Tumour volumes were calculated by a summation-of-areas technique, which, although laborious, is more accurate than multiplying orthogonal tumour diameters and is particularly important for irregularly shaped tumours such as those found in HNSCC. For this reason, we did not evaluate unidimensional or bidimensional indices, such as those recommended in the World Health Organization (WHO) and Response Evaluation Criteria in Solid Tumours (RECIST) guidelines for reporting tumour response [22, 23]. The total number of patients in the study was limited owing to patient enrolment from another recent prospective study, although this was necessary because MRI is only performed 2 weeks into CCRT for research purposes at our institution. As a consequence of the limited sample size, subgroup analysis of different primary sites was also limited. We acknowledge that a 1 year minimum follow-up of patients in this study might have been too short to detect all eventual primary site failures. Finally, we did not assess the influence of primary volume on disease-specific and actuarial survival as the study has not matured. Future studies can address these limitations by increasing the study size and follow-up interval and by evaluating other tumour sites separately.

In conclusion, the present study shows that primary site volumetry performed pre-treatment, early during treatment and 6 weeks after treatment with definitive CCRT for HNSCC correlates with primary site outcome. Of these time points, the 6-week post-CCRT scan is most predictive of local failure and, although it cannot be used to modify treatment, it could be important for selecting patients for more aggressive monitoring. Primary tumour volumetry performed 2 weeks into treatment produces intermediate PPVs for all sites combined, suggesting that volumetry is unsuitable for identifying poor responders who would benefit from CCRT intensifications or conversion to definitive surgery. By contrast, high PPVs and sensitivities were achieved at 2 weeks for hypopharyngeal tumours, although there were no additional local failures identified using optimal discriminatory thresholds at this time point compared with pre-treatment volumetry. Tumour volumetry 2 weeks into CCRT achieves moderate to high specificity, suggesting that a subgroup of good responders could be identified at this early stage.

Acknowledgements

Data for part of this study were obtained from a prospective study partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (project number 4300/104).

References

- Garden AS, Harris J, Trotti A, Jones CU, Carrascosa L, Cheng JD, et al. Long-term results of concomitant boost radiation plus concurrent cisplatin for advanced head and neck carcinomas: a phase II trial of the radiation therapy oncology group (RTOG 99-14). *Int J Radiat Oncol Biol Phys* 2008;71:1351–5.
- McCollum AD, Burrell SC, Haddad RI, Norris CM, Tishler RB, Case MA, et al. Positron emission tomography with 18F-fluorodeoxyglucose to predict pathologic response after induction chemotherapy and definitive chemoradiotherapy in head and neck cancer. *Head Neck* 2004;26:890–6.
- Porceddu SV, Jarmolowski E, Hicks RJ, Ware R, Weih L, Rischin D, et al. Utility of positron emission tomography for the detection of disease in residual neck nodes after (chemo)radiotherapy in head and neck cancer. *Head Neck* 2005;27:175–81.
- Yao M, Graham MM, Hoffman HT, Smith RB, Funk GF, Graham SM, et al. The role of post-radiation therapy FDG PET in prediction of necessity for post-radiation therapy neck dissection in locally advanced head-and-neck squamous cell carcinoma. *Int J Radiat Oncol Biol Phys* 2004;59:1001–10.
- Green FL, Page DL, Fleming ID. *AJCC Cancer Staging Manual*. 6th edition. New York: Springer, 2002.
- Ljumanovic R, Langendijk JA, Hoekstra OS, Knol DL, Leemans CR, Castelijns JA. Pre- and post-radiotherapy MRI results as a predictive model for response in laryngeal carcinoma. *Eur Radiol* 2008;18:2231–40.
- Hamilton S, Venkatesan V, Matthews TW, Lewis C, Assis L. Computed tomographic volumetric analysis as a predictor of local control in laryngeal cancers treated with conventional radiotherapy. *J Otolaryngol* 2004;33:289–94.
- Pameijer FA, Hermans R, Mancuso AA, Mendenhall WM, Parsons JT, Stringer SP, et al. Pre- and post-radiotherapy computed tomography in laryngeal cancer: imaging-based prediction of local failure. *Int J Radiat Oncol Biol Phys* 1999;45:359–66.
- Hoebbers FJ, Pameijer FA, de Bois J, Heemsbergen W, Balm AJ, Schornagel JH, et al. Prognostic value of primary tumor volume after concurrent chemoradiation with daily low-dose cisplatin for advanced-stage head and neck carcinoma. *Head Neck* 2008;30:1216–23.
- Plataniotis GA, Theofanopoulou ME, Kalogera-Fountzila A, Haritanti A, Ciuleanu E, Ghilezan N, et al. Prognostic impact of tumor volumetry in patients with locally advanced head-and-neck carcinoma (non-nasopharyngeal) treated by radiotherapy alone or combined radiochemotherapy in a randomized trial. *Int J Radiat Oncol Biol Phys* 2004;59:1018–26.
- Gilbert RW, Birt D, Shulman H, Freeman J, Jenkin D, MacKenzie R, et al. Correlation of tumor volume with local control in laryngeal carcinoma treated by radiotherapy. *Ann Otol Rhinol Laryngol* 1987;96:514–18.
- Lee WR, Mancuso AA, Saleh EM, Mendenhall WM, Parsons JT, Million RR. Can pretreatment computed tomography findings predict local control in T3 squamous cell carcinoma of the glottic larynx treated with radiotherapy alone? *Int J Radiat Oncol Biol Phys* 1993;25:683–7.
- Chen SW, Yang SN, Liang JA, Lin FJ, Tsai MH. Prognostic impact of tumor volume in patients with stage III-IVA hypopharyngeal cancer without bulky lymph nodes treated with definitive concurrent chemoradiotherapy. *Head Neck* 2009;31:709–16.
- Tsou YA, Hua JH, Lin MH, Tsai MH. Analysis of prognostic factors of chemoradiation therapy for advanced hypopharyngeal cancer — does tumor volume correlate with central necrosis and tumor pathology? *ORL J Otorhinolaryngol Relat Spec* 2006;68:206–12.
- Mancuso AA, Mukherji SK, Schmalzuss I, Mendenhall W, Parsons J, Pameijer F, et al. Preradiotherapy computed tomography as a predictor of local control in supraglottic carcinoma. *J Clin Oncol* 1999;17:631–7.
- Hermans R, Van den Bogaert W, Rijnders A, Baert AL. Value of computed tomography as outcome predictor of supraglottic squamous cell carcinoma treated by definitive radiation therapy. *Int J Radiat Oncol Biol Phys* 1999;44:755–65.
- Mukherji SK, Schmalzuss IM, Castelijns J, Mancuso AA. Clinical applications of tumor volume measurements for predicting outcome in patients with squamous cell carcinoma of the upper aerodigestive tract. *AJNR Am J Neuroradiol* 2004;25:1425–32.
- van den Broek GB, Wildeman M, Rasch CR, Armstrong N, Schuurin E, Begg AC, et al. Molecular markers predict outcome in squamous cell carcinoma of the head and neck after concomitant cisplatin-based chemoradiation. *Int J Cancer* 2009;124:2643–50.
- Lukits J, Timar J, Juhasz A, Dome B, Paku S, Repassy G. Progression difference between cancers of the larynx and hypopharynx is not due to tumor size and vascularization. *Otolaryngol Head Neck Surg* 2001;125:18–22.
- Liu M, Lawson G, Delos M, Jamart J, Chatelain B, Remacle M, et al. Prognostic value of cell proliferation markers, tumour suppressor proteins and cell adhesion molecules in primary squamous cell carcinoma of the larynx and hypopharynx. *Eur Arch Otorhinolaryngol* 2003;260:28–34.
- Homma A, Furuta Y, Oridate N, Nakano Y, Kohashi G, Yagi K, et al. Prognostic significance of clinical parameters and biological markers in patients with squamous cell carcinoma of the head and neck treated with concurrent chemoradiotherapy. *Clin Cancer Res* 1999;5:801–6.
- Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981;47:207–14.
- Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009;45:228–47.