


RESEARCH ARTICLE

Open Access



# Three-dimensional (3D) magnetic resonance volume assessment and loco-regional failure in anal cancer: early evaluation case-control study

Hema Sekhar<sup>1\*</sup> , Rohit Kochhar<sup>2</sup>, Bernadette Carrington<sup>2</sup>, Thomas Kaye<sup>3</sup>, Damian Tolan<sup>3</sup>, Mark P. Saunders<sup>4</sup>, Matthew Sperrin<sup>5</sup>, David Sebag-Montefiore<sup>6</sup>, Marcel van Herk<sup>1†</sup> and Andrew G. Renehan<sup>1†</sup>

## Abstract

**Background:** The primary aim was to test the hypothesis that deriving pre-treatment 3D magnetic resonance tumour volume (mrTV) quantification improves performance characteristics for the prediction of loco-regional failure compared with standard maximal tumour diameter (1D) assessment in patients with squamous cell carcinoma of the anus undergoing chemoradiotherapy.

**Methods:** We performed an early evaluation case-control study at two UK centres (2007–2014) in 39 patients with loco-regional failure (cases), and 41 patients disease-free at 3 years (controls). mrTV was determined using the summation of areas method (Vol<sub>sum</sub>). Reproducibility was assessed using intraclass concordance correlation (ICC) and Bland-Altman limits of agreements. We derived receiver operating curves using logistic regression models and expressed accuracy as area under the curve (ROC<sub>AUC</sub>).

**Results:** The median time per patient for Vol<sub>sum</sub> quantification was 7.00 (inter-quartile range, IQR: 0.57–12.48) minutes. Intra and inter-observer reproducibilities were generally good (ICCs from 0.79 to 0.89) but with wide limits of agreement (intra-observer: –28 to 31%; inter-observer: –28 to 46%). Median mrTVs were greater for cases (32.6 IQR: 21.5–53.1 cm<sup>3</sup>) than controls (9.9 IQR: 5.7–18.1 cm<sup>3</sup>,  $p < 0.0001$ ). The ROC<sub>AUC</sub> for mrT-size predicting loco-regional failure was 0.74 (95% CI: 0.63–0.85) improving to 0.82 (95% CI: 0.72–0.92) when replaced with mrTV (test for ROC differences,  $p = 0.024$ ).

**Conclusion:** Preliminary results suggest that the replacement of mrTV for mrT-size improves prediction of loco-regional failure after chemoradiotherapy for squamous cell carcinoma of the anus. However, mrTV calculation is time consuming and variation in its reproducibility are drawbacks with the current technology.

**Keywords:** Anal cancer, Tumour volume, Loco-regional failure, Prognosis

\* Correspondence: [hsekhar1982@gmail.com](mailto:hsekhar1982@gmail.com)

†Marcel van Herk and Andrew G Renehan are joint senior authors.

<sup>1</sup>Division of Molecular & Clinical Cancer Sciences, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Wilmslow Road, Manchester M20 4BX, UK

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Chemo-radiotherapy (CRT) is the primary treatment for patients with squamous cell carcinoma of the anus (SCCA) [1]. The key aim of treatment is to achieve loco-regional control whilst maintaining sphincter preservation. However, up to a quarter of those treated report loco-regional failure (LRF) within 3 years [2]. These patients may be offered a second attempt of cure with salvage surgery, but this is a radical operation associated with significant morbidity and poor long-term outcomes [3].

Until recently, CRT for SCCA adopted a broad “one size fits all” strategy, typically comprising 5-fluorouracil and mitomycin concurrent with 50 to 55 Gy radiotherapy, despite a wide spectrum of the loco-regional disease stages at presentation. Stratified approaches are now being evaluated, as exemplified by the UK PLATO (Personalising Radiotherapy Dose in anal cancer, ISRCTN88455282) trial [4], a single protocol ‘platform’ comprising the ACT3, 4 and 5 trials with the aim of personalising radiotherapy dose across the disease spectrum. Presently, pre-treatment MR assessment of T-size (mrT-size), together with nodal status, are the key determinant of treatment stratification but this strategy has imperfections. For example, there is variation in maximum tumour dimensions determined using T2 weighting imaging [5]. One-dimensional tumour diameter may not adequately represent tumour biology that tends to be infiltrative in nature and irregularly shaped. Furthermore, based on using 7th edition AJCC staging (where the cut-off from T2 to T3 is 5 cm), Gunderson et al. [6] illustrated that there is a considerable ‘jump-up’ in 3-year LRF rates between T2N0 (10%) and T3N0 (22%), suggesting that T-size alone may not optimally capture the heterogeneity radio-resistance and risk of LRF.

An alternative imaging approach to pre-treatment staging is magnetic resonance quantification of tumour volume (mrTV). This has been explored in other tumour sites, such as in head and neck [7–10] and lung cancer [11–14]. However, methods and results have been inconsistent. Some studies use 3D quantification techniques while other have used semi-automated estimates of ellipsoid volume derived from 1D tumour dimensions. It is unclear whether this latter approach is a valid estimate.

The primary aim was to test the hypothesis that deriving pre-treatment 3D mrTV quantification improves performance characteristics for the prediction of LRF compared with standard maximal tumour diameter (1D) assessment in patients with SCCA undergoing CRT. We chose to quantify 3D mrTV using a summation of areas method, as conceptually, this better captures the complex 3-dimensional nature of an anal tumour compared with estimation methods derived from 1D tumour dimensions. But in turn, this method may be labour intensive and there may be considerable variation in reproducibility. Thus, secondary aims were to evaluate whether 3D mrTV quantification

directly measured on scans by the summation of area method is reproducible; and whether volume estimation derived from less labour intensive 1D tumour dimensions is a valid estimation of tumour volume.

## Methods

In accordance with the CRUK/EORTC imaging biomarker consensus statement, this was a two-centre Domain 2 validation study evaluating performance characteristics, reproducibility and whether the biomarker is ‘fit for purpose’ [15].

### Patients and treatment

We performed a case-control study at two UK centres, the Christie NHS Foundation Trust, Manchester, and Leeds Teaching Hospitals Trust, Leeds (LTHT). Patients were included if they had histologically confirmed SCCA; T1 to T4 disease (AJCC 7th edition) [16]; and had received CRT with curative intent for non-metastatic disease. For the control group, patients were free of LRF for at least 3 years follow-up. Patients with histologies other than SCC were excluded, as were patients where T-stage was undeterminable (Tx disease,) as mrTV and mrT-size parameters could not be quantified.

All patients were treated between 2007 and 2014 prior to the introduction of IMRT (Intensity Modulated Radiotherapy). The treatment protocol followed that used in the ACT II trial [17] – namely, radiotherapy of 50.4 Gy was delivered over 5.5 weeks with a two phase technique, without a mandatory break. Phase 1 included 30.6 Gy in 17 daily fractions with non-conformal rectangular parallel-opposed fields. Phase 2 required conformal planning and delivered 19.8 Gy in 11 daily fractions over 15 days to the primary tumour with a 3 cm margin and any involved lymph nodes. Chemotherapy regimens were administered concurrently with radiotherapy as either: mitomycin-C (MMC) 12 mg/m<sup>2</sup> on day 1, and continuous infusion of 5-fluorouracil (5-FU) 1000 mg/m<sup>2</sup> on days 1–4 and days 29–32.

### Selection for case-control study

From the retrospective two-centre clinical databases, all 40 patients with LRF from 2007 to 2014 undergoing CRT and with measureable anal tumours were selected as cases (one outlier volume later excluded). Forty-one patients without LRF at 3 years were controls. Control selection was at random from those patients in the databases with available MR images satisfying criteria listed next.

### Tumour volume quantification

All tumour quantification used routinely collected pre-treatment MR imaging, performed on a 1.5 Telsa MR employing optimal pelvic phase-array body coil (acquisition protocols are detailed in the supplementary material, Table S1). For inclusion, scans had to meet the following criteria:

(i) include a small field of view high resolution T2-weighted (T2W) sequence in the axial plane as minimum, with a slice thickness  $\leq 4$  mm; (ii) field of view extending above and below the tumour in two orthogonal planes to allow complete tumour assessment including TV quantification and assessment of T-size; and (iii) where the tumour required more than one series to assess the whole tumour, then these series had to overlap sufficiently such that the entire tumour was imaged and could be quantifiable.

To determine mrT-size, three primary orthogonal measurements of the maximal diameters were taken in the anterior-posterior (AP), left-right (LR) and cranio-caudal (CC) planes along the axis of the tumour measured on the high resolution T2W images. The AP and LR diameters were recorded on the axial plane at the point of maximal dimension. The longest diameter was noted and the next dimension was taken at an axis perpendicular to the above. CC dimension was measured in either the coronal or sagittal plane. The largest of these three diameters was considered to be the tumour size (cm). Assessors (RK and BC) were blind to LRF status.

mrTV measurements were performed using World-Match (in-house written software from MvH [18]) that allowed simultaneous contouring on several sequences of different planes. Pre-treatment MR images were imported in anonymised DICOM format and the primary tumour was manually contoured. Delineations were checked and adjusted accordingly using coronal and sagittal planes. All contiguous areas of tumour were contoured together including nodal masses that had coalesced with the tumour or contiguous areas of extramural vascular invasion (EMVI). This was required to account for difficulties in defining a plane between the entities and to allow consistency of approach. Separate or discrete nodal volumes were not included. The assessor (HS) was blind to LRF status.

For the main analysis, mrTV was derived by using a summation of areas method ( $\text{Vol}_{\text{sum}}$ ), which sums-up the area contoured on serial image slices while taking into account the distance between the slices; i.e. the slice thickness of the scans (Fig. 1a & b). The time taken to contour the TV on MR images, was recorded for the first 22 patients ('training') and compared with the remaining patients.

#### Intra- and inter-observer variability

Intra- and inter-observer variability of mrTV quantification was assessed in ten (randomly selected) patients and compared using intra-class concordance correlations (ICC) with scores  $< 0.5$  representing poor;  $0.5$  to  $< 0.75$  moderate;  $0.75$  to  $< 0.9$  good; and  $\geq 0.9$  excellent agreement [19]. 95% confidence intervals were derived from z-transformations. Bland-Altman plots were employed to further assess agreement [20]. We standardised all plots so that y-axis (mean

difference between measure modalities) ranges were equivalent to those of the x-axis (mean value from both measure modalities). We then examined each plot for: (i) mean values closeness to zero; (ii) levels of agreement; and (iii) that the pattern across the range of means was proportionate i.e. evaluating for trends across the range. For levels of agreement, we expressed as percentage of the x-axis range of values and reported as these as 'wide' if the limits fell outside  $\pm 10\%$  of the average mean difference.

#### Ellipsoid and elliptical cylinder volume estimation

We tested whether volume estimation derived from 1D tumour dimensions is a valid estimate of TV, using ellipsoid and elliptical cylinder equations and measured 1D diameters (Fig. 1c & d), where  $d_1$ ,  $d_2$  and  $d_3$  are the maximal diameters of the primary tumour measured in the AP, LR and CC planes. The ellipsoid equation was:  $\frac{4}{3} \times \pi \times \frac{d_1}{2} \times \frac{d_2}{2} \times \frac{d_3}{2}$ ; the elliptical cylinder equation was:  $(\pi \times d_1 \times d_2 \times d_3)/4$ . These were assessed for reproducibility as above. We also evaluated for accuracies compared with  $\text{Vol}_{\text{sum}}$  using the  $\text{ROC}_{\text{AUC}}$  method (we expected ellipsoid and elliptical to be equivalent as they are derived from the same parameters).

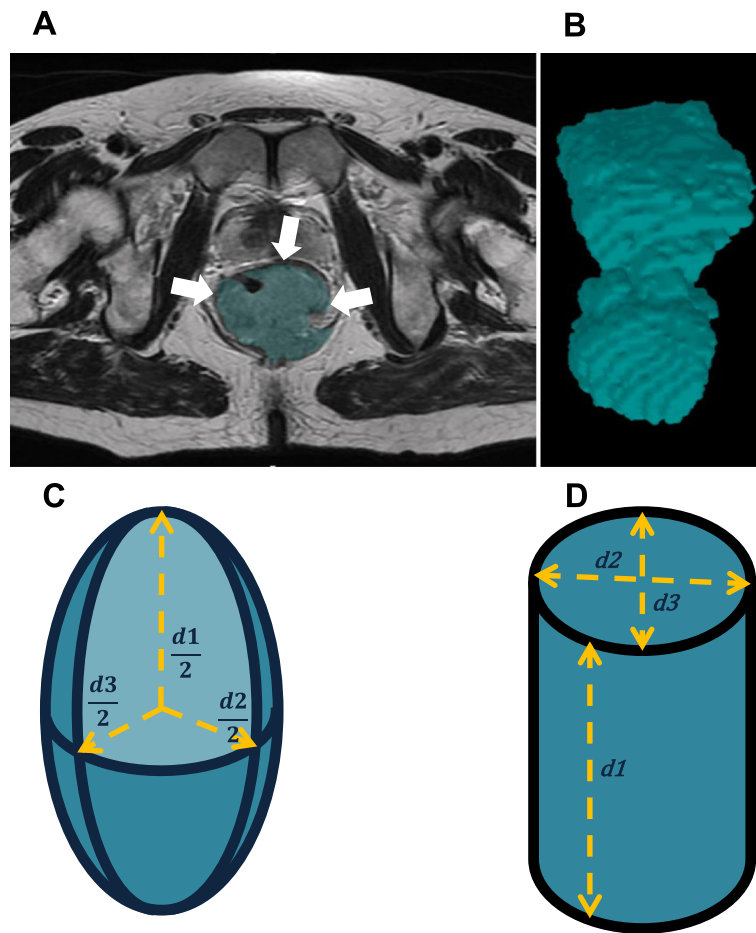
#### Statistical analysis

Stata software, version 14 (Stata Corp., Tx, USA) was used for all statistical analyses. Continuous data were summarised as medians with inter-quartile ranges (IQR) and categorical data were presented as proportions. Comparisons were by Chi-square and Fisher's Exact tests and non-parametric Mann-Whitney U tests, respectively.

Assessment of the discriminatory potential of the different tumour quantifications was tested with receiver operator characteristic curves (ROCs) and estimation of accuracy using the AUC and compared against each other using the method of DeLong et al. [21]. Multivariable models used logistic regression, and derived  $\text{ROC}_{\text{AUC}}$  using post-estimation commands.

In our power calculation, we posited that 10% would be a meaningful clinical difference for areas under the curve  $\text{ROC}_{\text{AUC}}$ . We added 3% as case-control studies tend to overestimate performance characteristics [22]. Thus, we concluded that 42 cases of LRF and 42 controls of non-LRF patients would be required to reach  $\text{ROC}_{\text{AUC}}$  difference of 13% at  $\alpha < 0.001$ .

We utilised other indicators of performance characteristics, using the methods described by Pencina et al. [23] which derives two characteristics – the Integrated Discriminatory Improvement (IDI), an index of improvements in sensitivity relative to specificity, and Net Reclassification Improvement (NRI), an index of net change in events versus non-events detected, which in turn focuses on medical decision making.



**Fig. 1** Method of quantification of tumour volume: **a** pelvic axial section of MR images depicting an anal canal tumour contoured in blue (arrows); **b** sequential contours rendered to create 3D representation of the tumour, demonstrating its irregular and complex shape; and **c** estimation of volume using orthogonal tumour size measurements under the assumption that the tumour volume approximates an ellipsoid. **d** estimation of volume using orthogonal tumour size measurements under the assumption that the tumour volume approximates an elliptical cylinder. Figures **(a)** and **(b)** were created using World-Match (inhouse written software). Figures **(c)** and **(d)** were created using Microsoft Office Powerpoint

## Results

### Patient characteristics

Eighty-one patients were initially included in the study. One case with a tumour volume of 652 cm<sup>3</sup> was excluded as an extreme outlier, leaving 80 patients as characterised in Table 1. Sixty-one patients were treated at The Christie and 19 at LTHT. The cases and controls were well-balanced for age and gender. As expected, the LRF group had more patients presenting with T3/4 disease compared with the non-LRF group and more patients presenting with node positivity. There were no differences noted in these baseline variables between the patients from The Christie and LTHT (Table S2).

### Volume quantification

mrTV measurements took a median of 7 min, but with wide variation from less than 1 min to over an hour. Contours

performed in the 22 preliminary patients took longer than subsequent determinations [median: 10.5 (IQR: 7.2–14.3) minutes versus 5.6 (IQR: 4.3–12.2) minutes,  $p = 0.018$ ].

The median volume in all 80 patients determined by summation of areas was 20.1 (IQR: 9.1–39.0) cm<sup>3</sup>. The median mrTV increased with mrT-stage as follows: T1, 5.7 cm<sup>3</sup>; T2, 9.1 (IQR: 4.6–16.7) cm<sup>3</sup>; T3, 38.4 (IQR: 21.6–50.2) cm<sup>3</sup>; T4, 34.5 (IQR: 18.1–66.5) cm<sup>3</sup>,  $p < 0.001$  (Figure S1).

### Reproducibility

Reproducibility was generally good. Intra-observer ICC scores were 0.89 (95% CI: 0.75 to 1.00); inter-observer ICC scores were 0.79 (95% CI: 0.55 to 1.00). Agreement was further explored with Bland-Altman plots, which revealed wide limits of agreement from - 7.93 cm<sup>3</sup> (- 28%) to 8.47 cm<sup>3</sup> (31%) for intra-observer variability and -

**Table 1** Patient characteristics of those with and without locoregional failure

	Controls Non-LRF	Cases LRF	p value
<b>N</b>	41	39	
<b>Gender</b>			
Men (%)	14 (34)	16 (41)	0.525 <sup>a</sup>
Women (%)	27 (66)	23 (59)	
<b>Median age (IQR), years</b>	60 (53–69)	57 (50–69)	0.563 <sup>b</sup>
<b>mrT-stage (%)</b>			0.018 <sup>c</sup>
T1	1 (2)	0	
T2	24 (59)	11 (28)	
T3	9 (22)	16 (41)	
T4	7 (17)	12 (31)	
<b>mr Nodal status</b>			
LN <sup>-</sup> (%)	26 (63)	14 (36)	0.014 <sup>a</sup>
LN <sup>+</sup> (%)	15 (37)	25 (64)	

LRF Locoregional Failure, non-LRF Without evidence of locoregional failure after 3-years follow-up, IQR Interquartile Range, LN+ Nodal Involvement; <sup>a</sup>Chi-square test; <sup>b</sup>Mann-Whitney U Test; <sup>c</sup>Fisher's Exact Test

7.65 cm<sup>3</sup> (- 28%) to 12.59 cm<sup>3</sup> (46%) for inter-observer variability (Fig. 2).

**Tumour size versus volume**

Median mrT-size was larger in cases compared with controls [5.2 (IQR: 4.3–6.2) cm versus 3.6 (IQR: 2.9–5.2) cm, p = 0.0004]. Similarly, mrTV determined by summation of

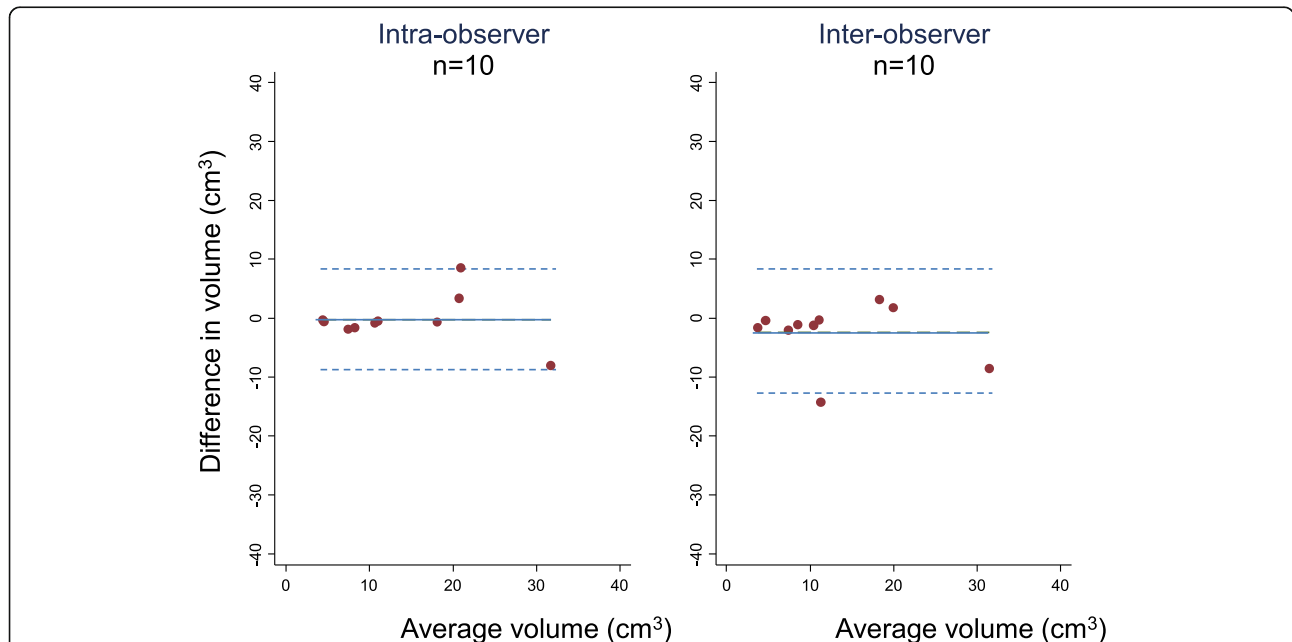
areas was greater in cases compared with controls [32.6 (IQR: 21.5–53.1) cm<sup>3</sup> versus 9.9 (IQR: 5.7–18.1) cm<sup>3</sup>, p = 0.0001] (Fig. 3a).

The ROC<sub>AUC</sub> for mrT-size predicting LRF was 0.74 (95% CI: 0.63–0.85). This improved to 0.82 (95% CI: 0.72–0.92) when mrTV replaced mrT-size (p = 0.024) (Fig. 3b). As nodal status is a predictor for LRF [2], we adjusted for this in two separate multivariable logistic regression models for mrT and mrTV. The ROC<sub>AUC</sub> did not materially change (Table 2).

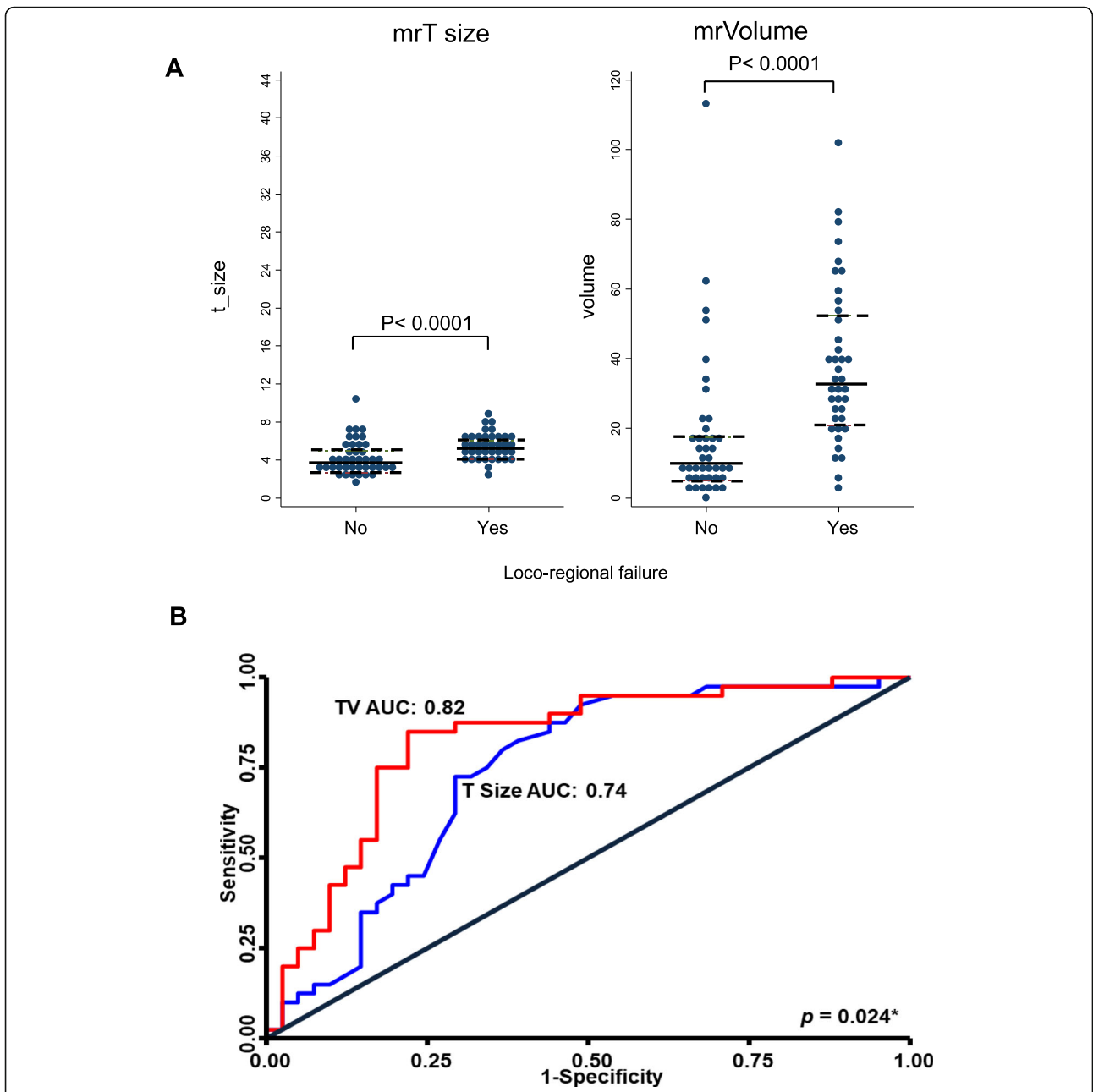
**Ellipsoid and elliptical cylinder estimates**

The median volumes estimated from the ellipsoid equation was 20.6 (IQR: 8.6–44.7) cm<sup>3</sup>. That for volume estimated from elliptical cylinder equation was higher at 30.8 (IQR: 13.0–67.0) cm<sup>3</sup>. Agreement was excellent between Vol<sub>sum</sub> and volume determined by the ellipsoid eq. [ICC scores =0.91 (95% CI: 0.86 to 0.94)] and good between Vol<sub>sum</sub> and volume determined by the elliptical cylinder eq. [ICC scores =0.73 (95% CI: 0.66 to 0.80)]. From the Bland-Altman plots, the limits of agreement for Vol<sub>sum</sub> and volume determined by the ellipsoid equation were from - 20.4 (- 16%) to 20.1 (17%). For Vol<sub>sum</sub> versus volume determined by the elliptical cylinder, there was a disproportionate trend with increasing volume (Figure S2).

The ROC<sub>AUC</sub> for mrTV determined by the ellipsoid volume estimate was 0.79 (95% CI: 0.68 to 0.89); that for the elliptical cylinder volume estimate was identical (Table S3).



**Fig. 2** Bland-Altman plots demonstrating intra- and inter-observer (modest) agreement of tumour volume quantification in 10 patients: **a** average mean difference of -0.27 cm<sup>3</sup> and limits of agreement from -7.93 cm<sup>3</sup> (- 28%) to 8.47 cm<sup>3</sup> (31%) for intra-observer variability; and **b** average mean difference -2.5 cm<sup>3</sup> and limits of agreement -7.65 cm<sup>3</sup> (- 28%) to 12.59 cm<sup>3</sup> (46%) for inter-observer variability. Figures were created using Stata software, version 14



**Fig. 3 a** Dotplots demonstrating the association of tumour volume and T-size with loco-regional failure (LRF) with the patient with the large volume outlier removed from both analyses: The horizontal black lines represent the median tumour quantification values; the dashed black lines represent the limits of the interquartile range. The y-axis has been standardised so that the medians in the controls are aligned **(b)**: Receiver Operator Characteristic curves comparing the discriminatory performance of measured tumour volume and mrT-size. \*AUCs compared by the method of DeLong et al. Figures were created using Stata software, version 14

**IDI and NRI**

For mrTV (measured as Vol<sub>sum</sub>) versus mrT-size, the IDI was 9.6% (standard error, se: 3.0%), which was statistically different ( $p = 0.0015$ ). By contrast, the NRI was potentially large at 34% (se: 22%), but not statistically significant ( $p = 0.123$ ).

**Discussion**

**Summary of Main findings**

mrTV quantification determined by summation of areas may improve accuracy to predict for loco-regional failure when replacing mrT-size. However, this method is time consuming and variation in reproducibility are drawbacks

**Table 2** Univariable and multivariable AUCs from ROCs for outcome of loco-regional failure ( $N = 80$ )

Variable	Univariable	Multivariable <sup>a</sup>
	AUC (95% CI)	AUC
Gender	0.534 (0.427–0.642)	
Age (years)	0.463 (0.333–0.592)	
T-Size (cm)	0.731 (0.617–0.846)	0.738 <sup>b</sup>
Tumour Volume (cm <sup>3</sup> )	0.817 (0.720–0.915)	0.801 <sup>b</sup>
Nodal status	0.638 (0.531–0.744)	

CI Confidence Interval

<sup>a</sup>These are two variable models – the variable listed in the left-hand column plus nodal status

<sup>b</sup>No 95% CIs are returnable in these post-estimations

with current technology. The use of tumour volume estimates that utilise routinely obtained 1D tumour diameter measurements (ellipsoid and elliptical cylinder equations) had imperfect levels of agreement with volume determined by summation of areas but it is unknown whether this influences performance characteristics defined by AUC on ROC analyses.

### Context of other literature

To our knowledge, this is the first study to evaluate the performance characteristics of mrTV and LRF in SCCA. Many studies [24–29] in patients with SCCA have reported on TNM-related parameters as treatment predictors. Of the six published trials in this field, three [30–32] reported tumour predictive information, using either AJCC T-stage or a single one-dimension tumour size cut-off. In contrast, tumour biology occurs at a 3D level, with the chance of cure being dependent on factors such as increasing number of tumour clonogens that require sterilising and the extent of tumour hypoxia, both of which increase with the bulk of the tumour. Thus, tumour volume seems a ‘closer to real-life’ model for tumour biology.

MR imaging is the current standard of care for the pre-treatment staging of SCCA [33], providing high-resolution imaging. It is tempting to speculate that extension of this imaging platform to tumour volume quantification could improve treatment prediction. MR imaging additionally allows analysis of tumour heterogeneity, another parameter that may be a better representation of tumour biology and offer complementary information to current staging parameters. One study of forty patients [34] found tumour heterogeneity is associated with disease recurrence.

### Limitations and strengths

The study has several strengths. First, we used data from two well-characterised UK treatment centres with uniform treatment protocols. Second, to minimise the risk of heterogeneity in scanning parameters and quality over time and between centres, scan quality parameters were evaluated. Third, in-house written software was employed,

which allowed contouring in multiple planes, facilitating accurate volume quantification.

The study has several limitations. First, with only 80 patients, it is a relatively small sample size. Second, to enrich for events we used a case-control study design. This approach is biased in that it overestimates performance characteristics [22] and hence we powered using a very conservative  $\alpha < 0.001$ . Third, we fell just short of our target sample size. An improvement of 9% in the ROC<sub>AUC</sub> for mrTV over mrT-size was statistically significant at  $p = 0.024$ , and might be clinically meaningful. A similar 9% was noted with the IDI method, and while the NRI was potentially promising at 34%, it had much uncertainty. Fourth, as a retrospective study, there are potential biases and confounding. To mitigate against some of these, the assessors were blinded. Finally, the study recognises that ‘drawing’ tumour volume on MR is currently imprecise. It is challenging to differentiate peri-tumour oedema and fibrosis from tumour tissue and pathological validation of TV measurements was not possible. Indeed, tumour delineation is one of the most uncertain aspects of RT planning and as such, assessment of tumour volume for the purpose of prediction is subject to the same issues [35–39]. This is reflected in the moderate-good intra- and inter-observer concordance. In a parallel set of data (unpublished), we found there is modest intra- and inter-observer variability for T-size (ICC for intra-observer variability: 0.86; ICCs for inter-observer variability ranging from 0.73 to 0.82).

### Clinical implications and future research

The findings from this study suggest that 3D MR-based tumour volume quantification may enhance prediction of LRF risk over current MR-based methods. We speculate that TV better reflects tumour biology and could facilitate risk stratification to improve the precision of personalised treatment. The main limitation is that tumour ROI definition on MR is not currently routinely performed, and can be time consuming. MR-based RT planning, currently investigated in tumours such as rectal adenocarcinoma [40], may become routine in the future for SCCA, requiring tumour delineation for treatment planning which can then be reused as a prognostic factor. Techniques for automated contouring may also offer a time-sparing solution in the future, once developed and validated. Our hypothesis must now be evaluated – for example by, ‘piggy-backing’ mrTV definition onto ongoing trials.

Tumour regression grading has been found to predict for early local failure in SCCA [41]. Further work examining the role of measured tumour volume reduction following CRT may provide more prognostic information; however, differentiating tumour from fibrosis and oedema following CRT is a significant challenge.

PET-CT is increasingly used in the pre-treatment assessment of anal cancer and small studies demonstrate the

potential prognostic role of metabolic tumour volume (MTV) [42–45]. Large scale exploration of the prognostic role of MTV could be useful, and since PET-CT does not necessarily detect areas of tumour necrosis [46], which are unlikely to contain clonogenic cells, an important factor in the probability of cure, examining the predictive role of MTV relative to TV would be of interest.

Lastly, we acknowledge that looking at a single parameter to predict LRF is simplistic and it is more realistic that a number of patient, biological, tumour and radiological features each contribute to prognostic outcomes in anal cancer [47], and are likely to be incorporated into future risk-adapted treatment strategies. A multiparametric prognostic model incorporating all of these features would be ideal, but investigating this issue is outside of the remit of this preliminary study. Larger-scale studies that facilitate this and are in progress. We have, however, identified mrTV to be a potentially useful prognostic tool, to be studied within such future models.

## Conclusions

These preliminary results suggest that 3D MR-based tumour volume quantification may improve prediction of LRF in patients with SCCA following CRT over currently employed measurements. Further work will be required to refine ROI contouring, both in regards to facilitating a less time consuming and more reproducible process, and to explore the role of this tumour volume within multiparametric prognostic models for LRF in SCCA.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12885-020-07613-7>.

### Additional file 1.

## Abbreviations

3D: Three Dimensional; 5-FU: 5-Fluorouracil; AJCC: American Joint Committee on Cancer; AP: Anterior-posterior; AUC: Area Under the Curve; CC: Cranio-caudal; CI: Confidence Interval; CM: Centimetre; CRT: Chemoradiotherapy; CRUK: Cancer Research United Kingdom; DICOM: Digital Imaging and Communications in Medicine; EMVI: Extra-mural Vascular Invasion; EORTC: European Organisation for Research and Treatment of Cancer; Gy: Gray; ICC: Intraclass Concordance Correlation; IDI: Integrated Discriminatory Improvement; IMRT: Intensity Modulated Radiotherapy; IQR: Inter-quartile Range; LR: Left-right; LRF: Loco-regional Failure; LTHT: Leeds Teaching Hospitals Trust; MMC: Mitomycin-C; MR: Magnetic Resonance; MTV: Metabolic Tumour Volume; NHS: National Health Service; NRI: Net Reclassification Improvement; PET-CT: Positron Emission Tomography - Computed Tomography; PLATO: Personalising Radiotherapy Dose in anal cancer; ROC: Receiver Operator Characteristic; ROI: Region Of Interest; RT: Radiotherapy; SCCA: Squamous Cell Carcinoma of the Anus; se: Standard Errors; T2W: T2-weighted; TNM: Tumour-Nodal-Metastases; TV: Tumour Volume; UK: United Kingdom; VOLsum: Volume from summation of areas

## Acknowledgements

AGR is supported by the Manchester NIHR Biomedical Research Centre (IS-BRC-1215-20007).

## Authors' contributions

HS, RK, MvH, DS-M and AGR contributed to conception and design. Collection and assembly of data were performed by HS, RK, BC, TK, and DT. Data analyses were performed by HS, MS and AGR. Data interpretation was performed by HS, RK, MPS, DS-M, MvH and AGR. All authors contributed to manuscript writing, final approval of manuscript, and are accountable for all aspects of the work.

## Funding

We acknowledge the generous funding from Bowel Disease Research Fund and the Royal College of Surgeons, England. The funders have had no involvement in the design of the study or collection, analysis, interpretation of data or in the writing of the manuscript.

## Availability of data and materials

The datasets supporting this article are stored in a secured research database and may be available upon presentation of formal approval.

## Ethics approval and consent to participate

This study was approved by the ethics committee of East of Scotland Research Ethics Service (16/ES/0800) and of West Midlands Solihull Research Ethics Committee (16/WM/0486). As the study data was handled in an anonymised format, individual consent from patients was not deemed necessary by the ethics committee. The study was performed in accordance with the Declaration of Helsinki. Participating sites included The Christie NHS Foundation Trust, Manchester, UK and St James' University Hospital, Leeds, UK.

## Consent for publication

Not applicable.

## Competing interests

AGR has received lecture honoraria from Merck Serona and Janssen-Cilag, and independent research funding and lecture honoraria from Novo Nordisk and Sanofi Pasteur MPS, unrelated to this study. All other authors have no conflicts of interest to declare.

## Author details

<sup>1</sup>Division of Molecular & Clinical Cancer Sciences, School of Medical Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Wilmslow Road, Manchester M20 4BX, UK. <sup>2</sup>Department of Radiology, The Christie NHS Foundation Trust, Manchester, UK. <sup>3</sup>Department of Clinical Radiology, St James' University Hospital, Leeds, UK. <sup>4</sup>Department of Clinical Oncology, The Christie NHS Foundation Trust, Manchester, UK. <sup>5</sup>Division of Informatics, Imaging and Data Sciences, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK. <sup>6</sup>Leeds Institute of Cancer & Pathology, University of Leeds, St James' University Hospital, Leeds, UK.

Received: 15 July 2020 Accepted: 3 November 2020

Published online: 30 November 2020

## References

- Glynn-Jones R, Renehan A. Current treatment of anal squamous cell carcinoma. *Hematol Oncol Clin North Am.* 2012;26(6):1315–50.
- Sekhar H, Malcomson L, Kochhar R, Sperrin M, Alam N, Chakraborty B, et al. Temporal improvements in loco-regional failure and survival in patients with anal cancer treated with chemo-radiotherapy: treatment cohort study (1990–2014). *Br J Cancer.* 2020;122(6):749–58.
- Renehan AG, Saunders MP, Schofield PF, O'Dwyer ST. Patterns of local disease failure and outcome after salvage surgery in patients with anal cancer. *Br J Surg.* 2005;92(5):605–14.
- PLATO\_trial. Personalising Anal cancer radioTherapy dOse – Incorporating ACT3, ACT4 and ACT5 [http://medhealth.leeds.ac.uk/info/430/solid\\_tumours/2210/plato](http://medhealth.leeds.ac.uk/info/430/solid_tumours/2210/plato) Accessed 20 Sept 2018.
- Prezzi D, Mandegaran R, Gourtsoyianni S, Owczarczyk K, Gaya A, Glynn-Jones R, et al. The impact of MRI sequence on tumour staging and gross tumour volume delineation in squamous cell carcinoma of the anal canal. *Eur Radiol.* 2018;28(4):1512–9.
- Gunderson LL, Moughan J, Ajani JA, Pedersen JE, Winter KA, Benson AB 3rd, et al. Anal carcinoma: impact of TN category of disease on survival, disease relapse, and colostomy failure in US gastrointestinal intergroup RTOG 98-11 phase 3 trial. *Int J Radiat Oncol Biol Phys.* 2013;87(4):638–45.



7. Baghi M, Mack MG, Hambek M, Bisdas S, Muerthel R, Wagenblast J, et al. Usefulness of MRI volumetric evaluation in patients with squamous cell cancer of the head and neck treated with neoadjuvant chemotherapy. *Head Neck*. 2007;29(2):104–8.
8. Chen C, Fei Z, Pan J, Bai P, Chen L. Significance of primary tumor volume and T-stage on prognosis in nasopharyngeal carcinoma treated with intensity-modulated radiation therapy. *Jpn J Clin Oncol*. 2011;41(4):537–42.
9. Feng M, Wang W, Fan Z, Fu B, Li J, Zhang S, et al. Tumor volume is an independent prognostic indicator of local control in nasopharyngeal carcinoma patients treated with intensity-modulated radiotherapy. *Radiat Oncol*. 2013;8(1):208.
10. Sze WM, Lee AW, Yau TK, Yeung RM, Lau KY, Leung SK, et al. Primary tumor volume of nasopharyngeal carcinoma: prognostic significance for local control. *Int J Radiat Oncol Biol Phys*. 2004;59(1):21–7.
11. Basaki K, Abe Y, Aoki M, Kondo H, Hatayama Y, Nakaji S. Prognostic factors for survival in stage III non-small-cell lung cancer treated with definitive radiation therapy: impact of tumor volume. *Int J Radiat Oncol Biol Phys*. 2006;64(2):449–54.
12. Etiz D, Marks LB, Zhou SM, Bentel GC, Clough R, Hernando ML, et al. Influence of tumor volume on prognosis in patients irradiated for non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys*. 2002;53(4):835–46.
13. Martel MK, Strawderman M, Hazuka MB, Turrisi AT, Fraass BA, Lichter AS. Volume and dose parameters for survival of non-small cell lung cancer patients. *Radiother Oncol*. 1997;44(1):23–9.
14. Willner J, Baier K, Caragiani E, Tschammler A, Flentje M. Dose, volume, and tumor control prediction in primary radiotherapy of non-small-cell lung cancer. *Int J Radiat Oncol Biol Phys*. 2002;52(2):382–9.
15. O'Connor JP, Aboagye EO, Adams JE, Aerts HJ, Barrington SF, Beer AJ, et al. Imaging biomarker roadmap for cancer studies. *Nat Rev Clin Oncol*. 2017;14(3):169–86.
16. Edge SB, Byrd DR, Compton CC, Fritz AG, Greene FL, Anus TA. *AJCC Cancer staging manual*. 7th ed. New York: Springer-Verlag; 2010. p. 165–73.
17. James RD, Glynne-Jones R, Meadows HM, Cunningham D, Myint AS, Saunders MP, et al. Mitomycin or cisplatin chemoradiation with or without maintenance chemotherapy for treatment of squamous-cell carcinoma of the anus (ACT II): a randomised, phase 3, open-label, 2x2 factorial trial. *Lancet Oncol*. 2013;14(6):516–24.
18. Wolthaus JW, Schneider C, Sonke JJ, van Herk M, Belderbos JS, Rossi MM, et al. Mid-ventilation CT scan construction from four-dimensional respiration-correlated CT scans for radiotherapy planning of lung cancer patients. *Int J Radiat Oncol Biol Phys*. 2006;65(5):1560–71.
19. Koo TK, Li MY. A guideline of selecting and reporting Intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155–63.
20. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307–10.
21. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics*. 1988;44(3):837–45.
22. Leeflang MM, Moons KG, Reitsma JB, Zwinderman AH. Bias in sensitivity and specificity caused by data-driven selection of optimal cutoff values: mechanisms, magnitude, and solutions. *Clin Chem*. 2008;54(4):729–37.
23. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008;27(2):157–72 discussion 207–12.
24. Svensson C, Goldman S, Friberg B. Radiation treatment of epidermoid cancer of the anus. *Int J Radiat Oncol Biol Phys*. 1993;27(1):67–73.
25. Tomaszewski JM, Link E, Leong T, Heriot A, Vazquez M, Chander S, et al. Twenty-five-year experience with radical chemoradiation for anal cancer. *Int J Radiat Oncol Biol Phys*. 2012;83(2):552–8.
26. Das P, Bhatia S, Eng C, Ajani JA, Skibber JM, Rodriguez-Bigas MA, et al. Predictors and patterns of recurrence after definitive chemoradiation for anal cancer. *Int J Radiat Oncol Biol Phys*. 2007;68(3):794–800.
27. Mai SK, Welzel G, Herrmann B, Bohrer M, Wenz F. Long-term outcome after combined radiochemotherapy for anal cancer - retrospective analysis of efficacy, prognostic factors, and toxicity. *Onkologie*. 2008;31(5):251–7.
28. Olivatto LO, Cabral V, Rosa A, Bezerra M, Santarem E, Fassizoli A, et al. Mitomycin-C- or cisplatin-based chemoradiotherapy for anal canal carcinoma: long-term results. *Int J Radiat Oncol Biol Phys*. 2011;79(2):490–5.
29. Roohipour R, Patil S, Goodman KA, Minsky BD, Wong WD, Guillem JG, et al. Squamous-cell carcinoma of the anal canal: predictors of treatment outcome. *Dis Colon Rectum*. 2008;51(2):147–53.
30. Bartelink H, Roelofs F, Eschwege F, Rougier P, Bosset JF, Gonzalez DG, et al. Concomitant radiotherapy and chemotherapy is superior to radiotherapy alone in the treatment of locally advanced anal cancer: results of a phase III randomized trial of the European organization for research and treatment of cancer radiotherapy and gastrointestinal cooperative groups. *J Clin Oncol*. 1997;15(5):2040–9.
31. Glynne-Jones R, Sebag-Montefiore D, Adams R, Gollins S, Harrison M, Meadows HM, et al. Prognostic factors for recurrence and survival in anal cancer: generating hypotheses from the mature outcomes of the first United Kingdom coordinating committee on Cancer research anal Cancer trial (ACT I). *Cancer*. 2013;119(4):748–55.
32. Gunderson L, Winter K, Ajani J, Pedersen J, Moughan J, Benson A, et al. Long-term update of US GI intergroup RTOG 98–11 phase III trial for anal carcinoma: survival, relapse, and colostomy failure with concurrent chemoradiation involving fluorouracil/mitomycin versus fluorouracil/cisplatin. *J Clin Oncol*. 2012;30(35):4344–51 Available from: <http://online.library.wiley.com/doi/10.1002/jco.22737>.
33. Glynne-Jones R, Nilsson PJ, Aschele C, Goh V, Peiffert D, Cervantes A, et al. Anal cancer: ESMO-ESSO-ESTRO clinical practice guidelines for diagnosis, treatment and follow-up. *Radiother Oncol*. 2014;111(3):330–9.
34. Owczarczyk K, Prezzi D, Cascino M, Kozarski R, Gaya A, Siddique M, et al. MRI heterogeneity analysis for prediction of recurrence and disease free survival in anal cancer. *Radiother Oncol*. 2019;134:119–26.
35. Hermans R, Feron M, Bellon E, Dupont P, Van den Bogaert W, Baert AL. Laryngeal tumor volume measurements determined with CT: a study on intra- and interobserver variability. *Int J Radiat Oncol Biol Phys*. 1998;40(3):553–7.
36. Hurkmans CW, Borger JH, Pieters BR, Russell NS, Jansen EP, Mijnheer BJ. Variability in target volume delineation on CT scans of the breast. *Int J Radiat Oncol Biol Phys*. 2001;50(5):1366–72.
37. Tai P, Van Dyk J, Yu E, Battista J, Stitt L, Coad T. Variability of target volume delineation in cervical esophageal cancer. *Int J Radiat Oncol Biol Phys*. 1998;42(2):277–88.
38. Vorwerk H, Beckmann G, Bremer M, Degen M, Dietl B, Fietkau R, et al. The delineation of target volumes for radiotherapy of lung cancer patients. *Radiother Oncol*. 2009;91(3):455–60.
39. Franco P, Arcadipane F, Trino E, Gallio E, Martini S, Iorio GC, et al. Variability of clinical target volume delineation for rectal cancer patients planned for neoadjuvant radiotherapy with the aid of the platform Anatom-e. *Clin Transl Radiat Oncol*. 2018;11:33–9.
40. Gwynne S, Mukherjee S, Webster R, Spezi E, Staffurth J, Coles B, et al. Imaging for target volume delineation in rectal cancer radiotherapy—a systematic review. *Clin Oncol (R Coll Radiol)*. 2012;24(1):52–63.
41. Kochhar R, Renehan AG, Mullan D, Chakrabarty B, Saunders MP, Carrington BM. The assessment of local response using magnetic resonance imaging at 3- and 6-month post chemoradiotherapy in patients with anal cancer. *Eur Radiol*. 2017;27(2):607–17.
42. Mohammadkhani Shali S, Schmitt V, Behrendt FF, Winz OH, Heinzl A, Mottaghy FM, et al. Metabolic tumour volume of anal carcinoma on (18)FDG PET/CT before combined radiochemotherapy is the only independent determinant of recurrence free survival. *Eur J Radiol*. 2016;85(8):1390–4.
43. Bazan JG, Koong AC, Kapp DS, Quon A, Graves EE, Loo BW Jr, et al. Metabolic tumor volume predicts disease progression and survival in patients with squamous cell carcinoma of the anal canal. *J Nucl Med*. 2013;54(1):27–32.
44. Cotter SE, Grigsby PW, Siegel BA, Dehdashti F, Malyapa RS, Fleshman JW, et al. FDG-PET/CT in the evaluation of anal carcinoma. *Int J Radiat Oncol Biol Phys*. 2006;65(3):720–5.
45. Trautmann TG, Zuger JH. Positron emission tomography for pretreatment staging and posttreatment evaluation in cancer of the anal canal. *Mol Imaging Biol*. 2005;7(4):309–13.
46. Krengli M, Millia ME, Turri L, Mones E, Bassi MC, Cannillo B, et al. FDG-PET/CT imaging for staging and target volume delineation in conformal radiotherapy of anal carcinoma. *Radiat Oncol*. 2010;5(1):10.
47. Jones CM, Goh V, Sebag-Montefiore D, Gilbert DC. Biomarkers in anal cancer: from biological understanding to stratified treatment. *Br J Cancer*. 2017;116(2):156–62.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.