Assessment of the Impact of Deformable Registration of Diagnostic MRI to Planning CT on GTV Delineation for Radiotherapy for Oropharyngeal Carcinoma in Routine Clinical Practice

Alice Taylor 1, Mehmet Sen 2* and Robin J. D. Prestwich 2,*

1 School of Medicine, Worsley Building, University of Leeds, Leeds LS2 9JT, UK; um13a2t@leeds.ac.uk
2 Department of Clinical Oncology, St. James’s University Hospital, Leeds Cancer Centre, Beckett Street, Leeds LS9 7TF, UK; mehmet.sen@nhs.net
* Correspondence: Robin.Prestwich@nhs.net; Tel.: +44-113-2067838; Fax: +44-113-2067886

Received: 24 October 2018; Accepted: 20 November 2018; Published: 24 November 2018

Abstract: Background: Aim of study was to assess impact of deformable registration of diagnostic MRI to planning CT upon gross tumour volume (GTV) delineation of oropharyngeal carcinoma in routine practice. Methods: 22 consecutive patients with oropharyngeal squamous cell carcinoma treated with definitive (chemo)radiotherapy between 2015 and 2016, for whom primary GTV delineation had been performed by a single radiation oncologist using deformable registration of diagnostic MRI to planning CT, were identified. Separate GTVs were delineated as part of routine clinical practice (all diagnostic imaging available side-by-side for each delineation) using: CT (GTV CT), MRI (GTV MR), and CT and MRI (GTV CTMR). Volumetric and positional metric analyses were undertaken using contour comparison metrics (Dice conformity index, centre of gravity distance, mean distance to conformity). Results: Median GTV volumes were 13.7 cm³ (range 3.5–41.7), 15.9 cm³ (range 1.6–38.3), 19.9 cm³ (range 5.5–44.5) for GTV CT, GTV MR and GTV CTMR respectively. There was no significant difference in GTV CT and GTV MR volumes; GTV CTMR was found to be significantly larger than both GTV MR and GTV CT. Based on positional metrics, GTV CT and GTV MR were the least similar (mean Dice similarity coefficient (DSC) 0.71, 0.84, 0.82 for GTV CT–GTV MR, GTV CTMR–GTV CT and GTV CTMR–GTV MR respectively). Conclusions: These data suggest a complementary role of MRI to CT to reduce the risk of geographical misses, although they highlight the potential for larger target volumes and hence toxicity.

Keywords: radiotherapy; deformable image registration; magnetic resonance imaging; oropharyngeal squamous cell carcinoma; gross tumour volume

1. Introduction

Intensity modulated radiotherapy (IMRT) provides steep dose gradients, facilitating lower doses to adjacent structures, but also has potential for geographical misses if tissues containing tumour are not accurately encompassed within the target volume [1–3]. However, the delineation of primary tumours is associated with considerable uncertainty [4,5], and variation in this process can have a major impact upon dose to primary tumour and to organs at risk (OAR) [6].

Recent international consensus guidelines for delineation of laryngeal, hypopharyngeal, oropharyngeal, and oral cavity primary tumour clinical target volumes (CTV) have recommended a volumetric approach based on the Danish Head and Neck Cancer Group (DAHANCA) approach of a ‘5 + 5 mm’ expansion margin (5 mm to high dose and a further 5 mm to intermediate or prophylactic dose) from gross tumour volume (GTV), with subsequent anatomical modification based upon tumour...
site and T stage [7]. This volumetric approach, with limited margins to CTV, places considerable importance upon accurate GTV delineation. As highlighted within these guidelines, the integration of clinical examination/endoscopy is a vital step in accurate GTV delineation; however, the use of the planning CT remains fundamental to primary tumour delineation.

CT imaging provides geometric accuracy and accurate dosimetric data based on tissue electron density, which is calibrated from Hounsfield units [8]. However, accurate delineation of an oropharyngeal tumour can be very challenging on CT, with limited soft tissue contrast and susceptibility to dental artefacts [9]. Consequently, CT-based delineation of oropharyngeal tumours is associated with considerable inter-observer [9,10] and intra-observer variability [11].

Integrating multi-modality imaging into the radiotherapy planning process provides an opportunity to improve upon delineation directly on CT, with side-by-side assessment of diagnostic imaging. Compared with CT, MRI provides superior soft tissue contrast and less susceptibility to dental artefacts [12], and diagnostic MRI is considered superior for visualising oropharyngeal tumours [7,9,12]. The use of MRI for delineation has been shown to reduce both inter-observer [10,13,14] and intra-observer variability [11].

A ‘gold-standard’ for incorporating MRI directly into the radiotherapy planning pathway is by acquisition of a dedicated MRI for planning in the radiotherapy planning position, using the same immobilisation devices, with subsequent rigid registration to the planning CT [13,15,16]. However, the majority of radiotherapy departments do not have access to a dedicated MRI scanner for radiotherapy planning. An alternative is registration of a diagnostic MRI scan, acquired optimally for radiotherapy purposes, to the planning CT scan. In a planning study, we have previously shown that deformable registration of a diagnostic MRI offers a significant advantage in the accuracy of GTV delineation compared with planning CT alone [11].

In the absence of a dedicated MRI acquired for radiotherapy planning, we have implemented the pathway of using deformable registration of a diagnostic MRI, acquired optimally for radiotherapy purposes, to planning CT into routine clinical practice. The purpose of this report is to assess the impact of the integration of a registered MRI upon GTV delineation of oropharyngeal carcinoma in routine clinical practice.

2. Methods

2.1. Study Population

This single centre retrospective study was approved by the institutional review board. Inclusion criteria for the study were: (1) histologically proven primary oropharyngeal squamous cell carcinomas (SCC); (2) treated with definitive radiotherapy +/− chemotherapy; (3) primary tumour identifiable on diagnostic MRI; (4) separate GTVs delineated in routine clinical practice using CT (GTV_{CT}), MRI (GTV_{MR}), and CT and MRI (GTV_{CTMR}); (6) all patients outlined by the same radiation oncologist; and (7) planning CT between February 2015 and December 2016.

2.2. Image Acquisition

2.2.1. MRI

Diagnostic MRI images were obtained on a 1.5T Siemens Magnetom Avanto (Siemens Healthcare, Erlangen, Germany) using axial T1-weighted sequences with a TR = 831 ms, TE = 8.6 ms, 105 2-mm thick contiguous slices and a voxel size = 0.9 × 0.9 × 2.0 mm. The sequence was a TIW fast spin echo sequence, with a number of scan averages of 2; the echo train length was 2 and parallel imaging using the GRAPPA method was used with a parallel imaging factor of 2. The patient was positioned on a standard curved diagnostic couch without an immobilisation mask. The use of a 2 mm contiguous...
slice thickness on MRI was a change to standard diagnostic practice in order to optimise the image registration to planning CT.

2.2.2. CT

Planning CT (pCT) images were optimised for use in treatment planning. CT imaging was performed on a Siemens Somatom Sensation 64-section (Siemens Healthcare, Erlangen, Germany). Images were acquired with the patient on a flat top radiotherapy couch in a five-point thermoplastic radiotherapy immobilization mask, and were contrast enhanced after a 25-s delay following a bolus of 100 mL of iodinated contrast (Omnipaque 300, Bracco Ltd., High Wycombe, UK) injected at 3 mL/s using the following settings: 120 kV, variable mA (min 10, max 600, noise index 12.2), tube rotation 1.0 s per rotation, pitch 0.9, with a 2.0 mm section reconstruction. The voxel size used was $1.0 \times 1.0 \times 2.0$ mm.

2.3. Deformable MRI–CT Image Registration

The T1 weighted MRI and pCT images were registered using deformable image registration (DIR). All registrations were performed with Mirada RTx v1.4 (Mirada Medical, Oxford, UK) by a trained radiotherapy dosimetrist. An automatic rigid fast algorithm was used for initial rigid registrations, and as the starting point for the DIR. If the automatic rigid registration results were not deemed acceptable, then a manual rigid registration was performed. DIR was then performed based upon a multi-modal, mutual information-based algorithm. All registrations were checked by the treating radiation oncologist by visual interrogation of the fused images; this involved assessing overlay images of regions of anatomy e.g., spinal cord, bone, and mucosa-air interfaces in proximity to the GTV.

2.4. GTV Delineation

Three contours per patient were produced by a single radiation oncologist specialising in head and neck cancer on Monaco v 5.1 (Elekta AB, Stockholm, Sweden). All contours were delineated viewing all available diagnostic imaging (including MRI) side-by-side. For each patient, three GTV contours had been manually delineated as part of routine clinical practice, according to an institutional protocol in the following order during the same session: GTV on the CT only without registration (GTV$_{CT}$), GTV on the T1-MRI only (GTV$_{MR}$) without viewing registered planning CT, and a final GTV based on viewing the GTV$_{CT}$ and GTV$_{MR}$ and registered images (GTV$_{CTMR}$).

2.5. Analysis of GTVs

The GTVs were quantitatively analysed by overall volume, and positional metrics using five contour comparison metrics calculated using ImSimQA v3.1 (OSL, Shrewsbury, UK). Three sets of comparisons were made: (i) GTV$_{CTMR}$ (reference) vs GTV$_{CT}$, (ii) GTV$_{CTMR}$ (reference) vs. GTV$_{MR}$, and (iii) GTV$_{CT}$ (reference) vs. GTV$_{MR}$. The five contour comparison metrics used are described below.

2.5.1. Conformity Index (CI)

CI is the ratio of the area of the overlapping region to the total area of the contours [17]. A value between 0 and 1 is given. 0 shows no overlap between the two GTVs. 1 shows identical GTVs.

$$CI = \frac{B}{A + B + C}$$  \hspace{1cm} (1)

$B$ represents the volume of the overlapping regions. $A$ and $C$ represent the volume of each GTV which is not overlapping.
2.5.2. Mean Distance to Conformity (MDC)

MDC is mean distance of each outlying voxel of the evaluation contour surface to the closest voxel from the reference contour surface [17]. MDC quantitates global mismatch in the shape between the two contours. MDC for a perfect overlap would be 0 mm.

2.5.3. Dice Similarity Coefficient (DSC)

DSC also quantifies the overlap between the two contours, with 1 showing complete overlap and 0 showing no overlap [17,18].

\[ \text{DSC} = \frac{2 \cdot (A \cap B)}{A + B} \]

where \( A \) is the volume of the reference contour and \( B \) is the volume of the evaluation contour.

2.5.4. Centre of Gravity Distance (CGD)

CGD is the distance between the geometrical centres of the two GTVs being compared. CGD for a perfect match would be 0 mm [17].

2.5.5. Volumetric Analysis

The volumes were compared by looking at a percentage difference between the volumes of the contours. Perfectly correlating contours would have a percentage difference of 0.

2.6. Statistical Analysis

Wilcoxon signed rank tests were performed on StataSE 13 (StataCorp, College Station, TX, USA) to test for significant difference in volumes between (i) GTV\(_{CT}\) and GTV\(_{MR}\), (ii) GTV\(_{CTMR}\) and GTV\(_{CT}\), and (iii) GTV\(_{CTMR}\) and GTV\(_{MR}\). Results were considered statistically significant if the \( p \)-value was \(<0.02\). A \( p \)-value of 0.02 was used to minimise the risk of Type I errors arising from multiple testing.

3. Results

22 consecutive patients who underwent radiotherapy for oropharyngeal carcinoma, and for whom primary tumour GTVs were contoured by a single radiation oncologist, were identified. A total of 9/22 (41%) and 10/22 (45%) patients had a tonsil and base of tongue primary respectively. Contouring was performed with the aid of deformable registration of a diagnostic MRI to the planning CT in routine clinical practice between February 2015 and December 2016. Patient and tumour characteristics are summarised in Table 1. All tumours were non-metastatic. The median time from diagnostic MRI to planning CT was 22 days (range 0–50). All patients had a diagnostic PET-CT, which was available for side-by-side assessment (not registered to planning CT) at the time of delineation. Figure 1 provides three examples of patients with small, medium, and large primary tumours, for whom deformable image registration of the MRI to the planning CT was undertaken, and the GTV\(_{CT}\), GTV\(_{MR}\), and GTV\(_{CTMR}\) which were contoured as part of routine clinical practice; GTV\(_{CTMR}\) was used as the basis for delineation of subsequent clinical target volumes (CTV) and planning target volumes (PTV) for treatment planning.
Table 1. Table to show patient demographics and tumour characteristics.

| Patient Number | Sex | Age (Years) | Subsite       | T-Stage | N-Stage | Grade | p16 Status | GTV<sub>CT</sub> (cm<sup>3</sup>) | GTV<sub>MR</sub> (cm<sup>3</sup>) | GTV<sub>CTMR</sub> (cm<sup>3</sup>) |
|----------------|-----|-------------|---------------|---------|---------|-------|------------|------------------|------------------|------------------|
| 1              | M   | 52          | Tonsil        | 3       | 2b      | 3     | positive   | 39.1             | 32.5             | 44.6             |
| 2              | M   | 56          | Tonsil        | 2       | 2b      | 3     | Positive   | 4.7              | 1.6              | 6.3              |
| 3              | M   | 45          | Base of tongue| 4a      | 2b      | 3     | Positive   | 6.1              | 8.7              | 10.1             |
| 4              | M   | 61          | Base of tongue| 2       | 2b      | 3     | Positive   | 5.3              | 5.8              | 7.0              |
| 5              | M   | 53          | Tonsil        | 4b      | 3       | 3     | Positive   | 20.2             | 21.9             | 27.2             |
| 6              | M   | 59          | Base of tongue| 4       | 2b      | 3     | Positive   | 12.5             | 15.3             | 17.8             |
| 7              | M   | 57          | Tonsil        | 3       | 3       | 3     | Positive   | 17.7             | 15.7             | 23.3             |
| 8              | F   | 53          | Base of tongue| 2       | 2b      | 3     | Positive   | 6.4              | 7.1              | 8.5              |
| 9              | F   | 66          | Pharyngeal wall| 3       | 1       | 3     | Positive   | 21.2             | 18.5             | 23.0             |
| 10             | F   | 50          | Base of tongue| 2       | 1       | 3     | Negative   | 7.4              | 7.2              | 8.1              |
| 11             | M   | 55          | Tonsil        | 4a      | 2b      | 2     | Positive   | 22.3             | 29.1             | 30.8             |
| 12             | F   | 61          | Base of tongue| 4a      | 2b      | 3     | Unknown    | 24.6             | 28.1             | 31.0             |
| 13             | M   | 53          | Tonsil        | 2       | 2b      | 3     | Positive   | 14.8             | 9.6              | 17.7             |
| 14             | M   | 65          | Base of tongue| 2       | 2c      | 3     | Positive   | 4.7              | 11.0             | 11.6             |
| 15             | M   | 58          | Tonsil        | 4a      | 2b      | 3     | Negative   | 24.0             | 26.5             | 30.8             |
| 16             | M   | 49          | Tonsil        | 3       | 0       | 3     | Positive   | 24.9             | 22.7             | 30.8             |
| 17             | M   | 56          | Base of tongue| 4a      | 2b      | 3     | Positive   | 17.0             | 18.8             | 22.7             |
| 18             | F   | 62          | Pharyngeal wall| 1       | 2b      | 3     | Unknown    | 3.5              | 6.0              | 6.3              |
| 19             | F   | 62          | Base of tongue| 2       | 2       | 3     | Negative   | 7.1              | 9.0              | 11.8             |
| 20             | M   | 53          | Base of tongue| 1       | 2b      | 2     | Unknown    | 3.8              | 3.9              | 5.5              |
| 21             | M   | 66          | Tonsil        | 2       | 2b      | 3     | Unknown    | 41.7             | 38.3             | 41.9             |
| 22             | F   | 66          | Soft palate   | 4a      | 1       | 3     | Unknown    | 8.3              | 11.5             | 14.7             |
Figure 1. Examples of delineation of primary tumour GTV using CT and deformably coregistered MRI in patients with different size primary tumours. (i) Example of large primary right tonsil carcinoma T3N2bM0 (Patient 1 in Table 1). (ii) Example of small primary right base of tongue carcinoma T1N2bM0 (Patient 20 in Table 1). (iii) Example of medium sized primary base of tongue carcinoma T4aN2bM0 (Patient 6 in Table 1). GTV\textsubscript{CT} (red), GTV\textsubscript{MR} (blue) and GTV\textsubscript{CTMR} (yellow) are shown. For each patient in (i), (ii), and (iii) axial, coronal, and sagittal images of planning CT are shown in (A–C) respectively; (D–F) show axial, coronal, and sagittal images of deformably coregistered T1-weighted diagnostic MRI respectively.
3.1. Comparison of GTV Volumes

The volumes of GTV\textsubscript{CT}, GTV\textsubscript{MR} and GTV\textsubscript{CTMR} for the 22 patients detailed in Table 2 are summarised in Figure 2. GTV\textsubscript{CTMR} was the largest volume for all 22 patients. Median GTV volumes were 13.7 cm\textsuperscript{3} (range 3.5–41.7), 13.4 cm\textsuperscript{3} (range 1.6–38.3), and 17.8 cm\textsuperscript{3} (range 5.5–44.5) for GTV\textsubscript{CT}, GTV\textsubscript{MR}, and GTV\textsubscript{CTMR}, respectively. As shown in Table 2, there was no significant difference in GTV\textsubscript{CT} and GTV\textsubscript{MR} volumes. GTV\textsubscript{CTMR} was found to be significantly larger than both GTV\textsubscript{MR} and GTV\textsubscript{CT}. GTVs for tonsil primary lesions were significantly higher than for non-tonsil primary lesions. Mean and median GTV\textsubscript{CTMR} for tonsil primary sites were 28.2 cm\textsuperscript{3} and 30.8 cm\textsuperscript{3}, versus 13.7 cm\textsuperscript{3} and 11.6 cm\textsuperscript{3}, respectively, for non-tonsil primary sites ($p = 0.01$).

**Table 2.** Summary of volume of GTVs contoured using CT, MR, and deformably registered MRI to CT. Statistically significant differences in bold ($p < 0.02$).

| Modality | Modality GTV Volumes (cm\textsuperscript{3}) |
|----------|---------------------------------------------|
|          | Mean | Median | Mean St. Dev. | Range | GTV Comparison (Wilcoxon Signed Rank Test) |
| CT       | 15.3 | 13.7   | 11.1          | 41.7  | Min | Max | GTV\textsubscript{CT}–GTV\textsubscript{MR}, $p = 0.47$ |
| MR       | 15.9 | 13.4   | 10.2          | 38.3  | 1.6 | 1.6 | GTV\textsubscript{CTMR}–GTV\textsubscript{CT}, $p = 0.018$ |
| CT–MR    | 19.6 | 17.8   | 11.8          | 44.6  | 5.5 | 5.5 | GTV\textsubscript{CTMR}–GTV\textsubscript{MR}, $p = 0.010$ |

![Figure 2](image)

*Figure 2.* Box plot summarising volumes of GTV\textsubscript{CT} (red), GTV\textsubscript{MR} (blue) and GTV\textsubscript{CTMR} (yellow). The x denotes the mean, and the horizontal line in each box the median value.

3.2. Comparison of Positional Metrics

The analysis of positional contour variability is summarised in Table 3. Based on all four positional metrics, GTV\textsubscript{CT} and GTV\textsubscript{MR} were the least similar (for example, mean DSC 0.71, 0.84, 0.82 for GTV\textsubscript{CT}–GTV\textsubscript{MR}, GTV\textsubscript{CTMR}–GTV\textsubscript{CT} and GTV\textsubscript{CTMR}–GTV\textsubscript{MR}, respectively). Positional metrics (DSC, conformality index, MDC and CGD) comparing GTV\textsubscript{CT}–GTV\textsubscript{MR} were significantly different when compared with those derived from either GTV\textsubscript{CTMR}–GTV\textsubscript{CT} or GTV\textsubscript{CTMR}–GTV\textsubscript{MR} analysis. By contrast, there were no significant differences in positional metrics obtained from comparing
CT_{CTMR}−GTV_{MR} and those comparing CT_{CTMR}−GTV_{CT} (Table 3). There were no significant differences in DSC, conformity index, MDC, or CGD comparing GTV_{CT}−GTV_{MR}, GTV_{CTMR}−GTV_{CT} and GTV_{CTMR}−GTV_{MR} for tonsil versus non-tonsil primaries.

Table 3. Intermodality positional GTV analysis.

| Metric                      | DSC (Mean (SD)) | Conformity Index (Mean (SD)) | MDC (mm) (Mean (SD)) | CGD (mm) (Mean (SD)) |
|-----------------------------|-----------------|------------------------------|----------------------|----------------------|
| GTV_{CTMR}−GTV_{CT}         | 0.84 (0.10)     | 0.66 (0.12)                 | 3.74 (1.19)          | 2.02 (1.40)          |
| GTV_{CTMR}−GTV_{MR}         | 0.82 (0.10)     | 0.65 (0.14)                 | 4.03 (1.10)          | 1.86 (1.59)          |
| GTV_{CT}−GTV_{MR}           | 0.71 (0.15)     | 0.50 (0.15)                 | 5.09 (1.89)          | 3.08 (2.37)          |

Statistical Comparisons of Positional Metrics: p Values

| GTV_{CT}−GTV_{MR} versus GTV_{CTMR}−GTV_{CT} | <0.01 | <0.01 | <0.01 | <0.01 |
| GTV_{CT}−GTV_{MR} versus GTV_{CTMR}−GTV_{CT} | <0.01 | <0.01 | <0.01 | <0.01 |
| GTV_{CTMR}−GTV_{MR} versus GTV_{CTMR}−GTV_{CT} | 0.71  | 0.82  | 0.15  | 0.70  |

4. Discussion

Accurate GTV delineation is a critical issue in head and neck radiotherapy, with the use of IMRT with steep dose gradients [12] and the introduction of new guidelines with tight volumetric and anatomical margins to CTVs [7]. Target volume delineation remains a major source of variability in the planning process [6,19,20]. Recurrence analyses have shown that geographical misses remain a key issue in the era of IMRT; in an analysis of locoregional recurrences referred to a tertiary centre, 41% were classified as marginal, [3] implying inadequate GTV delineation. The integration of multimodality imaging into the radiotherapy planning process holds the potential to considerably improve the accuracy and reproducibility of GTV delineation. Delineation of oropharyngeal carcinomas is particularly challenging on CT, with very limited soft tissue resolution and problems with dental artefacts [9,12], with MRI offering superiority in diagnostic pathways [7]. MRI-based delineation is widely considered to be more reproducible [10,11,21]. Although MRI is associated with geometric spatial distortion, this is very small for a limited cross-sectional area such as the head and neck region [13]. However, MRI does not provide the electron density information required for dosimetric calculation, and current clinical use requires registration to the planning CT.

Registration of an MRI acquired in treatment position to planning CT is considered the preferred method for the use of MRI in treatment planning. In common with many centres, we do not have dedicated access to an MRI acquired for planning purposes. In a prospective planning study in 14 patients [11] (of whom 11 had an oropharyngeal carcinoma), we defined a GTV based on a dedicated MRI acquired in the treatment position as a reference standard; deformable registration of a diagnostic MRI to planning CT led to GTV delineation which was significantly closer than based on planning CT alone (with side-by-side assessment of diagnostic imaging). Based upon these findings we have introduced deformable registration of a diagnostic T1-weighted MRI acquired at 2 mm contiguous slice thickness to the planning CT into clinical practice. This analysis reports on the impact of this pathway upon delineation of a final GTV used for planning in the reality of clinical practice.

In the retrospective analysis, all GTVs were contoured as part of routine clinical practice. There was no overall significant difference in the volume of primary tumours delineated using CT or MR alone. Consistent with this data, a previous planning study of 11 patients with oropharyngeal carcinoma found no significant difference in overall volumes outlined by multiple observers comparing GTVs delineated on CT or MRI alone, although GTV delineation on MRI was associated with a significantly smaller volume standard deviation than on CT [10]. Data are variable in other planning studies across differing head and neck tumour sites comparing CT- and MRI-defined GTVs, with some showing similar overall GTVs [22–24], although some studies have suggested MRI-defined GTVs are smaller [25,26] or larger [9,27,28]. In this current study, GTVs delineated using both MRI and CT

Healthcare 2018, 6, 135
GTVCTMR were significantly larger than those delineated using CT or MRI alone. Our prior planning study also showed that GTVs based on CT were significantly smaller than those using combined CT–MRI, although there were non-significant differences between GTVs based on MRI alone and combined CT–MRI [10].

Positional metric analysis demonstrated only limited similarity between GTVCT and GTVMR (DSC 0.71), despite the availability of diagnostic MRI (and PET-CT in 100% of patients) to guide delineation. There was more similarity between GTVCTMR and both GTVCT and GTVMR (DSC 0.84 and 0.82 respectively). Positional metrics comparing GTVCT–GTVMR were significantly different when compared with those derived from either GTVCTMR–GTVCT or GTVCTMR–GTVMR analysis. There was no significant difference in positional metric analysis between tonsil and non-tonsil primary tumours.

The findings of the volume and positional analysis suggests that the radiation oncologist used both GTVCT and GTVMR in determining the final GTVCTMR, which was used as the basis for subsequent CTV and PTV delineation according to institutional protocols. Therefore, both CT and MRI appear to be potentially complimentary in GTV delineation. Overall, clinical implementation of deformable registration of diagnostic MRI into clinical practice led to delineation of significantly larger GTVs than those delineated on CT alone. Dental artefacts and lack of soft tissue contrast on CT provide a clear rationale to integrate MRI directly into the planning process. Based upon this premise, these data suggest that this pathway may reduce the risk of geometric misses by more accurately identifying the GTV. However, conversely, multimodality imaging which leads to larger GTVs will increase high dose planning target volumes (PTV) (assuming the same protocols are used in generating CTVs and PTVs); this has the potential to increase treatment-related toxicity.

Limitations of this study include the use of a T1-weighted MRI without gadolinium; although gadolinium may assist in delineation, we do not have a suitable diagnostic MRI sequence currently available with a slice thickness that facilitates accurate registration (due to increased scan time required and constraints on MRI availability). However, enhanced sequences are available for side-by-side comparison with the registered non-enhanced T1-sequence. The analysis is based upon GTV delineation by a single experienced observer who had received instruction via head and neck radiologists in the use of MRI. Inter-observer variation in GTV delineation is well recognised [9,10,20], and is likely to have been seen if further observers had been included. However, the overall results are consistent with those of our prior multi-observer planning study in oropharyngeal carcinoma [10]. The diagnostic MRI was acquired prior to the planning CT (median interval 22 days); there is potential for tumour growth in the intervening period. The underlying premise for the integration of MRI into the planning process is that this will enhance the accuracy of GTV delineation; however, there are no correlative pathological data available as a ‘ground truth’ for oropharyngeal carcinoma to evaluate the accuracy of GTVs constructed using CT alone, MRI alone or a combination of both.

5. Conclusions

This study of clinical practice shows that the use of deformable registration of diagnostic MRI to planning CT for oropharyngeal carcinomas produced significantly larger primary tumour GTVs compared to the use of CT alone, with significant positional differences. These data suggest a complementary role of MRI in the radiotherapy planning process for oropharyngeal carcinoma to reduce the risk of geographical misses. This is potentially of increased importance in light of recent guidelines of tighter GTV to CTV margins. These data also show that combined modality imaging is likely to lead to larger target volumes and consequently potential for increased toxicity.

Author Contributions: A.T.: data collection, analysis and interpretation, manuscript preparation; M.S.: concept, contouring, manuscript review; R.J.D.P.: data interpretation, manuscript preparation.

Funding: This research received no external funding.

Acknowledgments: We would like to acknowledge the assistance of the Radiotherapy Physics Department in Leeds Cancer Centre in software use for analysis.
Conflicts of Interest: The authors declare no conflict of interest.

References

1. Eisbruch, A.; Marsh, L.H.; Dawson, L.A.; Bradford, C.R.; Teknos, T.N.; Chepeha, D.B.; Worden, F.P.; Urba, S.; Lin, A.; Schipper, M.J.; et al. Recurrences near base of skull after IMRT for head-and-neck cancer: Implications for target delineation in high neck and for parotid gland sparing. *Int. J. Radiat. Oncol. Biol. Phys.* 2004, 59, 28–42. [CrossRef] [PubMed]

2. Chen, A.M.; Farwell, D.G.; Luu, Q.; Chen, L.M.; Vijayakumar, S.; Purdy, J.A. Marginal misses after postoperative intensity-modulated radiotherapy for head and neck cancer. *Int. J. Radiat. Oncol. Biol. Phys.* 2011, 80, 1423–1429. [CrossRef] [PubMed]

3. Chen, A.M.; Chin, R.; Beron, P.; Yoshizaki, T.; Mikaelian, A.G.; Cao, M. Inadequate target volume delineation and local-regional recurrence after intensity-modulated radiotherapy for human papillomavirus-positive oropharynx cancer. *Radiother. Oncol.* 2017, 123, 412–418. [CrossRef] [PubMed]

4. Njeh, C.F. Tumor delineation: The weakest link in the search for accuracy in radiotherapy. *J. Med. Phys.* 2008, 33, 136–140. [CrossRef] [PubMed]

5. Weiss, E.; Hess, C.F. The impact of gross tumor volume (GTV) and clinical target volume (CTV) definition on the total accuracy in radiotherapy theoretical aspects and practical experiences. *Strahlenther. Onkol.* 2003, 179, 21–30. [CrossRef] [PubMed]

6. Rasch, C.; Steenbakkers, R.; van Herk, M. Target definition in prostate, head, and neck. *Semin. Radiat. Oncol.* 2005, 15, 136–145. [CrossRef] [PubMed]

7. Grégoire, V.; Evans, M.; Le, Q.T.; Bourhis, J.; Budach, V.; Chen, A.; Eisbruch, A.; Feng, M.; Giralt, J.; Gupta, T.; et al. Delineation of the primary tumour Clinical Target Volumes (CTV-p) in laryngeal, hypopharyngeal, oropharyngeal and oral cavity squamous cell carcinoma: AIRO, CACA, DAHANCA, EORTC, GEORCC, GORTEC, HKNPCSG, HNCIG, IAG-KHT, LPRHHT, NCIC CTG, NCRI, NRG Oncology, PHNS, SBRT, SOMERA, SRO, SSHNO, TROG consensus guidelines. *Radiother. Oncol.* 2018, 126, 3–24. [PubMed]

8. Battista, J.J.; Rider, W.D.; Van Dyk, J. Computed tomography for radiotherapy planning. *Int. J. Radiat. Oncol. Biol. Phys.* 1980, 6, 99–107. [CrossRef]

9. Ahmed, M.; Schmidt, M.; Sohaib, A.; Kong, C.; Burke, K.; Richardson, C.; Usher, M.; Brennan, S.; Riddell, A.; Davies, M.; et al. The value of magnetic resonance imaging in target volume delineation of base of tongue tumours—A study using flexible surface coils. *Radiother. Oncol.* 2010, 94, 161–167. [CrossRef] [PubMed]

10. Bird, D.; Scarsbrook, A.F.; Sykes, J.; Ramasamy, S.; Subesinghe, M.; Carey, B.; Wilson, D.J.; Roberts, N.; McDermott, G.; Karakaya, E.; et al. Multimodality imaging with CT, MR and FDG-PET for radiotherapy target volume delineation in oropharyngeal squamous cell carcinoma. *BMC Cancer* 2015, 15, 844. [CrossRef] [PubMed]

11. Chuter, R.; Prestwich, R.; Bird, D.; Scarsbrook, A.; Sykes, J.; Wilson, D.; Speight, R. The use of deformable image registration to integrate diagnostic MRI into the radiotherapy planning pathway for head and neck cancer. *Radiother. Oncol.* 2017, 122, 229–235. [CrossRef] [PubMed]

12. Prestwich, R.J.; Sykes, J.; Carey, B.; Sen, M.; Dyker, K.E.; Scarsbrook, A.F. Improving target definition for head and neck radiotherapy: A place for magnetic resonance imaging and 18-fluoride fluorodeoxyglucose positron emission tomography? *Clin. Oncol.* 2012, 24, 577–589. [CrossRef] [PubMed]

13. Webster, C.J.; Kilgallon, J.E.; Ho, K.F.; Rowbottom, C.G.; Slevin, N.J.; Mackay, R.I. A novel imaging technique for fusion of high-quality immobilised MR images of the head and neck with CT scans for radiotherapy target delineation. *Br. J. Radiol.* 2009, 82, 497–503. [CrossRef] [PubMed]

14. Ahmed, M.; Hansen, V.N.; Harrington, K.J.; Nutting, C.M. Reducing the risk of xerostomia and mandibular osteoradionecrosis: The potential benefits of intensity modulated radiotherapy in advanced oral cavity carcinoma. *Med. Dosim.* 2009, 34, 217–224. [CrossRef] [PubMed]

15. Brunt, J.N. Computed tomography-magnetic resonance image registration in radiotherapy treatment planning. *Clin. Oncol.* 2010, 22, 688–697. [CrossRef] [PubMed]

16. Hanvey, S.; McJury, M.; Tho, L.M.; Glegg, M.; Thomson, M.; Grose, D.; James, A.; Rizwanullah, M.; Paterson, C.; Foster, J. The influence of MRI scan position on patients with oropharyngeal cancer undergoing radical radiotherapy. *Radiother. Oncol.* 2013, 8, 129. [CrossRef] [PubMed]
17. Jena, R.; Kirkby, N.F.; Burton, K.E.; Hoole, A.C.; Tan, L.T.; Burnet, N.G. A novel algorithm for the morphometric assessment of radiotherapy treatment planning volumes. Br. J. Radiol. 2010, 83, 44–51. [CrossRef] [PubMed]

18. Dice, L.R. Measures of the amount of ecologic association between species. Ecology 1945, 26, 297–302. [CrossRef]

19. Prestwich, R.J.; Subasinghe, M.; Gilbert, A.; Chowdhury, F.U.; Sen, M.; Scarsbrook, A.F. Delayed response assessment with FDG-PET-CT following (chemo) radiotherapy for locally advanced head and neck squamous cell carcinoma. Clin. Radiol. 2012, 67, 966–975. [CrossRef] [PubMed]

20. Cooper, J.S.; Mukherji, S.K.; Toledano, A.Y.; Beldon, C.; Schmalfuss, I.M.; Amédur, R.; Sailer, S.; Loevner, I.A.; Kousouboris, P.; Ang, K.K.; et al. An evaluation of the variability of tumor-shape definition derived by experienced observers from CT images of supraglottic carcinomas (ACRIN protocol 6658). Int. J. Radiat. Oncol. Biol. Phys. 2007, 67, 972–975. [CrossRef] [PubMed]

21. Dirix, P.; Haustermans, K.; Vandecaveye, V. The value of magnetic resonance imaging for radiotherapy planning. Semin. Radiat. Oncol. 2014, 24, 151–159. [CrossRef] [PubMed]

22. Daisne, J.F.; Duprez, T.; Weynand, B.; Lonneux, M.; Hamoir, M.; Reychler, H.; Grégoire, V. Tumor volume in pharyngolaryngeal squamous cell carcinoma: Comparison at CT, MR imaging, and FDG PET and validation with surgical specimen. Radiology 2004, 233, 93–100. [CrossRef] [PubMed]

23. Geets, X.; Tomsej, M.; Lee, J.A.; Duprez, T.; Coche, E.; Cosnard, G.; Lonneux, M.; Grégoire, V. Adaptive biological image-guided IMRT with anatomic and functional imaging in pharyngo-laryngeal tumors: Impact on target volume delineation and dose distribution using helical tomotherapy. Radiother. Oncol. 2007, 85, 105–115. [CrossRef] [PubMed]

24. Seitz, O.; Chambron-Pinho, N.; Middendorp, M.; Sader, R.; Mack, M.; Vogl, T.J.; Bisdas, S. 18F-Fluorodeoxyglucose-PET/CT to evaluate tumor, nodal disease, and gross tumor volume of oropharyngeal and oral cavity cancer: Comparison with MR imaging and validation with surgical specimen. Neuroradiology 2009, 51, 677–686. [CrossRef] [PubMed]

25. Rasch, C.; Keus, R.; Pameijer, F.A.; Koops, W.; Muller, S.; Touw, A.; Bartelink, H.; Lebesque, J.V. The potential impact of CT-MRI matching on tumor volume delineation in advanced head and neck cancer. Int. J. Radiat. Oncol. Biol. Phys. 1997, 39, 841–848. [CrossRef]

26. Chung, N.N.; Ting, L.L.; Lui, L.T.; Wang, P.M. Impact of magnetic resonance imaging versus CT on nasopharyngeal carcinoma: Primary tumor target delineation for radiotherapy. Head Neck 2004, 26, 241–246. [CrossRef] [PubMed]

27. Caldas-Magalhaes, J.; Kaspers, N.; Kooij, N.; van den Berg, C.A.; Terhaard, C.H.; Raaijmakers, C.P.; Philippens, M.E. Validation of imaging with pathology in laryngeal cancer: Accuracy of the registration methodology. Int. J. Radiat. Oncol. Biol. Phys. 2012, 82, e289–e298. [CrossRef] [PubMed]

28. Emami, B.; Sethi, A.; Petruzzielli, G.J. Influence of MRI on target volume delineation and IMRT planning in nasopharyngeal carcinoma. Int. J. Radiat. Oncol. Biol. Phys. 2003, 57, 481–488. [CrossRef]