Evaluating the impact of a rigid and a deformable registration method of pre-treatment images for hypoxia-based dose painting

M. Lazzeroni a,b,*, A. Ureba a,b, V. Rosenberg c, H. Schäfer d, A. Rühle d,e, D. Baltas d, I. Toma-Dasu a,b, A.L. Grosu d

a Department of Physics, Stockholm University, Sweden
b Department of Oncology and Pathology, Karolinska Institute, Stockholm, Sweden
c Royal Institute of Technology (KTH), Stockholm, Sweden
d Department of Radiation Oncology, Medical Center, Medical Faculty Freiburg, German Cancer Consortium (DKTK) Partner Site Freiburg, Germany
e University of Leipzig Medical Center, Department of Radiation Oncology, Leipzig, Germany

ARTICLE INFO

Keywords:
Hypoxia
Dose painting
Image registration

ABSTRACT

Purpose: To assess the impact of rigid and deformable image registration methods (RIR, DIR) on the outcome of a hypoxia-based dose painting strategy.

Materials and methods: Thirty head and neck cancer patients were imaged with $[^{18}F]$FMISO-PET/CT before radiotherapy. $[^{18}F]$FMISO-PET/CT images were registered to the planning-CT by RIR or DIR. The $[^{18}F]$FMISO uptake was converted into oxygen partial pressure ($pO_2$) maps. Hypoxic Target Volumes were contoured on $pO_2$ maps for the deformed (HTV def) and non-deformed (HTV) cases. A dose escalation strategy by contours, aiming at 95 % tumour control probability (TCP), was applied. HTVs were characterised based on geometry-related metrics, the underlying $pO_2$ distribution, and the dose boost level. A dosimetric and radiobiological evaluation of selected treatment plans made considering RIR and DIR was performed. Moreover, the TCP of the RIR dose distribution was evaluated when considering the deformed $[^{18}F]$FMISO-PET image as an indicator of the actual target radiosensitivity to determine the potential impact of an unalignment.

Results: Statistically significant differences were found between HTV and HTV def for volume-based metrics and underlying $pO_2$ distribution. Eight out of nine treatment plans for HTV and HTV def showed differences on the level 10 %/3 mm on a gamma analysis. The TCP difference, however, between RIR and the case when the RIR dose distribution was used with the deformed radiosensitivity map was below 2 pp.

Conclusions: Although the choice of the CT plan-to-PET registration method pre-treatment impacts the HTV localisation and morphology and the corresponding dose distribution, it negligibly affects the TCP in the proposed dose escalation strategy by contours.

1. Introduction

Modern radiotherapy has seen a rapid increase in the qualitative and quantitative use of multiple image data sets at all stages of the treatment process, from planning, delivering, monitoring and adapting the treatment [1]. At the planning stage, for example, the concomitant use of anatomical and functional information for accurate target definition is a well-established clinical practice [2]. The role of functional imaging in radiotherapy, such as positron emission tomography (PET), is rapidly expanding [3,4], PET may also be crucial in identifying the tumour sub-regions of increased radioresistance for a radiobiologically-guided dose escalation strategy [5-9]. Indeed, evidence is accumulating that the intrinsic radiosensitivity of individual patients and the spatial and temporal heterogeneities of the tumour radioresistance before and during treatment may be important causes of tumour recurrence and treatment failure for many cancer types [10]. As tumour hypoxia is considered one of the main determinants of tumour radioresistance for many solid tumours [11], PET imaging of hypoxia has been used to...
envisaged different dose sculpting strategies targeting those radio-resistant tumour sub-volumes evidenced from the images. These proposed dose escalation strategies include empirical approaches exclusively based on clinical experience [9], methods based on a linear conversion of radiotracer uptake into dose levels [9], and more sophisticated strategies guided by radiobiological modelling where the computation of the dose distribution required to achieve a desired level of tumour control probability (TCP) is considered [12].

Yet, using multiple images acquired at different time points is only possible when proper image registration is achievable, either rigid image registration (RIR, merely based on translations and rotations) or deformable image registration (DIR, where local distortions between the image sets are accounted for). Data from a recent international survey [13] showed that most of the registrations performed in the radiotherapy process are of RIR type. In most cases, DIR is performed with dedicated software and, in a notably smaller percentage of cases, by using DIR included in the treatment planning system (TPS). Respondent centres reported clinically implementing DIR in atlas-based segmentation, multi-modality treatment planning, and dose deformation. As also evidenced by the work of the AAPM Radiation Therapy Committee Task Group No. 132 [14], DIR seems to have a fundamental role in the radiotherapy process when anatomical changes, due, for example, to patient weight loss, tumour shrinkage, and physiological organ shape variation, are expected. Despite its potential, several factors hinder the broader adoption of DIR in clinical practice. These include the need for more guidelines, training, improved tools for commissioning DIR software, and the quality assurance of registration results—which should involve creating or advising which quantitative metrics to use [15-17]. Consequently, while DIR stands as a powerful and versatile tool in radiotherapy, its utilisation may be accompanied by challenging-to-quantify uncertainties [18], which naturally depend on the particular DIR algorithm in use [19].

However, the impact of using DIR between the planning Computed Tomography (CT) image and the hypoxia PET image acquired at the pre-treatment stage for defining the hypoxic compartment and, subsequently, the required dose escalation level has not received the same attention in the literature. Therefore, this study aimed to assess the impact of the method used for image registration on a dose painting approach based on $^{18}$F-fluoromisonidazole ($^{18}$F-FMISO)-PET. A tri-fold assessment taking into consideration the morphological changes of the hypoxic compartment, the dosimetric impact, as well as the radiobiological impact on the TCP is reported in this study.

## 2. Materials and methods

### 2.1. Image dataset

The study included 30 patients with locally advanced Head and Neck (H&N) Squamous Cell Carcinoma who received concomitant radiochemotherapy at the University Medical Center Freiburg in Germany. The Gross Tumour Volume, including both the primary tumour and the lymph nodes (GTVNX) was delineated by board-certified radiation oncologists using Computed Tomography (CT), Magnetic Resonance Imaging (MRI), and $^{18}$F-fluorodeoxyglucose ($^{18}$F-FDG)-PET images. The study is registered at the German Clinical Trial Register (DRKS00003830) and was approved by the independent Ethics Committee of the University of Freiburg (reference no. 479/12).

A planning CT (CTplan) and an $^{18}$F-FMISO-PET/CT were acquired at different time points before the start of the treatment (with a time span of 6 ± 4 days between the CTplan and $^{18}$F-FMISO-PET/CT). The $^{18}$F-FMISO-PET image acquisition started 160 min after 300 MBq of radiotracer injection with one-bed position covering the whole H&N region. Patients were imaged in radiotherapy position with an H&N mask – the mask was also used for CT treatment planning and during the radiation treatment. PET data were reconstructed using an ordered-subset expectation-maximization algorithm to voxels of (2x2x2) mm$^3$.

### 2.2. Image registration strategies

The $^{18}$F-FMISO-PET/CT scan was co-registered to CTplan in a research version of the treatment planning system RayStation (RS TPS, RaySearch Laboratories AB, Sweden). The study evaluated two methods for image registration: a rigid CTplan-to-CT registration method and a deformable CTplan-to-CT registration method. The RIR consisted of translation and/or rotation operations based on the patient’s skeletal anatomy to align the frame of references of the acquired images. On the other hand, the DIR consisted of a rigid registration based on the bony anatomy followed by a hybrid deformable registration, which combines image information with anatomical information as provided by contour image sets (ANAtomically CONstrained Deformation Algorithm, ANACONDA) [20]. The DIR algorithm is based on a hybrid objective function consisting of a weighted sum of three non-linear terms: image similarity term, grid regularisation term, and penalty term [1,20]. The DIR vector field of the CTplan-to-CT registration (grid size 0.25 × 0.25 × 0.25 cm/voxel) was then applied to the $^{18}$F-FMISO-PET image using an in-house developed tool using the scripting features available in RS. A deformed PET image was then generated and loaded as a DICOM in the TPS.

Fig. 1 shows the study’s workflow, illustrating all the steps from image registration to evaluation of the treatment plans.

### 2.3. $^{18}$F-FMISO uptake-to-pO$_2$ conversion and HTV delineation

The $^{18}$F-FMISO-PET uptake was converted into partial pressure of oxygen (pO$_2$) distribution by applying a sigmoid conversion function at the voxel level [21-23]:

$$
pO_2 = \frac{c(a - \text{Uptake}(r))}{b + \text{Uptake}(r) - a}
$$

where a, b, and c are $^{18}$F-FMISO reaction-specific parameters equal to 10.9, 10.7, and 2.5 mmHg, respectively. The parameter $Uptake$ in equation (1) was calculated as follows: the voxel values in the $^{18}$F-FMISO-PET images were divided by the average value in a Well Oxygenated Volume (WOV), and the results were multiplied by the tracer uptake predicted by the conversion function for the assigned pO$_2$ in the WOV. The deep neck muscle volume, delineated by an expert radiologist, was chosen as WOV with an assigned pO$_2$ equal to 30 mmHg [24,25]. The pO$_2$-HTV was then delineated by automatically thresholding the pO$_2$ maps at 10 mmHg and intersecting with the clinical target volume (CTV). As two sets of pO$_2$ images were available, the corresponding to the rigidly registered $^{18}$F-FMISO-PET and the one corresponding to the deformally registered $^{18}$F-FMISO-PET, the hypoxic compartment delineated in the first set was named HTV, while the second one was named HTV$_{rel}$. To note that, if not otherwise specified, the notation HTV (or HTV$_{rel}$) refers to the actual volume of the hypoxic compartment.

### 2.4. Dose prescription and treatment planning

The dose distribution at the voxel level, D(i), which should be administered to a target to achieve a specified level of TCP, was determined using the methodology suggested by Toma-Dasu et al. [12,22]:

$$
D(i) = \frac{\alpha}{\beta} f(i) \left( 1 + 4 \frac{\beta}{\alpha^2} \ln \left( \frac{V_{fp}(i)}{V_{TCP}} \right) + 1 \right)
$$

where n is the number of treatment fractions, $\alpha/\beta$ is the intrinsic radiosensitivity expressed as the ratio of the parameters in the LQ model [26] and set to 10 Gy, $\nu$ is the volume of interest (VOI), and $\alpha$ is set equal to 0.32 Gy$^{-1}$ [12,22]. The parameter $f$ represents a map of dose...
Modification factors extracted from \( pO_2 \) maps as:

\[
f(i) = \frac{OER_{\text{max}}(k + pO_2(i))}{k + OER_{\text{max}}pO_2(i)}
\]

and having parameters equal to \( OER_{\text{max}} = 3 \) as the maximum Oxygen Enhancement Ratio (OER) used to measure the radiation sensitivity depending on the oxygenation, \( k = 2.5 \) [12,22].

Starting from a heterogeneous dose distribution determined by equation (2), a dose escalation strategy based on delivering a uniform dose level in the different targets (i.e. the rim between the CTV and the GTV, CTV-GTV; the rim between the GTV and the HTV, GTV-HTV; the hypoxic compartment, HTV) and aiming at 95% TCP in the CTV, was applied (dose prescription by contours) [12,22]:

\[
D_P = \frac{\gamma}{1 - \exp\left(\frac{-\bar{D}}{\sigma_D}\right)}
\]

(4)

where \( \gamma \) is the slope of the TCP curve set equal to 4 [12,22], \( \bar{D} \) is the average dose in the volume of interest and \( \sigma_D \) is the standard deviation for the dose within the considered VOI.

Photon treatment plans with VMAT two arcs were performed using the RS TPS. Dose distributions were delivered in 35 fractions using an integrated boost strategy in the HTV and HTV\(_{\text{def}}\), respectively, according to the calculations for the dose prescription given in Equation (4). The dose prescription to the rim between the PTV and the CTV was 52.5 Gy (i.e. PTV-CTV). Treatment plans were made using the same objective functions for plans corresponding to the RIR and DIR strategies.

2.5. Tri-fold evaluation of the impact of the registration method

The impact of using a rigid or deformable CT\(_{\text{plan}}\)-to-CT registration method was assessed on three evaluation levels:

a) HTV characterisation
b) Dosimetric evaluation
c) Radiobiological evaluation

The HTV characterisation was based on geometry-related metrics, on the location of the HTVs, on the underlying \( pO_2 \) distribution, and on the corresponding prescribed dose boost levels. In particular, the following quantities were calculated between HTV and HTV\(_{\text{def}}\): the geometrical centre shift (cm), the absolute difference between the volumes (cm\(^3\)), the average and maximum distance transform (cm) [27], the DICE index...
[28], the difference in the mean pO₂ value of HTV with respect to HTV_{def} (mmHg), and the difference in the prescribed uniform dose boost level in HTV with respect to HTV_{def} aiming at 95 % TCP in CTV (equation (4) (Gy)). Moreover, the non-parametric Wilcoxon rank test [29] was performed to uncover statistically significant differences between the quantities of interest calculated for HTV and HTV_{def}. A Spearman’s rank correlation coefficient [30] was calculated to find possible dependencies between the variables. The statistical analysis in this study was performed using the software MedCalc (version 20.027, MedCalc Software Ltd). A subset of nine patients was selected for treatment planning and further dosimetric and radiobiological evaluation based on the following variables: 1) the volume of the hypoxic compartment, 2) the geometrical centre shift, 3) the mean value of the underlying pO₂ distribution within the hypoxic compartment and 4) the uniform dose boost level prescribed to the hypoxic target. Variables calculated for HTV and HTV_{def}, which did not show statistically significant differences from the non-parametric Wilcoxon rank test, were disregarded. Of those remaining variables, only the ones that resulted as uncorrelated from the results of Spearman’s analysis were kept for patient selection.

The clinical feasibility of applying the proposed dose prescription method was assessed based on a dosimetric evaluation of the treatment plans accounting for target coverage, constraints for the organs at risk (OARs), conformity and homogeneity indexes (Gs, Hs) [31]. Furthermore, an in-house developed module for gamma index analysis [32] was used to quantify the spatial and dosimetric differences in the planned dose distributions between the treatment plans with HTV or HTV_{def} as boost volume. Plan differences were assessed on the 10 %/3 mm, 5 %/3 mm, 3 %/3 mm, and 2 %/3mm levels.

In the subsequent radiobiological evaluation, the TCP was evaluated for the CTV using a Poisson-LQ approach considering the RIR and DIR dose distributions and their corresponding radiosensitivity maps (TCP RIR and TCP DIR, respectively, in Fig. 1) [12,22]. An initial cell density of 10⁷ cells per cm³ was considered for the calculations. The TCP was also evaluated for the specific case where the RIR dose distribution was used with the radiosensitivity of the deformed PET image (assumed to be the actual patient radiosensitivity) to determine the effect of an improper alignment between dose distribution and radiosensitivity map (TCP* in Fig. 1). Furthermore, Normal Tissue Complication Probabilities (NTCPs) were calculated using the Lyman-Kutcher-Burman model [33-35] for the main OARs (i.e. spinal cord, brain stem, parotids, mandible) for the RIR and DIR treatment plans. The considered NTCP endpoints were xerostomia for the parotid glands, myelopathy for the spinal cord, necrosis for the brain stem, and osteoradionecrosis for the mandible [33,34,36]. Statistically significant differences were sought using the Wilcoxon rank test to assess the differences in TCP and NTCP between the RIR and DIR treatment plans, as well as to assess differences between the TCP RIR and TCP* (Fig. 1).

3. Results

Of the 30 patients analysed, 21 patients presented an HTV defined by thresholding the pO₂ maps at 10 mmHg and were considered for further analysis. Twelve patients had an HTV larger than 1 cm² (corresponding to a sphere of radius of about 6 mm).

Fig. 2 presents the results of four exemplifying patient cases (A-D) where a 3D representation of the hypoxic compartment contoured on the rigidly registered pO₂ distribution (HTV) and the hypoxic compartment contoured on the deformably registered pO₂ distribution (HTV_{def}) are overlapped on the planning CT (panel I). In panels II and III, the main radiotherapy targets of interest are shown in the transversal and sagittal planes, respectively. The main geometry-related metrics describing the differences and positional shifts between the volumes are reported in the table.

The non-parametric Wilcoxon rank test evidenced that HTV and HTV_{def} showed statistically significant differences (p = 0.001) and also the mean value of the underlying pO₂ distributions (p = 0.03). The
median and ranges of the analysed quantities were: I) 0.13 [−1.74, 12.79] cm³ for the difference between HTV and HTV\textsubscript{def}; II) 0.71 [0.15, 1.32] cm for the geometrical centre shift of the HTVs; III) 0.13 [0.05, 0.62] cm and 0.44 [0.14, 1.68] cm for the mean and the maximum distance transform, respectively; IV) 0.38 [0.00, 0.88] for the DICE similarity coefficient; V) −0.28 [−0.68, 1.65] mmHg for the differences in the HTV mean pO₂ values. Uniform dose prescription levels, theoretically depending on tumour size and pO₂ values, resulted as negligibly different (p = 0.17) from the non-parametric Wilcoxon rank test. The median and range of the dose difference in HTVs were 0.1 [−0.9, 0.9] Gy for 52.5 Gy prescribed dose to the rim between the Planning Target Volume (PTV) and the CTV (PTV-CTV).

Fig. 3 shows the box plots of the metrics considered in the study for the whole patient dataset, which characterise the differences between HTV and HTV\textsubscript{def} in terms of geometry, spatial location, underlying pO₂ distribution, and prescribed uniform dose boost level to the hypoxic compartment aiming at 95 % TCP in the CTV. It should be noted that five of the analysed patients had an HTV\textsubscript{def} equal to 0, and, therefore, none of the considered metrics, except the volume difference (HTV-HTV\textsubscript{def}), are reported in Fig. 3 in those cases.

To select the patients for treatment planning, the following variables were considered: 1) the volume of the hypoxic compartment, 2) the geometrical centre shift, 3) the mean value of the underlying pO₂ distribution within the hypoxic compartment and 4) the uniform dose boost level prescribed to the hypoxic target. As the boost dose level in the hypoxic compartment did not show statistically significant differences from the non-parametric Wilcoxon rank test, this quantity was disregarded. Furthermore, Spearman’s coefficient indicated the mean value of the underlying pO₂ to be correlated with the volume of the hypoxic compartment and was therefore disregarded (R = −0.97 for the correlation HTV with mean pO₂ in HTV and p < 0.0001, R = −0.70 for the correlation HTV\textsubscript{def} with mean pO₂ in HTV\textsubscript{def} and p < 0.004).

Contrarily, R = 0.20 for the correlation of geometrical centre shift with HTV (p = 0.48) and R = 0.13 for the correlation of geometrical centre shift with HTV\textsubscript{def} (p = 0.64). Thus, both the volume of the hypoxic compartment and the geometrical centre shift were chosen for patient selection.

The difference in HTV-HTV\textsubscript{def} and the HTV geometrical centre shifts were normalised to their corresponding maximum value and categorised as low, medium, and high. Nine representative patients with different combinations of geometrical centre shifts and changes in the hypoxic compartment volume were selected for further planning. The patient selection criterion is further described in Table 1.

Differences in the prescription doses for the selected patients in the boost volumes (i.e., Dose(HTV)-Dose(HTV\textsubscript{def})) had a median of 0.1 and range [−0.9, 0.9] Gy.

Treatment plans aiming at escalating the dose in the HTV and HTV\textsubscript{def} targets, respectively, corresponding to the rigid and deformable registration strategies similarly fulfilled the quality and dosimetric criteria for target coverage, OAR constraints, CI and HI.

Fig. 4 shows an example for one of the patients considered in the study of the treatment plans aiming at escalating the dose in the HTV (a) and HTV\textsubscript{def} (b) targets, respectively.

Fig. 5 shows box plots of the CIs and HIs in the targets of interest (HTV, GTV-HTV, CTV-GTV) for treatment plans corresponding to the RIR and DIR strategies.

The TCP RIR and TCP DIR in the CTV were calculated considering the planned dose distribution and the corresponding underlying radiosensitivity (i.e., associated with the underlying pO₂ distribution) for the rigid and deformable registration strategy, respectively (Fig. 1). All the treatment plans for both the RIR and DIR cases rendered a TCP in the CTV above 95 %, except for one patient case where the RIR TCP and DIR TCP were 94 % and 93 %, respectively. The Wilcoxon rank test showed a non-statistically significant difference between the TCP values for the RIR and DIR cases. Moreover, the TCP\textsuperscript{*} was evaluated when considering the planned dose distribution for the rigid registration strategy and the underlying radiosensitivity for the deformable registration strategy (i.e., using the deformed pO₂ distribution) (Fig. 1). For this case, all the treatment plans showed a TCP\textsuperscript{*} above 95 % in the CTV. Furthermore, the percentual difference in tumour control probability (ΔTCP in per-centual points, pp) between the RIR TCP and the TCP\textsuperscript{*} resulted in a median value of 0.4 pp with a range [−0.6, 1.4] pp. The Wilcoxon rank test results confirmed a non-statistically significant difference between the TCP RIR and TCP\textsuperscript{*}. No statistically significant differences were also found in the NTCP for the main OARs between RIR and DIR treatment plans. The average difference in NTCP between the mentioned treatment plans was below 1 pp for all the considered OARs.

4. Discussion

In this study, we investigated the impact of the PET-to-planning-CT registration method (RIR/DIR) at the pre-treatment stage on a dose painting strategy targeting hypoxia. Despite delocalisation and morphological variation of the hypoxic compartment and corresponding

![Fig. 3. Characterisation of the Hypoxic Target Volumes (HTVs). Box plots of the main metrics considered in the study and characterising the differences in the Hypoxic Target Volumes in the deformed (HTV\textsubscript{def}) and non-deformed (HTV) cases. Outliers are indicated in the figure as full rhomboids, the mean value corresponds to the star symbol, and the thick horizontal line within the box indicates the median value. Box plot whiskers show the 1.5 interquartile range (IQR) value. The characterisation of the HTV was performed on the whole cohort of 30 patients.](image-url)
dose distribution observed, the impact on treatment outcome in terms of TCP was negligible.

The algorithm of the hybrid deformable registration available in the RS TPS was extensively tested and validated and found to perform well compared to other algorithms, as discussed in Ref. [26]. At the pre-treatment stage, substantial anatomical changes may not be expected in the H&N area, thus inherently mitigating the impact of potential uncertainties in the DIR algorithm. Nonetheless, in general, DIR uncertainties are expected to depend on the particular DIR algorithm used, as shown in the study by Zhang and co-workers [19], where different algorithms available in the RS TPS (i.e. Anaconda and Morfeus) and pre-setting parameters are compared. In our study, despite the use of immobilisation masks, average vector field displacements below the cm range were observed in the CTV. Upon examining the overall shape of the vector field, it appeared reasonable to infer that these changes primarily stem from morphological variations within the pharynx and oropharynx, attributable to physiological movements like swelling and respiration. Noteworthy deviations in the deformation grid were also observed in regions beyond the CTV area, such as the shoulder region. These discrepancies could be attributed to positional differences with respect to the radiotherapy position when patients were imaged in the PET/CT scanner.

The morphological changes and delocalisation between the HTVs contoured on pO$_2$ maps derived from $^{[18]}$F-FMISO-PET images registered either with DIR or RIR methods may have to be expected, in line with the available literature [1,13], indicating the importance of using DIR in radiotherapy applications.

However, to the best of the authors’ knowledge, this is the first time that a tri-fold assessment on the impact of DIR in PET-to-planning-CT registration was performed and that other steps in the investigation related to a dosimetric and radiobiological assessment were undertaken for selected patients.

When the deformable registration vector field between the CT plan-to-CT was applied to the $^{[18]}$F-FMISO-PET image to generate a deformed PET image, a spatial interpolation of the voxel uptake among adjacent voxels was performed to determine the uptake value in the new coordinate point. Due to this interpolation process and related averaging operation, the deformed PET image pO$_2$ values were necessarily slightly higher in the HTV (Fig. 3). The largest impact of this operation could be seen in those cases where patients presented a very small HTV in the rigidly registered $^{[18]}$F-FMISO-PET image and corresponding pO$_2$ map ($<0.07$ cm$^3$) having underlying pO$_2$ values very close to the threshold of 10 mmHg (minimum pO$_2$ value $> 9.4$ mmHg). In these cases ($n = 6$), no hypoxic compartment was found in the deformed pO$_2$ map.

Although differences in the volume of the hypoxic compartment and the underneath pO$_2$ values were observed (Fig. 3), the difference in the uniform dose prescription levels in the boost volumes (Dose(HTV)-Dose (HTV$\text{def}$)) to be delivered in 35 fractions, theoretically depending on both the oxygenation of the target and number of clonogenic cells in the
actual volume (Equations (2)–(4)), were found to be close, with median 0.1 Gy and range (0.9, 0.9) Gy (Fig. 3), and the dose prescription levels did not show statistically significant differences. In most cases, a slightly higher dose prescription in the HTV with respect to the HTV_{def} was found, also because of the aforementioned averaging operations when producing the deformed PET image.

Results of the gamma analysis evidenced dose distribution differences between treatment plans aiming at a dose escalation in the HTV and HTV_{def}, respectively, on the 10 %/3 mm level. This difference results from a combination of different factors, such as dose prescription levels, morphology, and size of the hypoxic compartment (and, correspondingly, of the derived rim GTV-HTV), as well as the spatial location of the hypoxic compartment.

As in clinical settings, a deformable registration algorithm may not always be available [13], the treatment outcome in terms of TCP values was assessed considering a scenario where the actual radiosensitivity would be the one corresponding to the deformed pO$_2$ map and the treatment would be planned on the rigidly registered images. In this scenario, it might be reasonable to expect a lower TCP with respect to the TCP value calculated considering the RIR dose distributions and the corresponding radiosensitivity maps, due to a mismatch between dose distribution and radiosensitivity (Fig. 1). However, the obtained results did not reflect the initial expectations and evidenced a negligible ΔTCP between the cases under consideration (≤1.4 pp) and a non-statistically significant difference.

The negligible sensitivity with respect to the performed registration method in the TCP calculation may be attributed to the inherent robustness of the dose prescription method by contours. In fact, the formula for dose prescription by contours, aiming at achieving a 95 % TCP in the CTV (equation (4)), is intrinsically robust with respect to changes in radiosensitivity when the number of fractions is relatively high, as in the case of this study (n = 35). Starting from a distribution of oxygen at the voxel level in the volume of interest, a heterogeneous dose distribution at the voxel level that would be required to counteract the radiosensitivity is calculated (equations (2) and (3)). Subsequently, the method assumes fluctuations due to acute hypoxia in the underlying radiosensitivity and determines the corresponding homogeneous dose that would account for these fluctuations and achieve 95 % of TCP in the CTV (equation (4)). This makes the formula for dose prescription by contours intrinsically robust with respect to oxygen and, thus, dose variations when considering a relatively large number of fractions in the treatment plan. In this respect, in fact, the worst-case magnitude impact in TCP determination with respect to the imaging registration method would come by considering a treatment plan with only one fraction and a dose prescription by numbers calculated as in equation (2). Specifically, when applying a heterogeneous dose prescription calculated for the RIR case to the radiosensitivity distribution and targets considered in DIR, the impact in the TCP calculation for one single fraction treatment would be high due to a critical mismatch between the radiosensitivity map and the prescribed dose distribution. Previous studies by our group have investigated the TCP dependence on the number of fractions for different dose painting strategies in H&N cancer patients [37].

Moreover, no significant differences in terms of NTCP values were also found between the RIR and DIR plans.

Even though the treatment plan quality was comparable and satisfactory for the RIR and DIR cases in terms of dosimetric evaluation of target and OAR doses, CI and HI (Fig. 5), the physical characteristics of the actual dose distribution and the relatively smooth dose gradient in the boundaries between the different targets (HTV, GTV-HTV, CTV-GTV) counterbalance the effect of a change in the underlying...
radioresensitivity and ensure that adequate target coverage is, nonetheless, reached. In this sense, this might indicate that dose painting strategies in relatively large targets (i.e., large brush dose painting) may ensure robustness in treatment delivery, with respect to changes in underlying radioresitivity, as compared to dose painting by voxel approaches or other dose clustering methods (i.e., small brush dose painting) [9].

In conclusion, at the pre-treatment stage, the choice of the registration method of PET-to-planning-CT images impacts the localisation and morphological variation of the hypoxic compartment and corresponding dose distribution but appears to have negligible effects on tumour control. Thus, in this context, using DIR over RIR at the pre-treatment stage does not seem justifiable. However, the head and neck fixation for CT-planning and PET is mandatory.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by the Cancer Research Funds of Radiumhemmet (211232), The Swedish Cancer Society, and by the Swedish Research Council (2020-04618). The 18F FMISO trial is supported by the German Cancer Consortium (DKTK).

References