

ORIGINAL ARTICLE

## Assessment of residual setup errors for anatomical sub-structures in image-guided head-and-neck cancer radiotherapy

MILOS DJORDJEVIC<sup>1</sup>, EMMELIE SJÖHOLM<sup>1</sup>, OWE TULLGREN<sup>2</sup> & BRUNO SORCINI<sup>1</sup>

<sup>1</sup>Department of Medical Physics, Karolinska University Hospital, Stockholm, Sweden and <sup>2</sup>Department of Oncology, Karolinska University Hospital, Stockholm, Sweden

### ABSTRACT

**Background.** To quantify residual setup errors (RSE) and required planning target volumes (PTV) margins in head-and-neck cancer (HNC) radiotherapy when using daily image guidance (IG) and less-than-daily IG protocols.

**Material and methods.** Daily on-line kV-image registrations of 80 HNC patients (2640 imaged treatment fractions) were retrospectively studied to analyze RSE. Less-than-daily imaging protocols, using different action levels, were simulated on the data. To quantify local RSE; single rigid bony structures were defined as landmarks. The RSEs and required PTV margins were computed for each sub-structure with and without daily IG.

**Results.** For less-than-daily IG protocols the setup accuracy was more dependent on frequent imaging throughout the treatment course than the number of initially imaged fractions. With daily IG the RSE of the sub-structures ranged from 0.6 mm to 2.3 mm (systematic) and from 1.0 mm to 1.7 mm (random). Required PTV margins for the sub-regions ranged from 4.5 mm to 9.3 mm with no IG and from 2.3 mm to 6.8 mm with daily IG.

**Conclusion.** Anatomical changes over the treatment course require frequent IG to achieve accurate dose delivery using highly conformal radiotherapy techniques. The current study shows that considerable local RSE may remain even with daily IGRT. The comprehension of local RSEs in HNC radiotherapy is important when designating PTV margins as well as tolerance levels for couch correction and plan adaption.

Radiotherapy treatment of head-and-neck cancer (HNC) patients typically involves anisotropic target volumes and nearby organs at risk (OAR). Intensity-modulated radiation therapy (IMRT) and volumetric-modulated arc therapy (VMAT) techniques have the ability to deliver a highly conformal dose to the target with steep dose gradients and thereby decreasing the irradiated volume and sparing OARs. However, the actual conformity of the dose distribution delivered to the target volume will depend on the accuracy in which the planned patient geometry can be reproduced over the treatment course.

Setup uncertainties will largely depend on immobilization technique [1] and image-guidance (IG) procedures. Further, during the treatment course the patient may lose weight and inflammatory areas may contract resulting in increased mobility and less stable fixation [2]. In HNC radiotherapy the common clinical practice is rigid registration of bony anatomy, whether using planar or cone beam computed tomography (CBCT) images.

Setup correction protocols, using less-than-daily imaging, have shown to reduce the systematic (preparation) error to some extent [3–6]. With daily IG a considerable reduction in the overall inter-fractional setup uncertainty could be expected [7]. By implementing image-guided radiation therapy (IGRT) setup uncertainties can clearly be reduced to some degree, however the flexibility of the neck will cause relative motion between rigid bony structures resulting in different local setup uncertainties at different parts of the target volume [8–13]. As relative motion of rigid structures will deform the global target region a couch correction cannot simultaneously correct for all local setup errors. The patient positioning of HNC patients is a topic of ongoing research; electronic portal imaging device (EPID) using MV-image projections [14,15] and CT/CBCT imaging [7–13] have been utilized to quantify setup errors.

In this study we used orthogonal 2D-kV daily on-line image verification to analyze local small structure variations in HNC treatments. Single rigid

bony structures were selected to avoid relative movements within a sub-region. The high contrast spatial resolution in kV-images was suitable for the measurements of the bony landmarks used in this study. The time and labor efficiency with kV-image verification and automatic couch correction, in addition to small dose contribution ( $< 0.3$  mGy/image), allowed us to include a large number of patients with every treatment fraction imaged.

The aim of this study was to quantify residual setup errors (RSE) in HNC radiotherapy. The study consisted of two parts; the first part evaluates the impact on the global RSE using less-than-daily image guidance, assuming only rigid movements. In the second part we analyzed local RSE when using daily image guidance and on-line corrections. The local RSEs were used to estimate variable planning target volumes (PTV) margins for the head and neck region. Further, time trends during the treatment course were evaluated both for the global target and for local sub-structures.

## Material and methods

This study included 80 HNC patients that were treated with daily image-guided IMRT at the Karolinska University Hospital between 2009 and 2011. Patients diagnosed with tumors in the base of tongue, the tonsils and the hypo-pharynx were retrospectively selected. The patients were immobilized with five-point thermoplastic masks (Orfit High Precision immobilization system) covering head, neck and shoulders together with individual neck rests for optimal fixation of the whole neck region. Daily on-line image verification with two orthogonal high resolution kV images and automatic couch corrections were performed with the On-Board Imaging (OBI) system (Varian Medical Systems, Inc., Palo Alto, CA). Prior to treatment bony structures were carefully selected and outlined in the reference images (DRRs) to define the region-of-interest (ROI) for the on-line image registration performed by the therapist. The outlined structures for the group of patients included in this study included the cervical vertebrae C1–C3 and the angle of the mandible. For the majority of the patients, manual image registration was performed and when automatic registration was used the result was always verified by the therapists. Retrospectively, the setup uncertainties based on the daily alignment data were used to compute the group mean error ( $\mu$ ), the group inter-fractional random ( $\sigma$ ) and systematic errors ( $\Sigma$ ).  $\Sigma$  is defined as the SD of all patient systematic (mean) errors and  $\sigma$  is the root mean square of the patient random errors. The RSEs were calculated separately along the three axes: left-right (LR), anterior-posterior (AP) and superior-inferior (SI).

## Endpoints

Three main endpoints were evaluated in this study: 1) the impact on the patient setup error using less-than-daily imaging protocols with corrections based on the mean error for a number of treatment fractions; 2) the size of local RSEs when using daily image guidance; and 3) time dependency of the setup uncertainty during the treatment course.

### Less-than-daily correction protocols

Using the data from all the daily setup corrections, based on the image registrations performed by the therapists, less-than-daily imaging protocols were simulated retrospectively. Three main protocols were evaluated. The notation P<sub>n</sub> is used, where P stands for protocol and n = the number of imaged initial fractions.

*Protocol P0.* Only alignment according to the marks on the thermoplastic mask and no image guidance. The random and systematic errors were computed from the setup correction data from all fractions.

*Protocols P3 and P5.* Action level (AL) protocols calculating the mean setup error for a fixed number of treatment fractions to correct all the remaining fractions, without further imaging. The initial three (P3) and the initial five (P5) fractions were used for the mean shift. On-line corrections were performed the treatment fractions with imaging. Three different correction action levels (AL) were used separately for the mean error of the initial fractions: 4 mm, 3 mm and 2 mm AL. If the mean shift  $\geq$  AL the remaining alignment data were corrected with the mean value. Thus, three scenarios for P3 and P5, respectively, were evaluated. The protocols P3 and P5 were compared with a paired t-test to evaluate statistical difference in correcting the systematic error, based on all fractions. The residual group random and systematic setup errors were computed based on all fractions of the corrected alignment data for each AL separately.

*Protocol P5 + weekly.* An extended AL protocol with IG the first five fractions and then subsequent weekly imaging (every fifth fraction). The first stage is the same as protocol P5 and the same action levels were used separately. For each action level the same value was used as tolerance for the subsequent weekly imaging, e.g. when 3 mm AL (mean error  $\geq$  3 mm) was used for the mean shift of the first five fractions the weekly threshold was 3 mm (action applied when setup error  $>$  3 mm). If the weekly threshold was exceeded, additional image verifications were performed the two subsequent fractions. A mean value

was calculated based on these three fractions and if the mean value was  $\geq$ AL (same AL as for the five first fractions) the remaining fractions were corrected accordingly. Thus, three scenarios were evaluated using 2 mm, 3 mm and 4 mm for the AL and the weekly threshold.  $\Sigma$  and  $\sigma$  were computed based on all fractions after applying the P5 + weekly protocol for each scenario separately.

#### *Residual setup errors of sub-structures using daily on-line image guidance*

The consecutively last treated 40 patients were selected, from the 80 patients included in the first part of the study, to quantify local RSEs when using daily on-line IG. Retrospectively the target region was divided into ROIs defined by separate rigid bony structures: the cervical vertebral bodies C1–C5, the palatine process of maxilla and the front of the mandible (Figure 1). The RSE of each sub-structure remaining after the on-line alignment performed by the therapists was measured. The manual measuring tool with a 0.1 mm scale was used instead of re-aligning each sub-structure as the system round to integer mm. However, as the uncertainty in each measurement was assumed larger than 0.1 mm, especially due to the blurred DRRs, each measurement was rounded to nearest 0.5 mm. For the vertebrae the edges of the vertebral bodies were used for the measurements. For the AP displacement the

anterior edges were used as the tumors [gross tumor volume (GTV)] are situated in front of the vertebrae. The mean displacement of the edges represented the local translational RSE. The region for the maxilla and mandible used in the measurements are shown in Figure 1. Adding the corrected shifts performed by the treatment staff to each sub-structure the local setup errors with no IG were obtained. The residual group mean errors, inter-fractional random and systematic errors were computed for each sub-structure using daily on-line IG and compared with corresponding setup errors without IG. No correction action level was used for the daily IG protocol.

To estimate the CTV-to-PTV margin required for each sub-region we used the van Herk formula [16]:

$$M_{\text{PTV}} = 2.5 \Sigma + 0.7 \sigma \quad (1)$$

(i.e.  $M_{\text{PTV}} = 2.5 \Sigma + 1.64 (\sigma_{\text{tot}} - \sigma_{\text{p}})$ ,  $\sigma_{\text{tot}}$  = the quadratic sum of all patients SD including the SD of the penumbra  $\sigma_{\text{p}}$ , and  $\sigma_{\text{p}} = 3.2$  mm) for a minimum of 95% of the prescribed dose covering the CTV for 90% of the patients. This margin recipe is derived for translational setup uncertainties of a rigid object and the model do not account for rotations and deformations. As the head and neck region is subject to shape variations, this is a lower limit approximation. In this study we therefore chose to define the sub-regions by rigid bodies. Even though the global target region of the head and neck is subject to shape variations, we quantify the setup uncertainties in

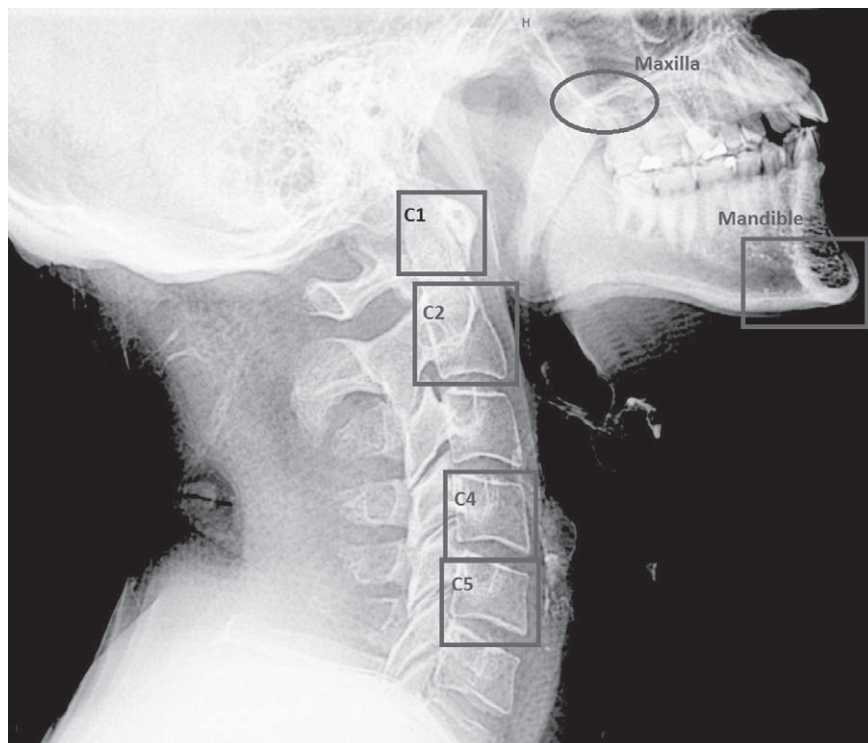


Figure 1. Region-of-interests used for the measurements of sub-structural displacements.

Table I. Patient set-up errors with the less-than-daily correction protocols P0, P5 and P5 + weekly, with different action levels (AL). Local variations are not included.

Correction protocol	Less-than-daily image verification								
	$\mu$ (mm)			$\Sigma$ (mm)			$\sigma$ (mm)		
	LR	AP	SI	LR	AP	SI	LR	AP	SI
No IG	0.0	-0.8	1.0	1.7	2.1	1.7	1.8	1.8	1.7
4mmAL	0.2	-0.5	0.4	1.5	1.8	1.6	1.6	1.9	1.6
4mmAL+ weekly	0.2	-0.4	0.4	1.4	1.7	1.6	1.6	1.9	1.6
3mmAL	0.1	-0.2	0.0	1.5	1.6	1.4	1.6	1.9	1.6
3mmAL+ weekly	0.0	0.0	0.1	1.2	1.0	1.3	1.5	1.7	1.6
2mmAL	0.0	-0.2	-0.3	1.1	1.4	1.2	1.6	1.9	1.6
2mmAL+ weekly	0.1	0.1	-0.2	0.9	0.9	1.2	1.5	1.7	1.6

AP, anterior-posterior; LR, left-right; SI, superior-inferior;  $\mu$ , mean systematic error;  $\Sigma$ , SD of systematic errors;  $\sigma$ , root-mean square of random errors.

small sub-regions defined by rigid bodies and only translational deviations are used to calculate required PTV margins for the sub-regions.

#### Time trends

To study time variations over the treatment course the five first fractions, five fractions in the middle of the treatment (fraction 16–20) and the five final fractions were compared. Two analyses were performed: 1) For each subset of five treatment fractions  $\Sigma$  and  $\sigma$  were computed for the translational shifts along the three axis based on the on-line corrections performed by the therapists. For each patient the difference in mean displacement of the five initial and five final fractions were statistically evaluated with an independent two sample t-tests. 2) To investigate time trends in deformations we analyzed changes in relative displacement between C1 and C5 in AP and SI directions, and between C5 and the maxilla in LR direction. An independent two-sample t-test was used for each patient to determine relative significant displacements between the sub-structures at the three time instances.

## Results

#### Less-than-daily correction protocols

The rigid RSEs, group mean error ( $\mu$ ), SD of patient mean errors ( $\Sigma$ ) and root mean square of the patients inter-fractional SD ( $\sigma$ ), are presented in Table I for the less-than-daily IG protocols P0, P5 and P5 + weekly with 4 mm, 3 mm and 2 mm ALs for protocol P5 and P5 + weekly, respectively. The RSEs in Table I are based on all the rigid shifts of all fractions for the 80 HNC patients after applying the correction protocols, local variations are not accounted for. The small changes in  $\sigma$  between the protocols are due to corrections performed the days with imaging. There were no statistical significant

difference ( $p < 0.05$ ) in reducing the group systematic error between protocol P5 and P3, therefore the result of protocol P3 is omitted in Table I. With  $AL = 3$  mm the systematic errors for P5 and P3 were:  $\Sigma_{P5} = (1.3; 1.6; 1.6)$  mm and  $\Sigma_{P3} = (1.5; 1.6; 1.5)$  mm in LR, AP and SI directions, respectively. Figure 2 shows the  $2\Sigma$  value for the AL protocols in Table I. If the distributions of the patients mean setup error is assumed being normal distributed,  $2\Sigma$  will geometrically encompass 95.4% of the patients mean errors. The improvement in setup accuracy (reduction in  $\Sigma$ ) with the protocol P5 + weekly was most prominent in the AP direction. With P5 + weekly and a 3 mm AL the PTV margins (using the formula  $M_{PTV} = 2.5 \Sigma + 0.7\sigma$  and the values in Table I) were reduced from 5.7, 6.5 and 5.6 mm, with no imaging, to 4.2, 3.7, and 4.5 mm in LR, AP and SI directions, respectively. Figure 3

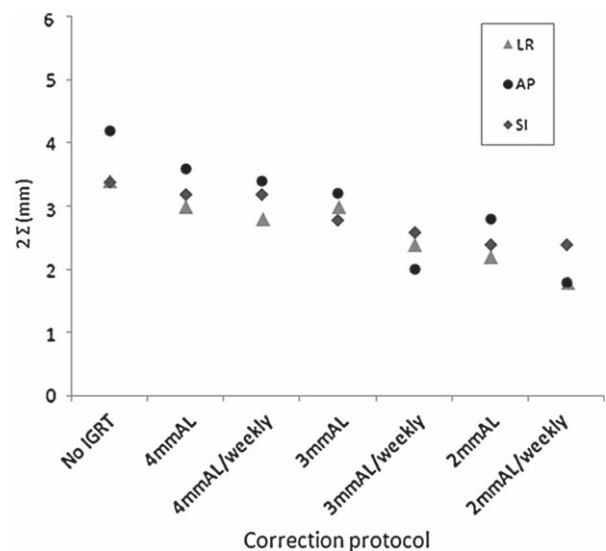


Figure 2.  $2\Sigma$  ( $2 \times$ SD of the patients mean errors) are shown for the protocols P0 (no IG), P5 and P5 + weekly; with 4 mm, 3 mm and 2 mm action levels (AL), in LR, AP and SI directions, based on the rigid registrations of the contoured ROIs.

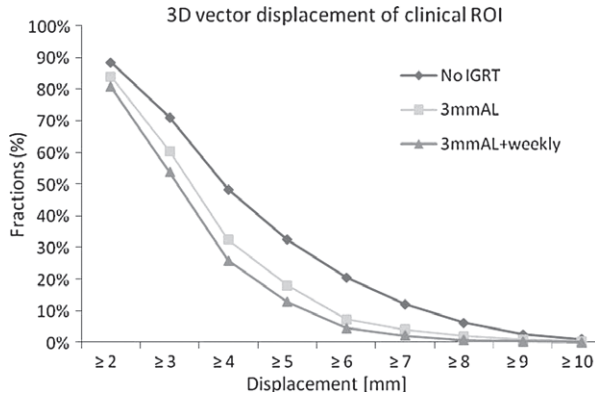


Figure 3. The cumulative distributions of the residual 3D vector displacement for all fractions of all patients with no IG and the correction protocols P5 and P5 + weekly with a 3mm AL.

shows the cumulative distributions of the residual three-dimensional (3D) vector displacement, for all fractions of all patients, of the global target for the protocols P0, P5 and P5 + weekly with 3 mm AL. With no corrections; the mean vector displacement of the global target was  $\geq 3$  mm for 51%,  $\geq 4$  mm for 27% and  $\geq 5$  mm for 8% of the patients.

#### Local residual setup errors with daily IG

To account for local variations within the target volume Table II summarizes the RSEs for the different sub-structures without IG and corresponding

residual errors with daily IG. Even with daily IG large residual setup uncertainties remains in the lower neck region (C5) and for the maxilla and mandible. Due to overlaying of skull and cheek structures in the upper neck region in the frontal images the measured lateral displacement of C1 and C2 had an increased uncertainty and are therefore not included in Table II. The largest variations of the vertebrae were observed in the AP direction. Without IG; the setup uncertainties in the AP direction for the lower neck and around the mandible (Table II) are clearly larger than the setup uncertainties based on the rigid registrations performed by the therapists (Table I). The approximate PTV margins, based on  $M_{PTV} = 2.5 \Sigma + 0.7\sigma$ , for each sub-structure are listed in Table III for no IG and daily IG. In Table IV the RSE in the AP direction for C2 has been subtracted from the RSEs of four of the sub-regions for each fraction of each patient, showing local AP-displacements due to cervical deformation relative C2. Using C2 as base-line reduced the systematic error for the maxilla but increased the systematic errors in the lower neck (vertebrae C4 and C5).

The mean vector displacement of C5 was  $\geq 3$  mm for 66% of the patients with no corrections and for 17% of the patients with daily IGRT. Figure 4 shows the cumulative distributions of the residual 3D vector displacement for the maxilla and the vertebra C5, with no IG and with daily IG.

Table II. Residual setup errors (RSE) for different bony structures in HNC radiotherapy; with and without daily IG.

	No IG								
	$\mu$ (mm)			$\Sigma$ (mm)			$\sigma$ (mm)		
	LR	AP	SI	LR	AP	SI	LR	AP	SI
Maxilla	0.4	-0.9	1.4	1.6	1.9	1.9	1.9	1.4	1.8
Mandible	0.3	-0.6	1.0	2.2	3.1	2.3	1.8	2.2	2.1
C1	N/A	-0.8	1.9	N/A	1.8	1.5	N/A	1.4	1.7
C2	N/A	-1.1	1.7	N/A	2.2	1.3	N/A	1.6	1.7
C4	-0.1	-1.3	1.8	1.7	2.9	1.4	1.9	2.3	1.8
C5	-0.3	-1.1	1.8	1.8	3.0	1.5	1.9	2.4	1.8
Residual errors with daily IG									
	$\mu$ (mm)			$\Sigma$ (mm)			$\sigma$ (mm)		
	LR	AP	SI	LR	AP	SI	LR	AP	SI
Maxilla	-0.5	-0.3	-0.2	1.6	1.9	1.3	1.6	1.5	1.3
Mandible	-0.6	-0.3	-0.6	1.1	2.3	2.1	1.5	1.7	1.5
C1	N/A	-0.3	-0.5	N/A	1.6	0.8	N/A	1.4	1.1
C2	N/A	-0.2	-0.3	N/A	0.7	0.6	N/A	1.1	1.0
C4	0.0	-0.5	-0.4	0.7	1.2	0.7	1.2	1.5	1.0
C5	0.2	-0.4	-0.4	0.7	1.5	0.7	1.2	1.6	1.0

AP, anterior-posterior; LR, left-right; SI, superior-inferior;  $\mu$ , mean systematic error;  $\Sigma$ , SD of systematic errors;  $\sigma$ , root-mean square of random errors. The image registrations were based on reference structures outlined in DRR images; including the cervical vertebrae C1 to C3 and the angle of the mandible.

Table III. Local CTV-to-PTV expansion margins with and without daily IG.

	Margins (mm)					
	No IG			Daily IG		
	LR	AP	SI	LR	AP	SI
Maxilla	5.2	5.6	5.9	5.1	5.9	4.2
Mandible	6.8	9.3	7.1	3.8	6.8	6.3
C1	N/A	5.5	4.8	N/A	4.9	2.7
C2	N/A	6.5	4.5	N/A	2.6	2.3
C4	5.5	8.9	4.8	2.5	4.1	2.5
C5	5.7	9.3	5.0	2.6	5.0	2.6

AP, anterior-posterior; LR, left-right; SI, superior-inferior. The margins are based on the van Herk formula ( $M, 2.5\Sigma + 0.7\sigma$ ).

### Time trends

Based on the rigid registrations performed by the therapists 13%, 42% and 16% of the patients had a statistical significant difference ( $p < 0.05$ ) in LR, AP and SI mean positions, respectively, between the first five and last five treatment fractions. Without any corrections, the group systematic errors based on the five final fractions were larger than for the five initial fractions in all directions (LR, AP, SI):  $\Sigma_{\text{END}} = (2.0; 2.7; 1.9)$  mm and  $\Sigma_{\text{START}} = (1.7; 2.2; 1.6)$  mm.

Comparing the relative mean positions between C5 and C2; statistical significant difference ( $p < 0.05$ ) relative the five first fractions were observed for 10% (AP) and 8% (SI) of the patients in mid-treatment and 8% (AP) and 10% (SI) in the end of treatment. Lateral displacement between C5 and the maxilla showed significant time changes for only 5% of the patients, both in the middle and in the end of the treatment course. Thus, very similar results were obtained for the middle and final fractions and most of the patients showing significant changes were the same at the two time instances, indicating that the relative movements of the sub-structures occurred during the first half of the treatment.

### Discussion

This study evaluates the improvement in global setup accuracy using less-than-daily IG protocols and

Table IV. The residual patient setup errors and margins in the anterior-posterior (AP) direction, for four bony structures, using the vertebra C2 as base-line.

	Maxilla (mm)	C1 (mm)	C4 (mm)	C5 (mm)
$\Sigma$	1.5	1.0	1.6	1.8
$\sigma$	1.3	1.3	1.4	1.6
$\mu$	0.4	0.7	-0.6	-0.3
Range of means	-3.4-2.8	-1.5-2.7	-5.2-2.5	-4.2-3.6
Margins	4.6	3.4	4.9	5.7

$\mu$ , mean systematic error;  $\Sigma$ , SD of systematic errors;  $\sigma$ , root-mean square of random errors.

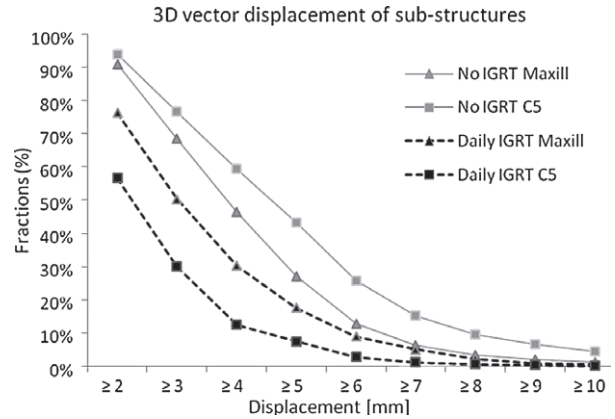


Figure 4. The cumulative distributions of the residual 3D vector displacement for the maxilla and the vertebra C5 with no IG and with daily IG.

quantifies the RSEs of bony sub-structures in the head and neck region with daily IGRT.

Correction protocols based on the mean shift of a number of initial fractions (P3 and P5) only improved the setup accuracy for patients with small random fluctuations and minor variations over time. Subsequent weekly IG was more important in reducing the setup error than the number of initially imaged fractions. Similar conclusion was made by Bertelsen et al. [17], comparing the three first fractions with the 10th and 20th fraction, recommending regular image guidance throughout the treatment courses. In agreement with Bertelsen et al. [17] we do not find the three or five first fractions representative for the whole treatment course. Imaging protocols using only a number of initial fractions may even increase the systematic error for patients with large time trends. The extended protocol with subsequent weekly IG (P5+weekly) did show predominant improvement in the AP direction which can be related to the large time dependency observed along this axis. Further, adding weekly IG had a greater impact on the setup accuracy than reducing the AL below 3 mm, especially in the AP-direction (Table I and Figure 2). Due to the accuracy of the iso-center marks on the mask, room laser calibration, imaging system and couch correction accuracy [18] we find that the AL is technically limited to 2 mm for the less-than-daily correction protocols.

Correlation clearly exists between image frequency and setup accuracy. Zeidan et al. [6] reported increased setup accuracy with more frequent imaging, in agreement with the present study. In addition the present study evaluates the impact of different action levels.

The time trend analyses showed that a large number of patients had a significant difference in AP position of the global target between start and end of treatment but a lot fewer patients showed significant

time dependent relative displacement between the vertebrae C1 and C5. The rather small changes in cervical curvature during the treatment course indicate that for most patients with large bony deformations the local deviations already existed at the start of the treatment course. Relatively large group mean errors ( $\mu$ ) were also observed for the sub-structures in AP and SI directions without daily IG (Table II), indicating a systematic deviation between treatment preparation (planning CT) and treatment execution geometry.

Based on the shifts of the rigid registrations performed by the therapists a PTV margin of 5 mm would be sufficient for protocol P5 + weekly with a 3 mm AL (using Equation 1). However, this assumes that all parts of the global target volume are accurately corrected for and no cervical deformations are present. The results for local sub-structural displacements demonstrated that even with daily IGRT considerable residual errors remained in some regions (Table II). Comparing Table I (for no IG) with Table II it is clear that without IG substantially larger setup errors were present in the lower neck and for the mandible than when assuming a rigid target volume. In the present study local residual errors with daily IG required larger margins than 5 mm in at least one direction for the lower neck region (AP), the maxilla (AP and LR) and the mandible (AP and SI). For several patients the large residual displacement of the mandible was actually caused by a backward pitch of the skull around the LR axis while the mandible remained relatively stable due to the constraint of the mask. In agreement with this study extensive local displacements of sub-regions in the head and neck area have been reported in other studies. Zhang et al. [8] used frequent CT-on-rail imaging for 14 patients, divided into groups with three different immobilizations, to analyze shifts of three ROIs, without corrections. Giske et al. [12] also used a version of CT-on-rail for a multiple ROI analyses of 45 patients using custom made immobilization device, very different from the thermoplastic mask used in the present study. Polat et al. [9] (11 patients with short face mask) and van Kranen et al. [10] (38 patients with similar immobilization as in the present study) used automatic image registration of multiple ROIs in CBCT images. The limited number of patients in the studies by Zhang and Polat results in an uncertain estimation of the systematic errors, which has the largest impact on target coverage [16]. In the study by van Kranen et al. [10] residual setup uncertainties are specific for the SAL off-line correction protocol, while in the present study the local RSEs were computed for daily imaging and corrections. Further, we defined each sub-region by a single rigid bony structure to avoid relative movements within a ROI. Even

though different immobilization devices were used in these reports similar results were obtained for the vertebral column showing a larger mobility of the lower neck, in agreement with the present study. In the present study the residual errors for the lower neck were much larger in the AP direction than in LR and SI directions. van Kranen et al. [10], using similar immobilization, reported local RSE of similar size in AP direction, but much larger RSEs in the other directions than in the present study.

Although, we selected small sub-regions to avoid deformations and extensive rotations, the estimated PTV-margins are approximations as the van Herk formula (Equation 1) was not developed for HNC radiotherapy. Corrections for the distances between the 95% and 50% isodose surfaces and  $\sigma_p$  can be performed to adjust the formula (the factor 0.7) for the treatment plans used in the clinic. The obtained  $\Sigma$  and  $\sigma$  can also be used with other margin recipes, e.g. Stroom et al. [19].

Graff et al. [13] used a different approach looking at the dose variations to target regions and OAR with different alignment procedures using MV-CBCT; recommending 7-mm margin in the floor of mouth, 5-mm margin near the hard palate and 4-mm for the lower neck, when using C1–C2-based alignment.

The current study provides a large data set in terms of both number of patients and number of fractions. The main limitation of planar kV-images is the overlaying of bony anatomy and rotations, such as roll, appear more clearly in CBCT images. However, with the high contrast spatial resolution in kV-images the bony landmarks used in this study was clearly visible, e.g. sharp edges of the vertebrae. With the workflow in our clinic 2D-kV imaging is very time efficient and provide a low amount of extra workload for the therapists, in addition to low dose contribution to the patient ( $< 0.3$  mGy/image). When bony deformations are observed CBCT images are acquired to evaluate if a new planning CT is required. In addition regular CBCT images are acquired to verify deformations of the GTV and the OAR, e.g. parotid glands. Off-line adaptive strategies using deformable registration to study dosimetric consequences in HNC IMRT has been presented [20]; so far on-line approaches are considered too time consuming.

The size of local RSEs will depend on patient setup correction method, including imaging modality and registration technique, e.g. size and location of ROIs used for the registration [13]. Different approaches to optimize the couch correction and derive criteria for re-planning have been reported [21–23]. With the knowledge of patient setup accuracy in the clinic, different safety margins may be assigned for the upper and lower neck as well as in

the different directions. Yang et al. [24] reported a method for creating variable margin around the clinical target volume (CTV) to account for local variations in HNC radiotherapy.

In the present study we evaluated the improvement in setup accuracy using less-than-daily imaging protocols, analyzed local RSEs with daily IG and time trends during the treatment course. The results of the present study suggest that a variable margin, related to the tumor site, is a reasonable approach to manage minor cervical deformations. However, frequent IG and tolerance levels for bony deformations, related to the PTV margin used in the clinic, is required to make decisions for treatment adaption. We recommend that the image registration is based on a small region according to GTV and OARs, and tolerance levels are used for local displacements at distal parts in the target volume.

**Declaration of interest:** The authors report no conflicts of interest. The authors alone are responsible for the content and the writing in the paper.

## References

- [1] Hansen CR, Christiansen RL, Nielsen TB, Bertelsen AS, Johansen J, Brink C. Comparison of three immobilization systems for radiation therapy in head and neck cancer. *Acta Oncol Epub* 2013 Sep 24.
- [2] Barker JL, Jr., Garden AS, Ang KK, O'Daniel JC, Wang H, Court LE, et al. Quantification of volumetric and geometric changes occurring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. *Int J Radiat Oncol Biol Phys* 2004;59:960–70.
- [3] de Boer HCJ, Heijmen BJM. A protocol for the reduction of systematic patient setup errors with minimal portal imaging workload. *Int J Radiat Oncol Biol Phys* 2001;50:1350–65.
- [4] de Boer HCJ, van Sörensens de Koste JR, Creutzberg CL, Visser AG, Levendag PC, Heijmen BJM. Electronic portal image assisted reduction of systematic set-up errors in head and neck irradiation. *Radiother Oncol* 2001;61:299–308.
- [5] de Boer HCJ, Heijmen BJM. eNAL: An extension of the NAL setup correction protocol for effective use of weekly follow-up measurements. *Int J Radiat Oncol Biol Phys* 2007;67:1586–95.
- [6] Zeidan OA, Langen KM, Meeks SL, Manon RR, Wagner TH, Willoughby TR, et al. Evaluation of image-guidance protocols in the treatment of head and neck cancers. *Int J Radiat Oncol Biol Phys* 2007;67:670–7.
- [7] Den RB, Doemer A, Kubicek G, Bednarz G, Galvin JM, Keane WM, et al. Daily image guidance with cone-beam computed tomography for head-and-neck cancer intensity-modulated radiotherapy: A prospective study. *Int J Radiat Oncol Biol Phys* 2010;76:1353–9.
- [8] Zhang L, Garden AS, Lo J, Ang KK, Ahamad A, Morrison WH, et al. Multiple regions-of-interest analysis of setup uncertainties for head-and-neck cancer radiotherapy. *Int J Radiat Oncol Biol Phys* 2006;64:1559–69.
- [9] Polat B, Wilbert J, Baier K, Flentje M, Gluckenberger M. Nonrigid patient setup errors in the head-and-neck region. *Strahlenther Onkol* 2007;183:506–11.
- [10] van Kranen S, van Beek S, Rasch C, van Herk M, Sonke JJ. Setup uncertainties of anatomical sub-regions in head-and-neck cancer patients after offline CBCT guidance. *Int J Radiat Oncol Biol Phys* 2009;73:1566–73.
- [11] Ahn PH, Ahn AL, Lee CJ, Shen J, Miller E, Lukaj A, et al. Random positional variation among the skull, mandible, and cervical spine with treatment progression during head-and-neck radiotherapy. *Int J Radiat Oncol Biol Phys* 2009;73:626–33.
- [12] Giske K, Stoiber EM, Schwartz M, Stoll A, Muentner MW, Timke C, et al. Local setup errors in image-guided radiotherapy for head and neck cancer patients immobilized with a custom-made device. *Int J Radiat Oncol Biol Phys* 2011;80:582–9.
- [13] Graff P, Kirby N, Weinberg V, Chen J, Yom SS, Lambert L, et al. The residual setup errors of different IGRT alignment procedures for head and neck IMRT and the resulting dosimetric impact. *Int J Radiat Oncol Biol Phys* 2013;86:170–6.
- [14] Mongioj V, Orlandi E, Palazzi M, Deponti E, Marzia F, Stucchi C, et al. Set-up errors analyses in IMRT treatments for nasopharyngeal carcinoma to evaluate time trends, PTV and PRV margins. *Acta Oncol* 2011;50:61–71.
- [15] Pehlivan B, Pichenot C, Castaing M, Auperin A, Lefkopoulos D, Arriagada R, et al. Interfractional set-up errors evaluation by daily electronic portal imaging of IMRT in head and neck cancer patients. *Acta Oncol* 2009;48:440–5.
- [16] van Herk M, Remeijer P, Rasch C, Lebesque JV. The probability of correct target dosage: Dose-population histograms for deriving treatment margins in radiotherapy. *Int J Radiat Oncol Biol Phys* 2000;47:1121–35.
- [17] Bertelsen A, Nielsen M, Westberg J, Jensen HR, Brink C. The representativeness of patient position during the first treatment fractions. *Acta Oncol* 2009;48:259–66.
- [18] Djordjevic M. Evaluation of geometric accuracy and image quality of an on-board imager (OBI), MSc thesis. Medicinsk Strålningsfysik. Stockholm: Stockholm University and Karolinska Institute; 2007. p. 30–50 and p.66.
- [19] Stroom JC, de Boer HCJ, Huizenga H, Visser AG. Inclusion of geometrical uncertainties in radiotherapy treatment planning by means of coverage probability. *Int J Radiat Oncol Biol Phys* 1999;43:905–19.
- [20] Elström UV, Wysocka BA, Muren LP, Petersen JBB, Grau C. Daily kV cone-beam CT and deformable image registration as a method for studying dosimetric consequences of anatomic changes in adaptive IMRT of head and neck cancer. *Acta Oncol* 2010;49:1101–8.
- [21] van Kranen S, van Beek S, Mencarelli A, Rasch C, van Herk M, Sonke JJ. Correction strategies to manage deformations in head-and-neck radiotherapy. *Radiother Oncol* 2010;94:199–205.
- [22] Stoiber EM, Schwartz M, Huber PE, Debus J, Bendl R, Giske K. Comparison of two IGRT correction strategies in postoperative head-and-neck IMRT patients. *Acta Oncol* 2013;52:183–6.
- [23] Stoiber EM, Schwartz M, Debus J, Bendl R, Giske K. An optimized IGRT correction vector determined from a displacement vector field: A proof of principle of a decision-making aid for re-planning. *Acta Oncol Epub* 2013 Apr 25.
- [24] Yang J, Garden AS, Zhang Y, Zhang L, Dong L. Variable planning margin approach to account for locoregional variations in setup uncertainties. *Med Phys* 2012;39:5136–44.