Physica Medica 102 (2022) 66–72

Contents lists available at ScienceDirect

Physica Medica

journal homepage: www.elsevier.com/locate/ejmp

Suitability of propagated contours for adaptive replanning for head and neck radiotherapy

David Nash [a,b,*], Antony L. Palmer [a], Marcel van Herk [b,c], Alan McWilliam [b,c], Eliana Vasquez Osorio [b,c]

[a] Department of Medical Physics, Portsmouth Hospitals University NHS Trust, Portsmouth, UK
[b] Division of Cancer Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester, UK
[c] The Christie NHS Foundation Trust, Manchester, UK

ARTICLE INFO

Keywords:
Deformable image registration
Non-rigid registration
Contour propagation
Adaptive radiotherapy

ABSTRACT

Purpose: Adaptive radiotherapy relies on rapid recontouring for replanning. Contour propagation offers workflow efficiencies, but the impact of using unedited propagated OAR contours directly during re-optimisation is unclear.

Methods: Plans for ten head and neck patients were created on the planning CT scan. OAR contours for the spinal cord, brainstem, parotids and larynx were then propagated to five shading-corrected CBCTs equally spaced throughout treatment using five commercial packages. Two reference contours were created on the CBCTs by (1) a clinician and (2) a geometric consensus from the propagated contours. Treatment plans were re-optimised on each CBCT for each set of contours, and the DVH statistic differences to the reference contours were calculated. The spread of DVH statistic differences between the 5th and 95th percentiles was quantified.

Results: The spread of DVH statistic differences was 3.7 Gy compared to the clinician contour and 3.3 Gy compared to the consensus contour for the brainstem (and PRV) and 2.4 Gy and 2 Gy for the spinal cord (and PRV), across all 5 auto-contouring solutions. The parotids and larynx showed differences of 3.7 Gy compared to the clinician and 0.9 Gy to the consensus contour, with the larger difference for the clinician possibly caused by uncertainty in the clinician standard due to poor image quality on the CBCTs.

Conclusions: Propagated OAR contours can be used safely for adaptive radiotherapy replanning, however, where organ doses are close to clinical tolerance then the contours should be reviewed for accuracy regardless of the propagation software used.

Introduction

Radiotherapy forms part of the treatment for approximately 80% of head and neck cancer patients [1]. They receive an intensity modulated radiotherapy (IMRT) treatment such as volumetric modulated arc therapy (VMAT), typified by sharp dose gradients. Such treatments have been demonstrated to reduce some of the side-effects of radiotherapy, specifically xerostomia [2]. These treatments rely on a high degree of patient immobilisation with head/neck shells and regular on-treatment imaging, usually cone beam computed tomography (CBCT). These images are used to correct for any rigid body translations and rotations, a process termed "image guided radiotherapy (IGRT)." However, IGRT is unable to correct for any anatomical changes, such as weight loss or changes in target location or shape [3]. Changes to the clinical target volume (CTV) often occur, for instance shrinking during treatment [3–7], and can result in suboptimal treatment, either by geometric miss or increased dose to normal tissues.

Adaptive radiotherapy (ART) is an evolution of IGRT where the treatment plan is modified to account for anatomical changes [3]. Offline ART occurs between treatment fractions, but can only correct for systematic anatomical or physiological changes. Online ART happens during the treatment fraction and can correct for random and systematic changes, assuming these stay constant during delivery and plan creation [5,8]. Online or offline ART requires repeat imaging, either CT, CBCT or MR-guided, along with time-consuming repeat contouring [9,10]. Imaging and replanning the patient at the treatment machine in the treatment position is ideal, using either CBCT or MR but this has obvious workload and workflow issues, in particular for contouring.

* Corresponding author at: Department of Medical Physics, Portsmouth Hospitals University NHS Trust, Portsmouth, PO6 3LY, UK.
E-mail address: david.nash@porthosp.nhs.uk (D. Nash).

https://doi.org/10.1016/j.ejmp.2022.09.002
Received 11 March 2022; Received in revised form 19 August 2022; Accepted 12 September 2022
Available online 18 September 2022

Non-rigid registration, or deformable image registration, offers a potential alternative to manual contouring by propagating the original planning CT contours to the repeat imaging. Significant work has been reported evaluating the geometric accuracy of propagated OARs on CBCT [11–17] (considering prostate cancer, pelvic, lung and head and neck), with relatively good results although manual correction is often required [18–20] (head and neck), with differing views on how often the contours need correcting in the head and neck region [20–22]. Previous studies have assessed the accuracy of dose calculation when using propagated contours for head and neck cancer [17,23], showing they are suitable for plan assessment with a view to schedule replanning, but no studies have looked at the impact of directly using propagated contours for replanning.

In this work, we investigated the impact of using propagated contours for replanning and whether use of propagated contours produced acceptable replans. Using the propagated OAR contours from five deformation packages, previously assessed for geometric and dosimetric differences when recalculated on CBCTs [24], we evaluated whether the plans created with these OAR contours were of similar quality to those based on gold and silver standards. This will inform whether non-edited propagated contours can directly be used to create acceptable replans based on on-treatment CBCT imaging, potentially reducing the workload pressures in an (online, day-to-day) ART pathway.

Methods and materials

Patient selection and propagation of contours

This work used ten arbitrarily selected head and neck cancer patients from a single institution. The patients consisted of 5 oropharyngeal, 2 oral cavity, 1 hypopharynx, 1 supraglottic and 1 of unknown primary (target below nasal region) patients with bilateral nodal treatment. Each patient was selected such that the brainstem was fully visualised on CBCT imaging. Each patient was retrospectively replanned with VMAT on the planning CT (pCT) with 65.1 Gy in 30 fractions to the high risk PTV and 54 Gy in 30 fractions to the low risk PTV. The plans were assessed by an experienced Medical Physics Expert with >5 years' experience reviewing head and neck plan quality, by comparing isodoses and against the planning goals. The planning goals are given in Table 1.

Five CBCTs spread through treatment were selected, and each underwent a "shading correction" [25] to improve the Hounsfield units to within 1 % of the true value and dose accuracy to within 1 % whilst also improving image quality [26]. The spinal cord, brainstem, parotid and larynx contours were propagated from the pCT to the five CBCTs, using five deformable image registration packages: (1) RayStation v7 (RaySearch Labs, Sweden), (2) ADMIRE v1.13 (Elekta, UK), (3) Mirada v1.6 (Mirada Medical Systems, UK), (4) ProSoma v4.1 (Medcom, Germany) and (5) Pinnacle^3 v16.0 (Philips, Netherlands). Note that one patient did not have a larynx contour.

Clinical gold standard reference contours were outlined on each of the CBCTs by a single oncologist who had access to the pCT contours to ensure consistency, whilst a second silver standard set of contours was generated by creating a STAPLE (simultaneous truth and performance level estimation) contour [27] from the 5 propagated contours. The terminology "silver standard" has been used in other fields [28,29], but we are adopting it for this work. This was due to deficiencies in the image quality of the CBCTs for contouring and can be considered as representative of the "average" propagated contour and are therefore assessing the variation in the resulting treatment plans because of the different registration algorithms.

Table 1
Planning PTV and OAR constraints.

Volume	Constraint	Objective
High dose PTV	D95%	>61.85 Gy
	D2%	<68.36 Gy
Low dose PTV	D95%	>51.30 Gy
Spinal cord	D1cc	Mandatory: <44 Gy
Spinal cord PRV	D1cc	Mandatory: <46 Gy
Brainstem	D1cc	Mandatory: <52 Gy
Brainstem PRV	D1cc	Mandatory: <54 Gy
Parotids	Average dose	Optimal: <26 Gy
Larynx	Average dose	As low as reasonably practicable

Planning target volumes (PTVs) were generated on the CBCTs by rigidly propagating the clinical target volumes (CTVs) from the pCT to each of the CBCTs and then applying a 4 mm isotropic expansion in RayStation. We chose to rigidly propagate the CTVs to keep a conservative approach to replanning as it has been shown that tumours may dissolve rather than shrink [30] and thus re-outlining the CTV may be inappropriate. This is the approach recommended by Sonke et al [3]. However, a full re-evaluation and propagation of target structures is not performed and is outside the scope of this study. PTVs were clipped 5 mm from the skin and these used for planning and reporting. Planning risk volumes (PRVs) were generated from the propagated contours by applying a 5 mm isotropic expansion. The DVH statistics for the original plans are reported.

For completeness, we report the geometric accuracy of the propagated contours [23], with the mean distance to agreement (mDTA) [31] and Dice similarity score (DSC) [32]. These were calculated for each of the organs compared to the gold and silver standards and reported using violin plots for each OAR. Additionally, we compared the gold and silver standards using the same metrics.

Replanning

Replanning on each of the CBCTs was accomplished via the RayStation scripting interface; the original pCT optimisation solution was copied to the CBCTs and contours were updated for each propagation contours set (RayStation, Pinnacle, ProSoma, Mirada and ADMIRE) or gold and silver standards. The plan was then optimised 5 times sequentially, and final coverage and organ at risk constraints were assessed. The same planning goals as the original plan were used (Table 1) and were prescribed to the median dose of the 65.1 Gy PTV. If the CTV coverage was not considered acceptable based on an isodose assessment, the optimisation solution was adapted. This gave a total of 350 plans; 5 plans per patient per CBCT for Mirada, ProSoma, Pinnacle, RayStation and ADMIRE and a further 2 plans for the gold and silver standards, for 10 patients. For the 65.1 Gy PTV the D95%, median and D2% DVH statistics were extracted and assessed for all patients (mean and 5th and 95th percentiles compared); for the 54 Gy PTV, the D95% and median were assessed. The number of plans that did not achieve the PTV coverage goals was reported and by how much.

Dose volume histogram (DVH) statistics comparison

The organ at risk (OAR) dose volume histogram (DVH) statistics of all structures in Table 1 were extracted for each of the plans. The parotids were divided into two groups – those where the planning tolerance was achieved on the pCT plan (the "spared" parotids) and those where the planning tolerance was exceeded (the "treated" parotids). The DVH statistics differences were then calculated between the plans created using the propagated contours and the plans based on the gold and silver standard contours. Note that all DVH statistics were assessed using the gold and silver standard contours, resulting in two sets of DVH statistics differences. Both sets of data for each standard is useful, as with any concerns regarding image quality this gives an understanding of the uncertainty in the dose estimates. Violin plots were created for all dose differences for each organ, and then grouped by CBCT and propagation package, again for each organ. The median, the 5th percentile and 95th percentile dose differences were calculated, along with the Wilcoxon sign-rank p-value. The DVH statistics for the plans created for the gold

and silver standards were plotted against the DVH statistics for the plans created with the propagated contours for each of the organs.

Clinical goals

Using the OAR DVH statistics listed in Table 1, the DVH statistics for the OARs for each of the replans using the gold and silver standard contours were reviewed. For the serial organs (spinal cord, and its PRV, and brainstem, and PRV) the D1cc statistics were compared against the clinical tolerance. Any instances of the organ failing tolerance were determined, and by what degree. The median and range between the 5th and 95th percentiles (termed the 90-percentile range) of data for each organ is presented. For reference, the volumes of the treated and spared parotids were extracted and plotted for the pCT and each CBCT.

Results

The DVH statistics for the original plans are shown in the supplementary materials, Table S1. Initially for the replans, nine out of 10 patients had acceptable CTV coverage based on visual assessment of isodoses, but one patient did not. This one failing patient showed considerable weight loss, and the optimisation solution had to be modified before continuing. The optimisation solution was then modified for a plan based on a selected CBCT and set of contours to achieve adequate coverage. This required the optimiser to be run three times without change to the optimisation solution between optimisations. The amended optimisation solution was then copied to all other plans for that patient and reoptimised. After this occurred, a further 45 plans (18 %) plans failed the D95% constraint for the 54 Gy PTV only, but upon review of the CTV coverage were considered acceptable and the PTV D95% was only missed by a small amount (5th percentile 51.1 Gy compared to the 51.3 Gy intended value, a loss of 0.2 Gy), thus all patients had acceptable CTV coverage. Table 2 gives the mean of the median dose, D95% and D2% for the 65.1 Gy PTV and the mean of the median and D95% for the 54 Gy PTV, along with 5th and 95th percentiles. All parameters were in tolerance except for the lower 5th percentile of the D95% for the 54 Gy PTV, which falls short of the tolerance by 0.2 Gy. All OAR (except treated parotids) tolerances were achieved when assessed using the propagated contours.

The median mDTA for the propagated contours compared to the gold standard was within <3 mm and within 1 mm for the silver standard, for all organs. The DSC was generally high, with DSC > 0.7 for most organs compared to the gold standard and >0.9 for the silver standard. Poorer results were seen compared to the gold, rather than silver standards (see figure S1 in supplementary materials). The mDTA between the gold and the silver standard contours was 2.1 ± 0.8 mm and the DSC was 0.70 ± 0.11 for all contours.

DVH statistics comparison

Fig. 1 summarises the change in DVH statistics for the standard contours between the plans created with the propagated contours and the gold and silver standard contours for each organ, whilst the 90-percentile, median and Wilcoxon p-values are shown in Table 3. Generally, the silver standard contours show improved dosimetric performance with the 90-percentile range within 2.03 Gy for all organs except for the brainstem and associated PRV compared to the propagated contours for all plans. The spinal cord and PRV have a wider dose range than the parotids and larynx, and the brainstem the widest dose range of 3.26 Gy. The gold standard shows broader spread of data compared to the propagated contours, with, excluding the brainstem, up to 3.21 Gy difference, particularly for the parotids which show a greater spread and shallower distribution (Fig. 1) than the silver standard. The brainstem is again the widest, at up to 3.46 Gy. Considering variation by CBCT and package, no significant differences were noted, suggesting the dose differences are not sensitive to time or package. The violin plots for the change in dose grouped by CBCT and package can be seen in the supplementary materials (figures S2 and S3 respectively). The plots of the DVH statistics for the plans created for the gold and silver standards against the DVH statistics for the plans created with the propagated contours for each of the organs can be found in the supplementary materials, figure S4.

Table 2
Coverage parameters for the 65.1 Gy PTV and the 54 Gy PTV, giving the median and 5th and 95th percentiles for all 350 plans.

	Dose (Gy)		
	Mean	5th percentile	95th percentile
65.1 Gy PTV			
Median	65.1	65.0	65.3
D95%	62.6	62.2	63.3
D2%	66.7	66.0	67.5
54 Gy PTV			
Median	54.2	54.0	54.7
D95%	51.9	51.1	52.7

Clinical goals

None of the D1cc statistics for the brainstem, brainstem PRV or spinal cord itself exceeded clinical tolerance when the plans created with the propagated contours were analysed with the gold and silver standard contours. However, the spinal cord PRV tolerance was exceeded on 7 plans when assessed with the gold standard contours (by 0.1 to 1.4 Gy) and on 2 plans (by 0.2 Gy) using the silver standard contours. It should be noted that on the plans created with the gold and silver standards these organs did not fail the clinical tolerance. The gold and silver standards failed to achieve the tolerance on different plans. Of the gold standards, all but one of the failing plans was within 0.7 Gy with the location of the 1.4 Gy failure located where the spinal cord transitions into the brainstem.

Fig. 2 shows the summary of the dose statistics for the plans, showing the median DVH statistics for the serial organs (D1cc shown) and the parallel organs (mean dose shown) and the 90 percentile ranges. The range of doses around the median dose constraint is variable depending on which organ is reviewed. The brainstem has the largest range, reflecting that the brainstem dose depends strongly on where the primary PTV is located ranging from very low doses to doses close to the clinical tolerance. The spinal cord dose has a much smaller range, reflecting the nature of these cases where the high dose PTV is generally close too or even wrapping around the spinal cord. The treated parotids show a wide range of possible doses, probably due to high dose gradients close to these organs. A decreasing trend of parotid volume was observed, as can be seen in the supplementary materials (figure S4).

Discussion

The majority of previous studies on propagated contours have presented a geometric comparison of contours [11–17], with a limited number also recalculating the original treatment plan on treatment imaging CBCT [17,23,33], but to our knowledge no studies have assessed the use of propagated contours for replanning and to evaluate the suitability of such contours for plan creation. In this work, for ten head and neck cancer patients, we reoptimized treatment plans on five CBCTs spread throughout treatment using various propagated contours. The dose distributions generated were assessed using the gold and silver standard delineations and dose statistics compared with plans generated with the gold and silver standard contours. Clinical goals were also assessed for the replans, again using the gold and silver standard delineations. We have shown that DVH statistic differences from the gold standard contours of up to 3.7 Gy are possible when using propagated

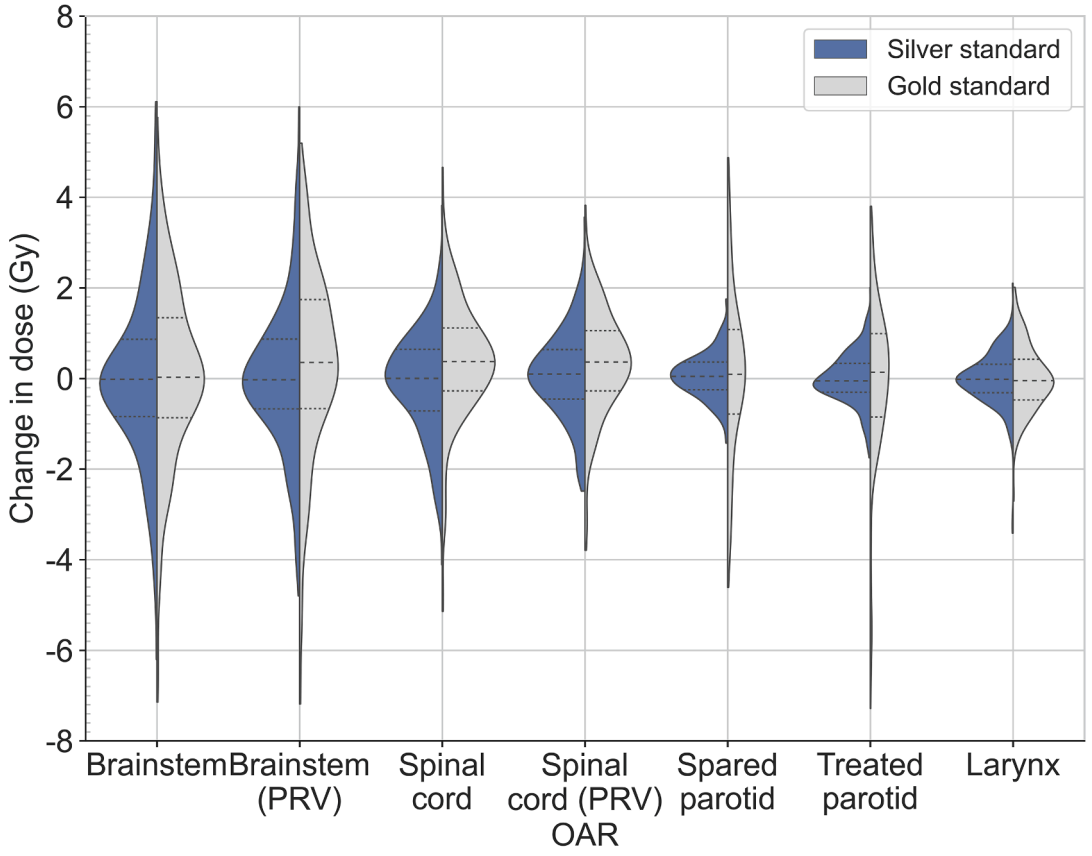


Fig. 1. Violin plots showing the DVH statistics for the gold standard contours between the plans created with the propagated contours and the silver and gold standard contours for each organ, using the gold and silver standards for 10 patients and 350 plans. Dashed lines indicate median and interquartile ranges.

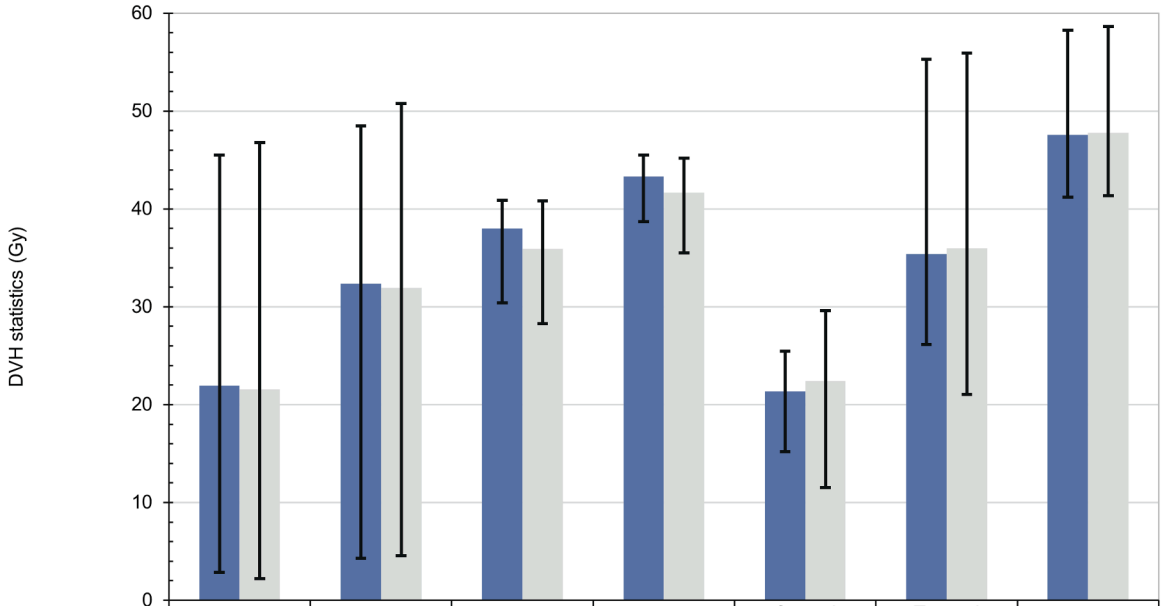
Table 3
The median and 90-percentile range, with p-values, for the dose differences for each organ for 10 patients and 350 plans analysed with the gold and silver standard contours.

Organ	DVH statistic	5th percentile (Gy)	Median (Gy)	95th percentile (Gy)	p-value
	Gold standard				
Spinal cord	D1cc	−1.84	0.37	2.37	<0.01
Spinal cord (PRV)	D1cc	−1.63	0.36	2.19	<0.01
Brainstem	D1cc	−3.13	0.03	3.26	0.289
Brainstem (PRV)	D1cc	−3.46	0.36	3.70	<0.01
Spared parotid	Mean	−3.27	0.10	3.21	0.283
Treated parotid	Mean	−3.66	0.14	2.39	0.419
Larynx	Mean	−0.89	0.18	1.86	<0.01
	Silver standard				
Spinal cord	D1cc	−2.03	0.01	1.36	0.19
Spinal cord (PRV)	D1cc	−1.55	0.11	1.79	0.017
Brainstem	D1cc	−3.22	0	3.08	0.92
Brainstem (PRV)	D1cc	−2.7	−0.01	3.26	0.614
Spared parotid	Mean	−0.7	0.09	0.81	0.04
Treated parotid	Mean	−0.86	−0.01	0.64	0.885
Larynx	Mean	−0.93	−0.02	0.9	0.981

contours. All of the plans created were considered to have acceptable coverage, with organ at risk tolerances being achieved when assessed with the propagated contours.

Whilst the median dose differences (Table 3) are generally small and rarely statistically significant (4 out of 14 comparisons), the spread of dose differences is quite wide. This illustrates that different contours can result in different dose distributions, leading up to 3.7 Gy difference in some DVH statistics. The replans when assessed with the gold standard contours generally show a greater spread than the silver standard contours, as would be expected considering the poor image quality of the CBCTs [23], but the silver standard contours do still show a broad range of DVH statistics. This is particularly marked for the serial organs (brainstem and spinal cord) where the replans assessed with the gold standard contours show differences up to 3.7 Gy while the silver standard is up to 3.26 Gy. These differences can likely be explained by sharp dose gradients in close proximity to the brainstem and spinal cords, with just a few millimetres difference on the contours enough to induce significant changes to the organ dose. The results imply that the spinal cord and brainstem propagated contours can reliably predict dose up to 3.7 Gy, in these 10 cases, of the dose tolerance, and that if an organ approaches approximately 3.7 Gy or is planned to less than this dose of the dose tolerance, close inspection of the contour may be warranted to ensure the organ is not being overdosed. Considering Fig. 2, the median doses to the spinal cord and brainstem are <3.7 Gy below tolerance so this would mean that the majority of plans and contours would not require further review, thus creating efficiencies in the process.

The brainstem and brainstem PRV show larger variation than the spinal cord and spinal cord PRV (up to 3.7 Gy compared against 2.4 Gy), for both standards. This is potentially explained by the variability of



	Brainstem-D1cc	Brainstem PRV-D1cc	Spinal cord-D1cc	Spinal cord PRV-D1cc	Spared parotid-Mean dose	Treated parotid-Mean dose	Larynx-Mean dose
Silver standard	21.9	32.4	38.0	43.3	21.3	35.4	47.5
Gold standard	21.6	32.0	35.9	41.7	22.4	36.0	47.8

Fig. 2. The median DVH statistics for each of the organs with the 90-percentile range indicated by the error bars for the gold and silver standard contours for 250 plans and 10 patients.

D1cc doses planned to the brainstem. This is shown in Fig. 2, where the brainstem's D1cc had a 90-percentile range of >40 Gy (e.g., going from 2.2 to 46.8 Gy for the gold standard) compared to approximately 10 Gy (e.g., from 28.3 to 40.8 Gy for the gold standard) for the spinal cord. This suggests there is much more potential for the dose gradients to shift relative to the brainstem than the spinal cord between patients depending on the accuracy of the contour.

The parotids show a larger spread for the gold standard than using the silver standard. In figure S4, it is demonstrated that the propagated parotids did not vary as much from the silver standard as they did from the gold standard in terms of volume, with the gold standard contours generally being smaller. This would therefore support the results seen here, whereby the silver standard DVH statistics show greater consistency. If there were concern about parotid dose during treatment, again doses within 3.7 Gy of tolerance should trigger inspection of the contours. However, as parotids are generally planned close to tolerance, this may mean most plans are reviewed and setting a different dose tolerance (e.g. 30 Gy rather than 27 Gy) may be appropriate for the parotids at replan, and potentially only spared parotids investigated, but this should be investigated as part of a clinical trial. A similar consideration should be applied to the larynx, but due to the more consistent dose bath present through the larynx smaller differences are seen for this, suggesting the larynx may benefit least from review.

A limitation of this study is that it is entirely retrospective, based on patients treated with different plans, for only 10 patients, none of which have large anatomy changes – but it is typical of what might be encountered in a day-to-day online ART workflow as encountered on the Varian (Palo Alto, CA) Ethos or Elekta Unity. This is not believed to have negatively impacted on the present study, but it does mean further work is required, probably within a clinical trial, to assess the impact in a prospective setting for more patients. However, the work does still support the use of non-rigid registration to improve the adaptive radiotherapy workflow as it demonstrates the uncertainty introduced by using propagated contours for planning is small.

A further limitation of this work is that the patient selection was arbitrary, with only one patient showing considerable weight loss during treatment. Besides this patient, weight loss was noted in 3 further patients. This is broadly in agreement with the literature, where about 32% of patients experience significant weight loss during treatment) [34]. Even though testing a clinically representative sample is essential to demonstrate applicability to a wider population, a similar study only including patients with significant changes can help us demonstrate the full potential benefit of adaptive radiotherapy. However, if the largest geometric changes occur only for bulky, fast responding, tumours, these may not directly impact the OARs. In particular, we believe that our results are directly applicable to relatively fixed organs (e.g., brainstem) but we could gain extra insights on organs prone to deformation, such as the parotid glands.

The small number of patients used in our study is also an issue. A limited number of patients (ranging 4–16) is typical of similar studies in the literature [11,12,14,16,18,20,35,36] due to the labour intensive steps required, e.g. planning/replanning. Extending this type of studies to larger samples would require the use of 'smarter' tools to replace these manual steps, for example using deep learning [37].

It should be noted that the standard contours are not identical, as shown by the mDTA and DSC – it was noted that due to the poor image quality of the CBCTs, the clinician struggled with low contrast details such as the parotids, spinal cord and brainstem and often had to rely on bony landmarks. In terms of the geometric accuracy of the propagated contours, these were more similar to the silver standard than to the gold standards (mDTA and DSC in figure S1). However, when comparing the plans created using the propagated contours to the plans created using the gold standards, small differences were seen (Fig. 1 and Table 3). For example, the brainstem 5th percentile is −3.13 Gy for the gold standard and −3.22 Gy for the silver standard, however, the mDTA is 2.9 mm and 0.6 mm respectively. This suggests that the local dose environment to the OARs has a greater impact than the geometric accuracy of the contour.

The use of the silver standard contour may be considered controversial, however interobserver variability in CT scans for head and neck

cancer is known to be in the range of millimetres, for example with mean Hausdorff distance (HD95%) ranging from 3.1 mm to 14.5 mm as shown by van der Veen et al [38]. This variability is expected to increase for CBCT contouring. On the other hand, CBCT to planning CT registrations have been shown to be relatively stable (for example, distance transform HD95% mean ranging 2.7 ± 1.0 mm in the study by Veiga et al [35]). This justified using a STAPLE contour comparison, as this provides the reader an insight of the variation of the data partially due to image quality concerns.

The use of rigid registration for the CTVs is also not ideal, as this does not reflect any changes in CTV size which may be expected, but recontouring the CTV was considered outside of the scope of this study. However, recontouring on account of presumed decrease in CTV size is presumptuous, as it has been shown that head and neck tumours dissolve rather than shrink [30], and therefore a rigid registration of a CTV may be most appropriate considering the uncertainties with image quality. This approach is recommended by Sonke et al [3] with CTV re-outlining only to be performed within clinical trials. If the CTV included nodal chains, as all the patients in this study did, it might be prudent to review the contours to ensure the nodes are still inside the CTV but this is difficult on CBCT imaging and is considered beyond the scope of this work. Another limitation is the silver standard contour is not fully independent of the individual package contours, which may explain part of the smaller differences observed.

The loss of coverage of some of the replans is a potential limitation of this study although only a small number did not achieve the planning tolerance. It is conceivable that using a tighter planning tolerance and assessing the replans to a lower standard (for instance, planning to D98% > 95% of prescription but reviewing replans to D95% > 95% of prescription) would have resulted in more plans achieving coverage, but an assessment of this was beyond the scope of this study. Nor is it believed this would affect the conclusions related to the adequateness of the propagated OAR contours for treatment planning, but further work where the CTV is manually outlined on the CBCTs (if considered appropriate) would allow for a more thorough assessment of this.

Using the CTV coverage to assess the quality of treatment plans is potentially sub-optimal, but with methods to assess quality of replans less clear, especially in daily adaptive radiotherapy this is likely acceptable. We took the approach we use clinically when assessing the need for replan, where essentially CTV coverage needs to be adequate. This is believed to be an appropriate approach for this work, as in daily adaptive radiotherapy the CTV coverage is prioritised. Arguably, in an offline approach the PTV (potentially with revised margins) could be the focus. In our results, and despite assessing only CTV coverage, the loss of PTV coverage was minimal at 0.2 Gy.

The use of absolute dose differences is a potential limitation of this work, as it may be specific to the dose prescription and distributions used in this study. However, radical head and neck prescriptions are usually similar (typically ranging 60 Gy in 30# compared to 70 Gy in 35# [39]) so it is likely that a simple scaling of the doses calculated here may be applicable, but further study is needed.

This work is of relevance to an adaptive radiotherapy workflow. The online adaptive radiotherapy solutions currently commercially available (MR linacs and Varian Ethos [9,40] systems) both require recontouring at the treatment machine, with it being reported as taking up to 26 min to outline and replan on a treatment MR scan [41], with both workflow and potential changes to anatomy implications. This work is of particular relevance to Ethos, as Ethos uses CBCT images for replanning. Using the work presented here, it would be feasible to propagate organ at risk contours to the treatment imaging, replan using the propagated contours and if the dose comes to within the suggested ranges given, put the organ under closer scrutiny as a clinical tolerance may be exceeded. Alternatively, using the dose ranges presented here, if the pCT DVH statistic is within the stated ranges (e.g. 3.7 Gy for the brainstem) of the tolerance dose, it is likely that the dose will still be within the range after replanning so the contour should be reviewed, but this requires further work. It may subsequently be possible to develop a "traffic light" metric for each organ based on how close the organ is to tolerance on the pCT, but this again would require further work.

Conclusion

It has been shown that using non-rigid registration and propagation of OAR contours from the pCT to CBCTs spaced throughout treatment, it is possible to estimate the dose to the gold and silver standard OARs to within 3.7 Gy. This has the potential to create a more efficient adaptive radiotherapy workflow, and the clinician or radiographer can be targeted over which OAR contours they manually adjust.

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 Cancer Research UK via funding to the Cancer Research Manchester Centre (C147/A25254) and the NIHR Manchester Biomedical Research Centre. Thanks need to go to Dr Shagun Juneja for assistance with the contouring in this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejmp.2022.09.002.

References

[1] Strojan P, Hutcheson KA, Eisbruch A, Beitler JJ, Langendijk JA, Lee AWM, et al. Treatment of late sequelae after radiotherapy for head and neck cancer. Cancer Treat Rev 2017;59:79–92.
[2] Nutting CM, Morden JP, Harrington KJ, Urbano TG, Bhide SA, Clark C, et al. Parotid-sparing intensity modulated versus conventional radiotherapy in head and neck cancer (PARSPORT): a phase 3 multicentre randomised controlled trial. Lancet Oncol 2011;12(2):127–36.
[3] Sonke JJ, Aznar M, Rasch C. Adaptive radiotherapy for anatomical changes. Semin Radiat Oncol 2019;29:245–57. https://doi.org/10.1016/j.semradonc.2019.02.007.
[4] Bhide SA, Davies M, Burke K, McNair HA, Hansen V, Barbachano Y, et al. Weekly volume and dosimetric changes during chemoradiotherapy with intensity-modulated radiation therapy for head and neck cancer: a prospective observational study. Int J Radiat Oncol Biol Phys 2010;76(5):1360–8.
[5] Castadot P, Lee JA, Geets X, Grégoire V. Adaptive radiotherapy of head and neck cancer. Semin Radiat Oncol 2010;20:84–93. https://doi.org/10.1016/j.semradonc.2009.11.002.
[6] Grégoire V, Jeraj R, Lee JA, O'Sullivan B. Radiotherapy for head and neck tumours in 2012 and beyond: Conformal, tailored, and adaptive? Lancet Oncol 2012;13:e292–300. https://doi.org/10.1016/S1470-2045(12)70237-1.
[7] Vásquez Osorio EM, Hoogeman MS, Bondar L, Levendag PC, Heijmen BJM. A novel flexible framework with automatic feature correspondence optimization for nonrigid registration in radiotherapy. Med Phys 2009;36:2848–59. https://doi.org/10.1118/1.3134242.
[8] Vickress JR, Battista J, Barnett R, Yartsev S. Online daily assessment of dose change in head and neck radiotherapy without dose-recalculation. J Appl Clin Med Phys 2018;19:659–65. https://doi.org/10.1002/acm2.12432.
[9] Moazzezi M, Rose B, Kisling K, Moore KL, Ray X. Prospects for daily online adaptive radiotherapy via ethos for prostate cancer patients without nodal involvement using unedited CBCT auto-segmentation. J Appl Clin Med Phys 2021;22:82–93. https://doi.org/10.1002/acm2.13399.
[10] Klüter S. Technical design and concept of a 0.35 T MR-Linac. Clin Transl Radiat Oncol 2019;18:98–101. https://doi.org/10.1016/j.ctro.2019.04.007.
[11] Thor M, Petersen JBB, Bentzen L, Høyer M, Paul ML. Deformable image registration for contour propagation from CT to cone-beam CT scans in radiotherapy of prostate cancer. Acta Oncol (Madr) 2011;50:918–25. https://doi.org/10.3109/0284186X.2011.577806.
[12] Thörnqvist S, Petersen JBBB, Høyer M, Bentzen LN, Paul ML. Propagation of target and organ at risk contours in radiotherapy of prostate cancer using deformable image registration. Acta Oncol (Madr) 2010;49:1023–32. https://doi.org/10.3109/0284186X.2010.503662.

[13] Riegel AC, Antone JG, Zhang H, Jain P, Raince J, Rea A, et al. Deformable image registration and interobserver variation in contour propagation for radiation therapy planning. J Appl Clin Med Phys 2016;17(3):347–57.
[14] Fabri D, Zambrano V, Bhatia A, Furtado H, Bergmann H, Stock M, et al. A quantitative comparison of the performance of three deformable registration algorithms in radiotherapy. Z Med Phys 2013;23(4):279–90.
[15] Faggiano E, Fiorino C, Scalco E, Broggi S, Cattaneo M, Maggiulli E, et al. An automatic contour propagation method to follow parotid gland deformation during head-and-neck cancer tomotherapy. Phys Med Biol 2011;56(3):775–91.
[16] Beasley WJ, McWilliam A, Aitkenhead A, Mackay RI, Rowbottom CG. The suitability of common metrics for assessing parotid and larynx autosegmentation accuracy. J Appl Clin Med Phys 2016;17:41–9. https://doi.org/10.1120/jacmp.v17i2.5889.
[17] Hvid CA, Elstrøm UV, Jensen K, Grau C. Cone-beam computed tomography (CBCT) for adaptive image guided head and neck radiation therapy. Acta Oncol (Madr) 2018;57:552–6. https://doi.org/10.1080/0284186X.2017.1398414.
[18] Tsuji SY, Hwang A, Weinberg V, Yom SS, Quivey JM, Xia P. Dosimetric evaluation of automatic segmentation for adaptive IMRT for head-and-neck cancer. Int J Radiat Oncol Biol Phys 2010;77:707–14. https://doi.org/10.1016/j.ijrobp.2009.06.012.
[19] Paganelli C, Meschini G, Molinelli S, Riboldi M, Baroni G. Patient-specific validation of deformable image registration in radiation therapy: overview and caveats. Med Phys 2018;45:e908–22. https://doi.org/10.1002/mp.13162.
[20] Kumarasiri A, Siddiqui F, Liu C, Yechieli R, Shah M, Pradhan D, et al. Deformable image registration based automatic CT-to-CT contour propagation for head and neck adaptive radiotherapy in the routine clinical setting. Med Phys 2014;41(12):121712.
[21] Teguh DN, Levendag PC, Voet PWJ, Al-Mamgani A, Han X, Wolf TK, et al. Clinical validation of atlas-based auto-segmentation of multiple target volumes and normal tissue (swallowing/mastication) structures in the head and neck. Int J Radiat Oncol Biol Phys 2011;81(4):950–7.
[22] Li X, Zhang YY, Shi YH, Zhou LH, Zhen X. Evaluation of deformable image registration for contour propagation between CT and cone-beam CT images in adaptive head and neck radiotherapy. Technol Heal Care 2016;24:S747–55. https://doi.org/10.3233/THC-161204.
[23] Nash D, Juneja S, McWilliam A, Palmer AL, Vasquez OE. Dosimetric and geometric evaluation of five commercial contour propagation tools for online adaptive radiotherapy. Med Phys 2020;47:e606.
[24] Nash D, Juneja S, Palmer AL, van Herk M, McWilliam A, Vasquez OE. The geometric and dosimetric effect of algorithm choice on propagated contours from CT to CBCTs. Phys Med 2022;100:112–9. https://doi.org/10.1016/s0167-8140(21)01745-x.
[25] Marchant TE, Moore CJ, Rowbottom CG, Mackay RI, Williams PC. Shading correction algorithm for improvement of cone-beam CT images in radiotherapy. Phys Med Biol 2008;53:5719–33. https://doi.org/10.1088/0031-9155/53/20/010.
[26] Yang C-C, Chen F-L, Lo Y-C, Zhang Q. Improving image quality of on-board cone-beam CT in radiation therapy using image information provided by planning multi-detector CT: a phantom study. PLoS ONE 2016;11(6):e0157072.
[27] Warfield SK, Zou KH, Wells WM. Simultaneous truth and performance level estimation (STAPLE): An algorithm for the validation of image segmentation. IEEE Trans Med Imaging 2004;23:903–21. https://doi.org/10.1109/TMI.2004.828354.
[28] Filannino M, Di Bari M. Gold standard vs. silver standard: the case of dependency parsing for Italian. Proc Second Ital Conf Comput Linguist CLiC-it 2015;2015 (2016):3–4. https://doi.org/10.4000/books.aaccademia.1475.
[29] Hahn U, Tomanek K, Beisswanger E, Faessler E. A proposal for a configurable silver standard. ACL 2010 - LAW 2010 4th Linguist Annot Work Proc 2010:235–42.
[30] Hamming-Vrieze O, van Kranen SR, Heemsbergen WD, Lange CAH, van den Brekel MWM, Verheij M, et al. Analysis of GTV reduction during radiotherapy for oropharyngeal cancer: Implications for adaptive radiotherapy. Radiother Oncol 2017;122(2):224–8.
[31] Brock KK, Mutic S, McNutt TR, Li H, Kessler ML. Use of image registration and fusion algorithms and techniques in radiotherapy: Report of the AAPM Radiation Therapy Committee Task Group No. 132. Med Phys 2017;44:e43–76. 10.1002/mp.12256.
[32] Dice LR. Measures of the amount of ecologic association between species. Ecology 1945;26:297–302. https://doi.org/10.2307/1932409.
[33] Ayyalusamy A, Vellaiyan S, Shanmugam S, Ilamurugu A, Gandhi A, Shanmugam T, et al. Feasibility of offline head & neck adaptive radiotherapy using deformed planning CT electron density mapping on weekly cone beam computed tomography. Br J Radiol 2017;90:20160420. https://doi.org/10.1259/bjr.20160420.
[34] Beaver MES, Matheny KE, Roberts DB, Myers JN. Predictors of weight loss during radiation therapy. Otolaryngol - Head Neck Surg 2001;125:645–8. https://doi.org/10.1067/mhn.2001.120428.
[35] Veiga C, Lourenço AM, Mouinuddin S, van Herk M, Modat M, Ourselin S, et al. Toward adaptive radiotherapy for head and neck patients: Uncertainties in dose warping due to the choice of deformable registration algorithm. Med Phys 2015;42:760–9. https://doi.org/10.1118/1.4905050.
[36] Veiga C, McClelland J, Moinuddin S, Lourenço A, Ricketts K, Annkah J, et al. Toward adaptive radiotherapy for head and neck patients: Feasibility study on using CT-to-CBCT deformable registration for “dose of the day” calculations. Med Phys 2014;41:031703. https://doi.org/10.1118/1.4864240.
[37] Wang M, Zhang Q, Lam S, Cai J, Yang R. A review on application of deep learning algorithms in external beam radiotherapy automated treatment planning. Front Oncol 2020;10. https://doi.org/10.3389/fonc.2020.580919.
[38] van der Veen J, Gulyban A, Willems S, Maes F, Nuyts S. Interobserver variability in organ at risk delineation in head and neck cancer. Radiat Oncol 2021;16:1–11. https://doi.org/10.1186/s13014-020-01677-2.
[39] The Royal College of Radiologists. Radiotherapy dose fractionation, third edition. London: 2019. 10.4997/JRCPE.2010.423.
[40] Archambault Y, Boylan C, Bullock D, Morgas T, Peltola J, Ruokokoski E, et al. Making on-line adaptive radiotherapy possible using artificial intelligence and machine learning for efficient daily re-planning. Med Phys Int J 2020;8:77–86.
[41] Acharya S, Fischer-Valuck BW, Kashani R, Parikh P, Yang D, Zhao T, et al. Online magnetic resonance image guided adaptive radiation therapy: first clinical applications. Int J Radiat Oncol Biol Phys 2016;94:394–403. https://doi.org/10.1016/j.ijrobp.2015.10.015.