Physics and Imaging in Radiation Oncology 9 (2019) 21–27

Contents lists available at ScienceDirect

Physics and Imaging in Radiation Oncology

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

Original Research Article

Monitoring anatomical changes of individual patients using statistical process control during head-and-neck radiotherapy

Nicholas J. Lowther[a,b], David A. Hamilton[a], Han Kim[a], Jamie M. Evans[a], Steven H. Marsh[b], Robert J.W. Louwe[a,*]

[a] Wellington Blood and Cancer Centre, Department of Radiation Oncology, Wellington, New Zealand
[b] University of Canterbury, School of Physical and Chemical Sciences, Christchurch, New Zealand

ARTICLE INFO

Keywords:
Anatomical changes
Head-and-neck
Radiotherapy
Statistical process control
Deformable image registration

ABSTRACT

Background and purpose: Reduced toxicity while maintaining loco-regional control rates have been reported after reducing planning target volume (PTV) margins for head-and-neck radiotherapy (HNRT). In this context, quantifying anatomical changes to monitor patient treatment is preferred. This retrospective feasibility study investigated the application of deformable image registration (DIR) and Exponentially Weighted Moving Average (EWMA) Statistical Process Control (SPC) charts for this purpose.

Materials and methods: DIR between the computed tomography for treatment planning (pCT) images of twelve patients and their daily on-treatment cone beam computed tomography (CBCT) images quantified anatomical changes during treatment. EWMA charts investigated corresponding trends. Uncertainty analysis provided 90% confidence limits which were used to confirm whether a trend previously breached a threshold.

Results: Trends in patient positioning reproducibility occurred before the end of treatment week four in 54% of cases. Using SPC process limits, only 24% of these were confirmed at a 90% confidence level before the end of treatment. Using an *a priori* clinical limit of 2 mm, absolute changes in patient pose were detected in 39% of cases, of which 82% were confirmed. Soft tissue trends outside SPC process limits occurring before the end of treatment week four were confirmed in 90% of cases.

Conclusion: Structure specific action thresholds enabled detection of systematic anatomical changes during the first four weeks of treatment. Investigation of the dosimetric impact of the observed deviations is needed to show the efficacy of SPC to timely indicate required treatment adaptation and provide a safety net for PTV margin reduction.

1. Introduction

The introduction of intensity modulated radiotherapy has enabled highly conformal dose deliveries which allows dose reduction to organs at risk (OAR) and reduced treatment toxicity [1–3]. These highly conformal treatments require image-guided radiotherapy (IGRT) to warrant accurate patient positioning and monitoring of changes in patient anatomy [4]. Planning target volume (PTV) margins of 5 mm are commonly applied to establish target coverage [5,6]. PTV margins and IGRT do not standardly account for non-rigid anatomy changes which are commonly observed in head-and-neck radiotherapy (HNRT) [7]. Nevertheless, two retrospective studies have suggested that 3 mm planning target volume (PTV) margins are sufficient if daily IGRT is used to correct for rigid patient position variations [8,9]. The group of Chen et al. has also clinically implemented reduced PTV margins from 5 mm to 3 mm and reported a reduction in late toxicity side-effects while maintaining equivalent two- and three-year loco-regional control rates [9–11]. However, these results may have been confounded by the application of adaptive radiotherapy (ART) for selected patients [12]. A recent study by Navran et al. also reported on favorable toxicity profiles for reduced 3 mm PTV margins while maintaining good tumor control rates [13]. The latter study applied an IGRT protocol that accounted for non-rigid patient positioning deviations as well as ART for selected patients. As we considered whether it was possible to implement reduced 3 mm PTV margins after achieving a large improvement in patient positioning in our department [14], it was recognized that many other aspects of HNRT including accuracy of target delineation, robustness of the planning solution for anatomical changes influence treatment outcomes [15] and should all be considered when the PTV margins are reduced. In addition, objective and generally applicable

* Corresponding author.
E-mail address: rob.louwe@ccdhb.org.nz (R.J.W. Louwe).

https://doi.org/10.1016/j.phro.2018.12.004
Received 19 October 2018; Received in revised form 27 November 2018; Accepted 7 December 2018

guidelines to select patients who will benefit from treatment adaptation are lacking [4,16], and most studies on ART apply subjective criteria to select patients for treatment adaptation [12,17–19]. In an effort to create a safety net for patients who exhibit large non-rigid deformations and objectively select patients who would benefit from treatment adaptation in the context of PTV margin reduction, this feasibility study investigated the first step to build such a framework. Specifically, the suitability of deformable image registration (DIR) and exponentially weighted moving average (EWMA) statistical process control (SPC) charts to objectively quantify and monitor individual patients' deformation, i.e., non-rigid changes in both pose and anatomy of the patient during treatment were investigated.

2. Materials and methods

2.1. Patient group

This retrospective study included twelve patients with cancers of the head-and-neck (HN) region that were treated radically. Comprehensive information regarding the patient cohort can be found in Supplementary Data A. The patients included in this retrospective study provided written consent to use their data in clinical audits.

2.2. pCT contouring

Nine bony anatomy (BA) structures were delineated on each patient's computed tomography for treatment planning (pCT) scan: C1-C3, C3-C5, C5-C7, mandible, maxilla, base of skull, hyoid, occipital and larynx. A pre-defined neck volume V^{neck} was defined as all tissue within the volume bounded by axial planes at the base of C3 at the anterior cortical boundary and base of C4 at the anterior cortical boundary (Fig. 1). The original parotid gland (PG) contours were reviewed by a radiation oncologist and corrected where necessary to minimize the uncertainty of the DIR results.

2.3. Deformable image registration

The pCT of each patient was deformed to match the anatomy of each daily cone beam computed tomography (CBCT) using SmartAdapt v.13 (Varian Medical Systems, Palo Alto CA, USA), resulting in 30 deformed CTs (dCT). The deformed structures of the dCT as well as the non-deformed structures of the pCT were propagated to the corresponding CBCTs. These data sets were exported in Digital Imaging and Communications in Medicine (DICOM) format for subsequent analysis with in-house developed software (MATLAB 2016b version 9.1, The MathWorks Inc.).

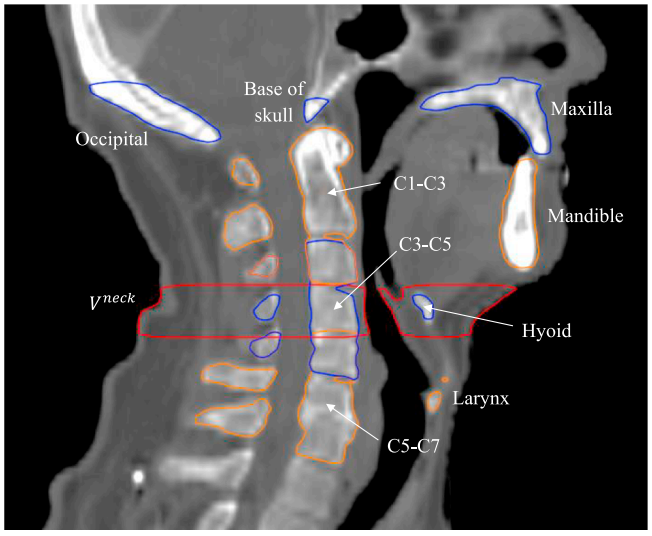

Fig. 1. Definition of bony anatomy (BA) structures and neck volume V^{neck} in the central sagittal plane. The superior end of the occipital bone contour was defined as to just include the external occipital protuberance.

2.4. Quantifying non-rigid deformations

The exported contours were used to quantify the non-rigid deformations: BA positioning reproducibility, centroid shifts of the PGs, volume changes of the PGs (V^{PG}) and V^{neck} changes. The BA positioning reproducibility was quantified by first calculating the change T in centroid position C of all BA structures with respect to its pCT position for each direction k = x, y, z, and each fraction f:

$$T_{f,k}^{BA} = C_{f,k,dCT}^{BA} - C_{f,k,pCT}^{BA} \quad (1)$$

Subsequently, the 3D-deviations of all BA structures were calculated relative to the reference structure C1-C3, using:

$$\partial_f^{BA} = \sqrt{(T_{x,f}^{BA} - T_{x,f}^{C1-C3})^2 + (T_{y,f}^{BA} - T_{y,f}^{C1-C3})^2 + (T_{z,f}^{BA} - T_{z,f}^{C1-C3})^2} \quad (2)$$

Similarly, the 3D-shifts of the PGs were calculated relative to C1-C3 according to Eq. (2).

To assist with readability, the 3D-deviation of BA structure X relative to C1-C3 will be referred to as "deviation of X" from here onwards.

2.5. Statistical process control

The trends of the metrics ∂_f^{BA}, ∂_f^{PG}, V_f^{neck}, V_f^{PG} over time were monitored using EWMA charts [20]. The values E_f of the EWMA statistic were calculated using:

$$E_f = \lambda X_f + (1 - \lambda) E_{f-1} \quad (3)$$

where λ is a constant defined as $0 < \lambda \leq 1$ that determines the depth of memory (smoothing) of the EWMA and X_f is the metric of interest for fraction f. As shown in Fig. 2, E_0 is the average metric of interest during the EWMA reference period. The first week of treatment (i.e. fractions 1–5) is used as the EWMA reference period, assuming that prospective monitoring of EWMA statistics would be conducted from fraction six onwards in a clinical scenario, with EWMA statistics for fractions 1–5 analyzed retrospectively.

The lower and upper process limits LPL and UPL, which indicate a statistically significant difference from the reference period, were calculated using:

$$\text{LPL, UPL} = \mu_0 \pm L\sigma\sqrt{\left(\frac{\lambda}{2-\lambda}\right)[1-(1-\lambda)^{2f}]} \quad (4)$$

The initial deviation μ_0 and variation σ were calculated as the mean and SD for the reference period, and L is a factor determining the width of the process limits. Considering that Q-Q plots indicated the metrics of interest were not always normally distributed, λ and L were set to $\lambda = 0.05$ and $L = 2.492$ to obtain similar type I and type II error probabilities as for normally distributed data [21].

PGs were grouped according to mean dose at planning: high- ($\leq$26 Gy, n = 7) and low-dose ($>$ 26 Gy, n = 17) consistent with the different impact of radiation dose predicted below or above this cut-off point by normal tissue complication probability (NTCP) models [22,23]. In cases where the superior or inferior extents of the CBCT did not contain a specific BA structure, that structure was excluded from the analysis for that patient.

The fraction numbers at which the EWMA statistics went outside the corresponding process limits were recorded. Where possible, EWMA trends that were obtained using DIR were compared with results obtained using an alternative manual method. ∂_f^{BA} results obtained using DIR were compared with results obtained using a manual method of registration (MMR) as described by various authors [14,24]. The MMR

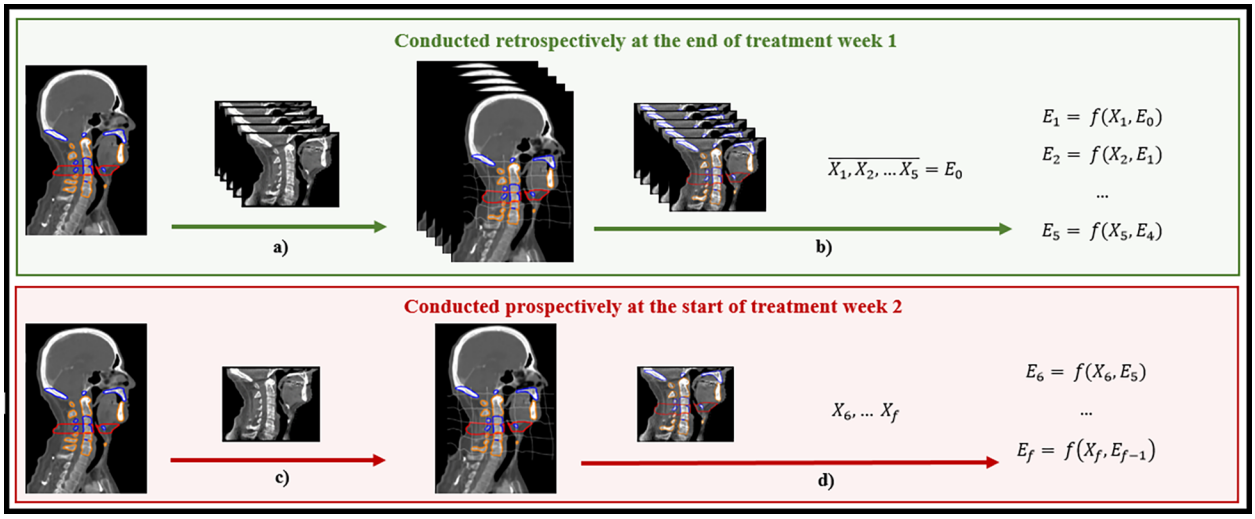

Fig. 2. Workflow detailing the exponentially weighted moving average (EWMA) statistics monitoring procedure based on DIR. (a) The planning CT (pCT) is deformed to match the anatomical configurations of each cone beam CT (CBCT) of treatment week 1; (b) Deformed pCT contours propagated to the CBCTs are used to calculate metrics of interest X_f for fractions 1–5. The average metric of interest is defined as E_0 and used to calculate the EWMA statistics $E_1 \ldots E_5$ as per Eq. (3); (c-d) At the beginning of treatment week two, propagation of deformed contours and EWMA calculations are conducted on a fraction-by-fraction basis.

used rigid image registrations (translations only; no rotations) of the individual structures to assess the shift of the structure centroids relative to the pCT. DIR results for V_f^{neck} trends were compared with the change of a single axial slice volume at the base of C2 that was manually contoured for all CBCTs as well as changes in patient weight. Considering that a change in patient weight may occur at any location of the patient's body, V_f^{neck} trends were expected to describe volume change of the treatment region more appropriately than a simple weight metric.

2.6. Detection of trends

In this study two methods to detect trends were investigated:

2.6.1. SPC limits for trend detection

A detailed uncertainty analysis was carried out to assess whether DIR and EWMA charts can accurately monitor non-rigid deformation trends of individual patients, and detect changes in patient positioning or anatomy in a timely manner. In these analyses including both the DIR and manual results, the uncertainty was calculated as per the Guide to the Expression of Uncertainty in Measurement (GUM) [25] (see Supplementary data B). The results of the uncertainty analyses were used to quantify the accuracy of trends as the average 90% level of confidence interval of the EWMA statistics for each metric over all treatment fractions. The sensitivity to detect a trend (i.e., the minimal detectable trend magnitude) with 90% confidence with respect to the reference period (i.e., E_0 set to zero) was defined as the difference between the 90% level of confidence of the upper EWMA process limit and the EWMA center line at the last treatment fraction. Differences between the sensitivity and accuracy of the two methods over all patients were assessed using paired *t*-tests at the 0.05 level of significance. The robustness of trend detection using SPC limits was quantified as the proportion of cases where the accuracy interval of the trend exceeded the 90% level of confidence of the process limits before the end of treatment relative to the number of cases where the EWMA statistic itself exceeded a process limit before the end of the fourth week of treatment.

2.6.2. Clinical limit for trend detection

Preliminary SPC trend results highlighted sub-mm process limits for specific structures which would likely not provide an efficient threshold for clinically relevant process changes. Also the exact moment that a BA deviation trend exceeded the control limits within a certain statistical confidence level was not easily defined, in particular for shallow trends in combination with narrow process limits of a structure. The EWMA trend analysis was therefore also carried out using an *a priori* 2 mm clinical limit as detection threshold for *absolute* BA deviations. The robustness of trend detection using a 2 mm clinical limit (including the observed E_0) was defined as the proportion of cases where the accuracy interval of the EWMA statistic exceeded this limit relative to the number of cases where the EWMA statistic itself exceeded this clinical limit before the end of the fourth week of treatment.

3. Results

3.1. SPC trends

Overall, 31% of the trends describing BA positioning reproducibility acquired with DIR remained within the EWMA process limits during treatment, while 54% and 15% of the trends exceeded a process limit before and after the end of the fourth week of treatment (i.e., fraction 20), respectively (Supplementary Data C). Fig. 3a shows the deviation of the mandible for patient three exceeding the upper DIR process limit at fraction 18.

3.1.1. SPC limits for trend detection

Both the average accuracy and sensitivity calculated over all patients were smaller than 1 mm for BA deviations. Table 1 summarizes the average accuracy and sensitivity recorded for the various structures and compares the results obtained for DIR and MMR. Except for the deviation of the occipital bone calculated with DIR, the average accuracies were smaller than the corresponding average sensitivities indicating that the minimal detectable trend with a 90% level of confidence was generally equal to the observed sensitivity. Statistically significant differences between DIR and MMR accuracies were observed for all BA structures except deviation of the maxilla. Conversely, only the sensitivity to detect a C5-C7 deviation was significantly different between the two methods.

The robustness to detect a statistically significant change relative to the reference period for each trend is summarized in Table 2 (individual patient results available in Supplementary Data D). For BA deviations, the overall robustness of trend detection was 24% (95% CI: 11–38%), with none of the potential trends for occipital, base of skull, maxilla and larynx deviations confirmed to be an actual trend at a 90% level of

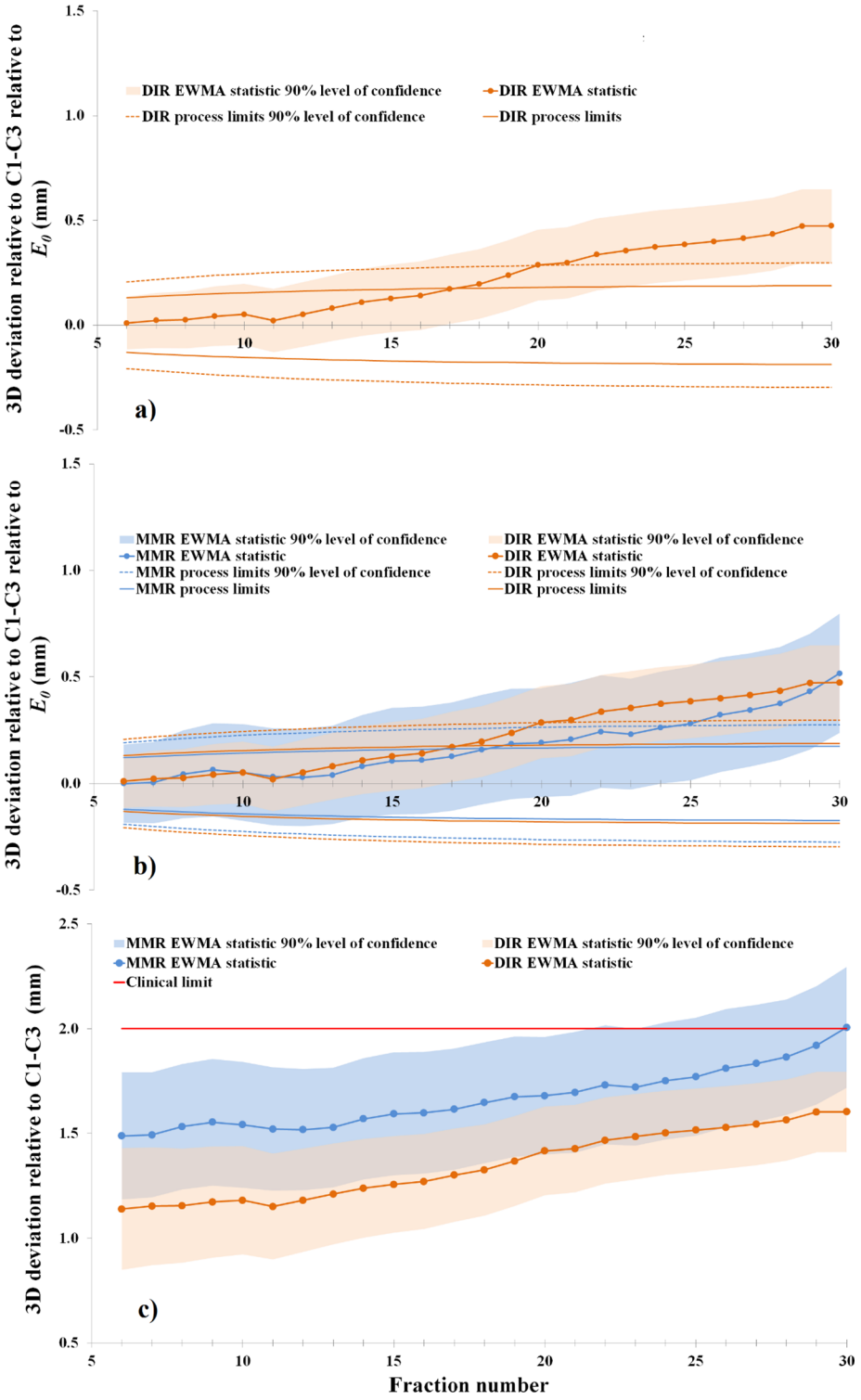

Fig. 3. Example of the exponentially weighted moving average (EWMA) charts used in this study of the mandible deviation for patient three. EWMA trends derived from deformable image registration (DIR) or the manual method of registration (MMR), are shown in orange and blue, respectively. Trends are plotted either as deviations relative to the SPC reference period (panels a and b) or absolute deviations including E_0 (panel c). (a) DIR derived EWMA chart including uncertainty analysis and corrected for initial deformation E_0; (b) comparison of DIR and MMR derived EWMA charts; (c) DIR and MMR derived EWMA charts including observed E_0 for the purpose of comparing the trends with a 2 mm clinical limit.

confidence. In contrast, the robustness of trend detection for soft tissue trends was 90% (95% CI: 80–99%), due to steeper gradients of these trends.

3.1.2. Clinical limit for trend detection

The *absolute* BA deviations exceeded the 2 mm clinical limit in 33/84 (39%) of the available cases (Table 3) (individual patient results available in Supplementary Data E). Fig. 3c provides an example of this analysis where the deviation of the mandible for patient 3 acquired using DIR did not exceed the 2 mm clinical limit during the treatment period. In the majority of cases (29/33) these deviations were already larger than 2 mm from the start of treatment in particular for the hyoid and larynx, indicative of the high mobility of these structures (see also Supplementary Data F). The overall robustness of trend detection for BA deviations larger than 2 mm was 82% (95% CI: 67–96%).

3.1.3. Trends returning to control

EWMA trends for BA deviations did not move back within either the process limits or clinical limits in any of the cases after previously exceeding control at the 90% level of confidence. In 2–3% of cases a prior

Table 1
Average sensitivity S and accuracy A of exponentially weighted moving average (EWMA) charts established from deformable image registration (DIR) and manually acquired raw data (MMR). Δ represents the difference between DIR and MMR results. Accuracy A = average 90% level of confidence interval over all treatment fractions. Sensitivity S = half the EWMA process limit 90% level of confidence at the last treatment fraction. *Indicates a statistically significant difference.

	DIR		MMR		Δ	
Structure	*A*	*S*	*A*	*S*	*A*	*S*
Occipital bone [mm]	0.3	0.3	0.2	0.3	0.1*	0.0
Base of skull [mm]	0.2	0.3	0.2	0.3	0.0*	0.0
Maxilla [mm]	0.2	0.3	0.2	0.4	0.0	−0.1
C3-C5 [mm]	0.1	0.3	0.2	0.3	−0.1*	0.0
C5-C7 [mm]	0.1	0.4	0.2	0.5	−0.1*	−0.1*
Hyoid [mm]	0.2	0.7	0.3	0.7	−0.1*	0.0
Larynx [mm]	0.2	0.8	0.3	0.8	−0.1*	0.0
Mandible [mm]	0.2	0.4	0.2	0.4	0.0*	0.0
High dose PG dev. [mm]	0.1	0.3	–	–	–	–
Low dose PG dev. [mm]	0.1	0.3	–	–	–	–
High dose PG vol. [cm^3]	0.3	0.5	–	–	–	–
Low dose PG vol. [cm^3]	0.3	0.5	–	–	–	–
V_f^{neck} [cm^3]	2.0	3.9	–	–	–	–

confirmation of a trend was negated as the lower boundary of the trend's accuracy interval moved back either within the upper process limit's 90% level of confidence or the clinical limit.

3.2. Soft tissue analysis

For eight patients (73%), the trend describing a volume change of V_f^{neck} exceeded a process limit before the end of the fourth week of treatment, and for one patient (9%), this occurred during the final two weeks of treatment. These trends usually represented a decrease of the neck volume, however for two of these nine patients, an increase in V_f^{neck} was observed.

A strong correlation ($r^2 = 0.87$; $p < 0.01$) was found between the changes in V_f^{neck} and patient weight. A slightly lower correlation ($r^2 = 0.76$; $p < 0.01$) was found between the changes in V_f^{neck} and those of the single axial slice volume at the base of C2. The 90% confidence intervals of the EWMA trends of the latter two metrics overlapped in 91% of cases.

Ninety-four percent of the 48 trends describing shift or shrinkage of a PG exceeded a process limit during HNRT treatment. EWMA trends describing shrinkage of high and low mean dose PG groups exceeded a process limit before the end of the fourth week of treatment in 94% and 86% of the cases, respectively. A decrease in volume was observed for all parotid glands. EWMA trends describing PG shift exceeded a process limit before the end of the fourth week of treatment in 76% and 71% of the cases, respectively. Supplementary Data G shows EWMA trends of average volume and shift for both PG groups.

The average volume change of the PGs from planning to end of treatment was not significantly different (p = 0.20) for the high dose PGs (-5.9 cm^3; range −3.3 to −10.0 cm^3) and the low dose PGs (−4.8 cm^3; range: −3.5 to −9.4 cm^3). Similarly, the average shift during treatment was not significantly different (p = 0.38) for the high dose PGs (2.1 mm; range 1.1–3.1 mm) and the low dose PGs (1.1 mm; range 1.1–3.3 mm).

4. Discussion

This study investigated the suitability of DIR and EWMA SPC to quantify and monitor individual patients' changes in pose and anatomy during HNRT. This method facilitates a standardized approach to quantify patient deformations with an estimated robustness derived from comprehensive uncertainty analyses. It is a first step in developing a safety net for PTV margin reduction in HNRT as well as objective guidelines to select patient for treatment adaptation. Considering that loss in target coverage during treatment might occur more often and might become more relevant with reduced mm PTV margins, EWMA-facilitated detection of non-rigid changes in pose and anatomy during treatment of an individual patient could allow for early treatment adaption.

Previous studies investigated the application of SPC charts to monitor the reproducibility of patient positioning based on MMR [14]. However, MMR is very labour and time intensive; hampering clinical implementation. The application of DIR and SPC to quantify and monitor individual patient deformations during treatment can largely be automated and is therefore attractive from an economic and efficiency perspective. Considering that a change in treatment is usually not feasible in the final weeks of treatment due to the time required for re-scanning the patient, contouring and re-planning, this study assumed that clinically relevant changes in pose and anatomy should be detected before the end of treatment week four.

Over all patients, DIR EWMA trends exceeded a process limit before the end of treatment week 4 in 54%, 73% and 83% of cases for BA deviations, volume changes of V_f^{neck} and PG changes, respectively (Supplementary Data C). However, these process limits represent the boundaries of the variation expected for a 'process' based on the observed variation during the reference period. The SPC limits therefore ignore the offset E_0 during the reference period relative to the desired patient position as defined during CT-planning which can be several mm's in case of highly mobile structures. Therefore, deformation trends exceeding an SPC process limit itself is not indicative for the clinical relevance of an observed BA deviation and would require additional interpretation rules. In addition, potential trends occurring during the reference period add to the uncertainty of the SPC limits. Furthermore, the exact moment that the trend of a BA structure deviation relative to the initial value during the reference period exceeded an SPC process limit was not easily defined due to the shallow trends that were observed, in spite of the sub-mm accuracies of the EWMA statistic and process limits. The latter problem is exemplified in Fig. 3a where the EWMA statistic exceeds the upper process limit at fraction 18, but the 90% level of confidence interval of the EWMA statistic does not cross

Table 2
Median and range of the fraction number where the exponentially weighted moving average (EWMA) statistic exceeded an SPC process limit before the end of treatment week 4, and the proportion of cases where the existence of a trend was confirmed by the accuracy interval of the trend exceeding the process limits 90% level of confidence before the end of treatment. This proportion is also expressed as percentage representing the robustness of the trend detection with the corresponding 95% confidence interval. EWMA parameters were established from deformable image registration (DIR) raw data. (PG = parotid gland; BOS = base of skull; Occ = occipital; Max = maxilla; Mand = mandible).

	∂_f^{Occ}	∂_f^{BOS}	∂_f^{Max}	∂_f^{C3-C5}	∂_f^{C5-C7}	∂_f^{Hyoid}	∂_f^{Larynx}	∂_f^{Mand}	∂_f^{PGhigh}	∂_f^{PGlow}	V_f^{PGhigh}	V_f^{PGlow}	V_f^{neck}
Median	13	15	11	14	14	14	19	15	14	9	10	12	13
Range	12–19	9–20	8–15	10–18	12–19	11–17	14–20	8–18	9–20	7–18	8–20	9–20	9–18
Proportion confirmed	0/4	0/6	0/5	3/7	3/5	2/7	0/4	3/7	12/13	4/5	14/16	6/6	7/8
Robustness	0%	0%	0%	43%	60%	29%	0%	43%	92%	80%	88%	100%	88%
95% CI	0–13%	0–8%	0–10%	0–87%	7–100%	0–69%	0–13%	0% −87%	74–100%	35–100%	68–100%	92–100%	58–100%

Table 3
Median and range of the fraction number where the exponentially weighted moving average (EWMA) statistic exceeded a 2 mm clinical limit before the end of treatment week 4, and the proportion of cases where the existence of a trend was confirmed by the accuracy interval of the trend exceeding the clinical limit before the end of treatment. This proportion is also expressed as percentage representing the robustness of the trend detection with the corresponding 95% confidence interval. EWMA parameters were established from deformable image registration (DIR) raw data. (BOS = base of skull; Occ = occipital; Max = maxilla; Mand = mandible).

	∂_f^{Occ}	∂_f^{BOS}	∂_f^{Max}	∂_f^{C3-C5}	∂_f^{C5-C7}	∂_f^{Hyoid}	∂_f^{Larynx}	∂_f^{Mand}
Median	1	1	6	NA	1	1	1	1
Range	1–1	1	1–11	NA	1–19	1	1	1–10
Proportion confirmed	3/4	0/1	1/2	0/0	3/3	9/9	8/9	3/5
Robustness	75%	0%	50%	NA	100%	100%	89%	60%
95% CI	20–100%	0–50%	0–100%	NA	83–100%	94–100%	63–100%	7–100%

the 90% level of confidence of the upper process limit during the treatment course. Similar behavior was observed for MMR derived EWMA charts (Fig. 3b). At the 90% level of confidence, BA deviation trends where the EWMA statistic exceeded the process limits during the first four weeks of treatment were only confirmed in 24% of the cases. For these reasons, it is more efficient to apply a clinical limit to detect BA deviations that could potentially have a clinical impact during treatment. This study showed that EWMA statistics representing BA deviation trends exceeding a generic *a priori* 2 mm clinical limit before the end of treatment week four were detected in 82% of cases at the 90% level of confidence. However, this clinical limit should be individually set for each BA structures' deviations based on expected impact to target coverage and/or OAR sparing, which is subject to future investigations. Alternatively, the robustness of SPC limits can possibly be improved by also considering the rate of the changes during the treatment to define SPC control limits.

In contrast to the observations for BA deviations, SPC process limits were useful to detect soft tissue trends that exceeded the process limits before the end of treatment week four which could be confirmed in 90% of cases.

DIR precision is important for correct anatomical mapping of the pCT to daily CBCTs. In the absence of a golden standard, trends acquired using DIR were compared with results obtained using an alternative manual method where possible in combination with an uncertainty analyses were conducted. Overall at the 90% level of confidence, BA deviation trends acquired with DIR and MMR overlapped in 77% of cases. This was slightly lower than expected (0.9^2) and may be caused by the assumption that various factors in the uncertainty analyses were normally distributed. Differences in E_0 between DIR and MMR ranged from −1.5 to 1.5 mm (Supplementary Data D). Further investigation into the differences between the two methods revealed two main causes: 1) Large deviations of mobile structures such as mandible could not always be recovered by DIR; 2) BA structures appeared to have expanded ('creep') after DIR due to structure boundaries not coinciding with anatomical boundaries such as the superior end of the occipital bone, and due to differences in pCT and CBCT Hounsfield Unit (HU) calibration. For the DIR algorithm used in this study, application of local rigidity constraints to the delineated BA may reduce the impact of 'creep' [26,27].

Precision of PG propagation using DIR was evaluated by comparing these results against independent radiation oncologist CBCT re-contours (Supplementary Data H) and were found to be in excellent agreement with values reported in literature [28–30]. Well aligned to literature [16], the PGs included in this study demonstrated an average volume decrease of 21% and an average medial shift of 2.6 mm during HNRT.

Considering the high correlation between these metrics, there was no indication that V_f^{neck} provides a more accurate metric to assess changes of the treatment volume than overall changes in patient weight.

Dosimetric analysis establishing the clinical relevance of the observed non-rigid changes in pose and anatomy is required to define objective decision rules and appropriate thresholds that can be applied in a safety net for patients who exhibit larger non-rigid deformations in the context of PTV margin reduction. We are therefore currently investigating these aspects in concert with the most efficient treatment adaptation approach for anatomical changes.

In conclusion, this study assessed the potential to quantify and monitor an individual patient's deformation, i.e., non-rigid changes in pose and anatomy during HNRT using DIR in combination with SPC. BA deviation *trends* occurring before the end of treatment week four could only be confirmed in 24% of cases when SPC process limits were used, whereas *absolute* BA deviations could be confirmed in 82% of cases when an *a priori* 2 mm clinical limit was used. SPC process limits were useful to detect soft tissue trends occurring before the end of treatment week four which could be confirmed in 90% of cases. The approach proposed in this study could facilitate timely treatment adaption through detection of problematic patient positioning reproducibility and anatomy changes for individual patients.

Acknowledgements

NL received grants from the New Zealand Department of Internal Affairs (Lottery Health Research) and from Universities New Zealand (Edward and Isabel Kidson Scholarship).

Conflict of interest statement

The authors have no conflicts of interest to report.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.phro.2018.12.004.

References

[1] Mendenhall WM, Amdur RJ, Palta JR. Intensity-modulated radiotherapy in the standard management of head and neck cancer: Promises and pitfalls. J Clin Oncol 2006;24:2618–23. https://doi.org/10.1200/JCO.2005.04.7225.
[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;12:127–36. https://doi.org/10.1016/S1470-2045(10)70290-4.
[3] Pacholke HD, Amdur RJ, Morris CG, Li JG, Dempsey JF, Hinerman RW, et al. Late Xerostomia After Intensity-Modulated Radiation Therapy Versus Conventional Radiotherapy. Am J Clin Oncol 2005;28:351–8. https://doi.org/10.1097/01.coc.0000158826.88179.75.
[4] 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.
[5] Rasch C, Steenbakkers R, Van Herk M. Target definition in prostate, head, and neck. Semin Radiat Oncol 2005;15:136–45. https://doi.org/10.1016/j.semradonc.2005.01.005.
[6] Suzuki M, Nishimura Y, Nakamatsu K, Okumura M, Hashiba H, Koike R, et al. Analysis of interfractional set-up errors and intrafractional organ motions during IMRT for head and neck tumors to define an appropriate planning target volume (PTV)- and planning organs at risk volume (PRV)-margins. Radiother Oncol 2006;78:283–90. https://doi.org/10.1016/j.radonc.2006.03.006.
[7] Barker JL, 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. https://doi.org/10.1016/j.ijrobp.2003.12.024.
[8] Houghton F, Benson RJ, Tudor GSJ, Fairfoul J, Gemmill J, Dean JC, et al. Assessment of action levels in imaging strategies in head and neck cancer using tomotherapy. are our margins adequate in the absence of image guidance? Clin Oncol 2009;21:720–7. https://doi.org/10.1016/j.clon.2009.08.005.
[9] Yu Y, Michaud AL, Sreeraman R, Liu T, Purdy JA, Chen AM. Comparison of daily versus nondaily image-guided radiotherapy protocols for patients treated with intensity-modulated radiotherapy for head and neck cancer. Head Neck 2014;36:992–7. https://doi.org/10.1002/HED.
[10] Chen AM, Farwell DG, Luu Q, Donald PJ, Perks J, Purdy JA. Evaluation of the planning target volume in the treatment of head and neck cancer with intensity-modulated radiotherapy: What is the appropriate expansion margin in the setting of daily image guidance? Int J Radiat Oncol Biol Phys 2011;81:943–9. https://doi.org/10.1016/j.ijrobp.2010.07.017.
[11] Chen AM, Yu Y, Daly ME, Farwell DG, Benedict SH, Purdy JA. Long term experience with reduced planning target volume margins and intensity modulated radiotherapy with daily image guidance for head and neck cancer. Head Neck 2014;36:1766–72. https://doi.org/10.1002/hed.
[12] Chen AM, Daly ME, Cui J, Mathai M, Benedict SH, Purdy JA. Clinical outcomes among patients with head and neck cancer treated by intensity-modulated radiotherapy with and without adaptive replanning. Head Neck 2014;36:1541–6. https://doi.org/10.1002/HED.
[13] Navran A, Heemsbergen W, Janssen T, Hamming-Vrieze O, Jonker M, Zuur C, et al. The impact of margin reduction on outcome and toxicity in head and neck cancer patients treated with image-guided volumetric modulated arc therapy (VMAT). Radiother Oncol 2018. https://doi.org/10.1016/j.radonc.2018.06.032.
[14] Moore SJ, Herst PM, Louwe RJW. Review of the patient positioning reproducibility in head-and-neck radiotherapy using Statistical Process Control. Radiother Oncol 2018;127:183–9. https://doi.org/10.1016/j.radonc.2018.01.006.
[15] Das IJ, Cheng CW, Chopra KL, Mitra RK, Srivastava SP, Glatstein E. Intensity-modulated radiation therapy dose prescription, recording, and delivery: Patterns of variability among institutions and treatment planning systems. J Natl Cancer Inst 2008;100:300–7. https://doi.org/10.1093/jnci/djn020.
[16] Brouwer CL, Steenbakkers RJHM, Langendijk JA, Sijtsema NM. Identifying patients who may benefit from adaptive radiotherapy: Does the literature on anatomic and dosimetric changes in head and neck organs at risk during radiotherapy provide information to help? Radiother Oncol 2015;115:285–94. https://doi.org/10.1016/j.radonc.2015.05.018.
[17] Yang H, Hu W, Wang W, Chen P, Ding W, Luo W. Replanning during intensity modulated radiation therapy improved quality of life in patients with nasopharyngeal carcinoma. Int J Radiat Oncol Biol Phys 2013;85:e47–54. https://doi.org/10.1016/j.ijrobp.2012.09.033.
[18] Ahn PH, Chen C-C, Ahn AI, Hong L, Scripes PG, Shen J, et al. Adaptive planning in intensity-modulated radiation therapy for head and neck cancers: single-institution experience and clinical implications. Int J Radiat Oncol Biol Phys 2011;80:677–85. https://doi.org/10.1016/j.ijrobp.2010.03.014.
[19] Schwartz DL, Garden AS, Thomas J, Chen Y, Zhang Y, Lewin J, et al. Adaptive radiotherapy for head-and-neck cancer: Initial clinical outcomes from a prospective trial. Int J Radiat Oncol Biol Phys 2012;83:986–93. https://doi.org/10.1016/j.ijrobp.2011.08.017.
[20] Hunter JS. The exponentially weighted moving average. J Qual Technol 1986;18:203–10.
[21] Borror CM, Montgomery DC, Runger GC. Robustness of the EWMA control chart to non-normality. J Qual Technol 1999;31:309–16. https://doi.org/10.1353/dic.2009.0000.
[22] Eisbruch A, Ten Haken RK, Kim HM, Marsh LH, Ship JA. Dose, volume, and function relationships in parotid salivary glands following conformal and intensity modulated irradiation of head and neck cancer. Int J Radiat Oncol 1999;45:278–9. https://doi.org/10.1016/S0360-3016(99)90269-9.
[23] Dijkema T, Raaijmakers CPJ, Ten Haken RK, Roesink JM, Braam PM, Houweling AC, et al. Parotid gland function after radiotherapy: The combined Michigan and Utrecht experience. Int J Radiat Oncol Biol Phys 2010;78:449–53. https://doi.org/10.1016/j.ijrobp.2009.07.1708.
[24] 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. https://doi.org/10.1016/j.ijrobp.2008.11.035.
[25] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, et al. Evaluation of measurement data—guide for the expression of uncertainty in measurement. Technical Report No. JCGM 100: 2008.
[26] König L, Derksen A, Papenberg N, Haas B. Deformable image registration for adaptive radiotherapy with guaranteed local rigidity constraints. Radiat Oncol 2016;11:1–9. https://doi.org/10.1186/s13014-016-0697-4.
[27] Staring M, Klein S, Pluim JPW. A rigidity penalty term for nonrigid registration. Med Phys 2007;34:4098–108. https://doi.org/10.1118/1.2776236.
[28] Hardcastle N, Tomé WA, Cannon DM, Brouwer CL, Wittendorp PW, Dogan N, et al. A multi-institution evaluation of deformable image registration algorithms for automatic organ delineation in adaptive head and neck radiotherapy. Radiat Oncol 2012;7:90. https://doi.org/10.1186/1748-717X-7-90.
[29] Hvid CA, Elstrøm UV, Jensen K, Alber M, Grau C. Accuracy of software-assisted contour propagation from planning CT to cone beam CT in head and neck radiotherapy. Acta Oncol 2016;55:1324–30. https://doi.org/10.1080/0284186X.2016.1185149.
[30] 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;41:121712https://doi.org/10.1118/1.4901409.