Review of a multidisciplinary team approach to patient positioning in head-and-neck cancer: a quantitative analysis

---

Sarah Joy Moore

2013/2014

A thesis submitted for the degree of

Bachelor of Radiation Therapy with Honours

at the University of Otago, Dunedin, New Zealand

Abstract

Accurate patient positioning is extremely important in radiation therapy for head-and-neck cancer. With the introduction of three-dimensional cone beam computed tomography (CBCT) at the Wellington Blood and Cancer Centre (WBCC), it was agreed that positioning accuracy required improvement. This led to the establishment of a multidisciplinary team (MDT) focused on patient positioning. Following a number of process changes made by the MDT, improvements in setup accuracy were observed but not formally quantified.

The aim of this thesis was to retrospectively quantify setup accuracy at WBCC, using CBCT images of 96 patients treated for head-and-neck cancer. On average 7 CBCT scans per patient were sequentially registered using each of the following match structures: C1-C3, C3-C5, C5-C7, C7-caudal, mandible occipital bone and the larynx. This enabled quantification of patient deformation as the measure of setup accuracy, by calculating the position of each structure relative to C1-C3. Statistical Process Control (SPC) was then used to assess trends in setup accuracy over time, allowing identification of specific time points where improvements occurred and correlation with process changes to be made.

The multiple rigid registration protocol and deformation values calculated for this patient cohort clearly demonstrated the relative movement of anatomical sub regions in the head-and-neck. SPC charts showed that a significant and consistent reduction in deformation was achieved since the instigation of the MDT. A reduction in the magnitude and variation of the patient systematic 3D-deformation vector was observed, from 2.8 mm ± 0.1 mm (1 S.D.) in 2011 to 0.9 ± 0.0 mm (1 S.D.) in late 2013/early 2014. Statistical correlation analysis revealed that the introduction of new

head supports (p = 0.003), as well as retraining of staff in making immobilisation equipment (p = 0.003) had a significant impact on patient systematic deformation. Both of these changes were actions initiated by the multidisciplinary team, which shows that a multidisciplinary approach to patient positioning had a positive impact on setup accuracy in our department.

Acknowledgements

It is a pleasure to thank those people who have helped to make this research project and thesis possible.

First and foremost, I offer my sincerest gratitude to my clinical supervisor, Dr Rob Louwe. Your enthusiasm for this project has been a key to its success, and the support you have provided in my learning and construction of statistical tools in the management of patient positioning has been incomparable. Your expert knowledge and time have been invaluable to my own development along with the development of this project.

Thank you to my academic supervisors, Dr Patries Herst and Paul Kane. Your extensive research experience has helped to shape this thesis and I am extremely grateful for the feedback and input you have both provided. I also extend my thanks to Dr James Stanley from the University of Otago, for the time you provided in reviewing the application of different statistical tools used in this project.

I would like to acknowledge the organisational support from Jennifer de Ridder, Sue Alexander and Lynne Greig at the Wellington Blood and Cancer Centre. The flexibility and support you have all extended to me during this time has been much appreciated. I would also like to thank Stephen Powell for the input you provided around immobilisation equipment used in the department.

Finally, I would like to thank the Wellington Blood and Cancer Centre Continuing Professional Development fund and the Wellington Cancer Centre Trust Fund for providing financial support toward the cost of the programme.

Table of contents

Appendices

List of tables

List of figures

List of equations

List of abbreviations

2D-kV	two dimensional kilo-voltage imaging
3D-CRT	three dimensional conformal radiation therapy
ANOVA	analysis of variance
C	cervical vertebra
CBCT	cone beam computed tomography
CL	centre line
CTV	clinical target volume
EWMA	exponentially weighted moving average
GTV	gross tumour volume
HNC	head-and-neck cancer
HS1	head support 1, Civco Vac-Lok Cushion Type-S
HS2	head support 2, Klarity Custom Head & Should Support Vacuum Bag (9 litre)
HS3	head support 3, Civco Vac-Lok Cushion Type-S + thermoplastic neck roll
HS4	head support 4, Klarity Custom Head & Should Support Vacuum Bag (7 litre)
IGRT	image guided radiation therapy
IMRT	intensity-modulated radiation therapy
Lat	lateral axis
Lng	longitudinal axis
LPL	lower process limit
MDT	multidisciplinary team
OAR	organ(s) at risk

pCT	planning computed tomography scan
PTV	planning target volume
QA	quality assurance
ROI	region of interest
RT	radiation therapist
SPC	statistical process control
SSN	suprasternal notch
TR0	patient cohort where no retrained RTs were involved in mask making
TR1	patient cohort where a minimum of one retrained RTs was involved in mask making
TR2	patient cohort where a minimum of two retrained RTs were involved in mask making
UPL	upper process limit
VMAT	volumetric modulated arc therapy
Vrt	vertical axis
WBCC	Wellington Blood and Cancer Centre

List of symbols

D	total treatment dose
d_2/d_3	SPC bias correction factors
$D_{f,k}^{ROI}$	deformation per ROI, fraction, translation axis
$E_{f,k}^{ref}$	setup error for reference structure C1-C3, per fraction, translation axis
$E_{f,k}^{ROI}$	setup error per ROI, fraction, translation axis
$EWMA$	exponentially weighted moving average
f	imaged fraction
F	total number of treatment fractions or fractionation
F_p^{ROI}	total number of imaged fractions for each patient per ROI
Gy	gray
H	Kruskal-Wallis test statistic
HU	Hounsfield unit
i	patient number
k	translational axis
L	width of control limits
λ	weighting constant for EWMA
mm	millimeter
m_p^{ROI}	patient systematic deformation per ROI
M_{ROI}	group mean of means per ROI
mR	moving range
$\overline{mR}$	mean moving range
n	subgroup size

N_{img}^{ROI}	total number of images registered per ROI
N_p^{ROI}	total number of ROIs per patient
P_{ROI}	total number of patient datasets per ROI
ref_date_i	reference date for pCT/mask making per patient
σ	measure of dispersion in SPC
σ_p^{ROI}	patient random deformation per ROI
σ_{ROI}	group random deformation per ROI
Σ_{ROI}	group systematic dispersion per ROI
v_f^{ROI}	3D vector per ROI, fraction
v_p^m	patient systematic overall deformation
v_p^σ	patient random overall deformation
W	Shapiro-Wilk test statistic
w_p^{EOT}	patient weight at end of treatment
w_p^{pCT}	patient weight at pCT
WL_p	percentage difference weight loss per patient
x	individual value
$\overline{x}$	mean individual value
x_p^{ROI}	individual value (deformation per axis or 3D deformation vector) per patient per ROI

1 INTRODUCTION

The term head-and-neck cancer (HNC) refers to a group of tumours, which may arise in the mucosa of the head-and-neck (including oral cavity and lip, pharynx, larynx and cervical oesophagus), the nasal cavity and paranasal sinuses, salivary glands and skin [1]. For this study, lymphoma patients requiring treatment to the head-and-neck region were also included in this grouping. HNC diagnoses (excluding skin, oesophagus, and lymphoma) account for approximately 2% of all cancer registrations in New Zealand, with an incidence in men that is twice as high as in woman [2]. While the total number of registrations for HNC is low, this population represents a cohort of patients with complex treatment needs, often requiring multidisciplinary input in treatment design [1,3].

Although the treatment approach for HNC is dependent on tumour site, radiation therapy plays an important role both as a primary or adjuvant therapy [3]. Typically, radiation therapy or surgery is employed for early stage disease (Stage I-II), while a combination of radiation therapy, surgery and/or chemotherapy is used for locally advanced disease (Stage II-IV) [3,4].

1.1 Radiation therapy for head-and-neck cancer

Radiation therapy is a local treatment, which for HNC is primarily delivered using external beam X-rays where the total dose of radiation is divided across a number of individual treatments or fractions. [5]. Linear accelerators are used to deliver the treatment and allow for precise dose deposition in the patient by varying the radiation field size and shape as the machine is positioned or rotated around the body [5,6]. One of the major challenges in radiation therapy is to adequately irradiate the tumour without causing toxicity to normal tissues [7]. This is especially true for HNC treatment, where target volumes regularly overlap or abut organs at risk (OAR), such as the salivary glands, oral cavity, brain stem and spinal cord [3,5,7]. In order to help deliver radiation to the targets and avoid OARs, the patient requires a treatment mask, which is used to immobilise the head-and-neck region.

1.1.1 Immobilisation

The treatment mask is individual to the patient, and is made prior to the planning computed tomography scan (pCT). Several different types of treatment masks are commercially available, each showing comparable repositioning accuracy [8–10]. Despite this, movement within treatment masks is not uncommon, and may be observed as variation in the position of different anatomical sub regions on a day-to-day basis [11–16]. At the Wellington Blood and Cancer Centre (WBCC), head, neck and shoulder thermoplastic masks are used for HNC patients. Thermoplastic is a material that softens in warm water, and is then placed over the patient's head-and-neck and locked onto the treatment couch (Figure 1-1) [8,10,17]. Once the mask has been made, it is reused for each fraction throughout the patient's treatment, unless the patient requires a repeat mask impression due to weight loss, tumour shrinkage

or setup issues. While thermoplastic material provides repositioning accuracy to some extent [8–10,18], this product has also demonstrated shrinkage [8]. In order to minimise shrinkage, specific guidelines around cooling time for thermoplastic are provided in our immobilisation protocols [17,19].

Head supports are also used in the immobilisation of HNC patients; these are typically standard shaped headrests or customisable materials that may be shaped to the patient's individual anatomy. Primarily, custom head supports are used at the WBCC, as these provide better reproducibility compared to standard head rests [20,21].

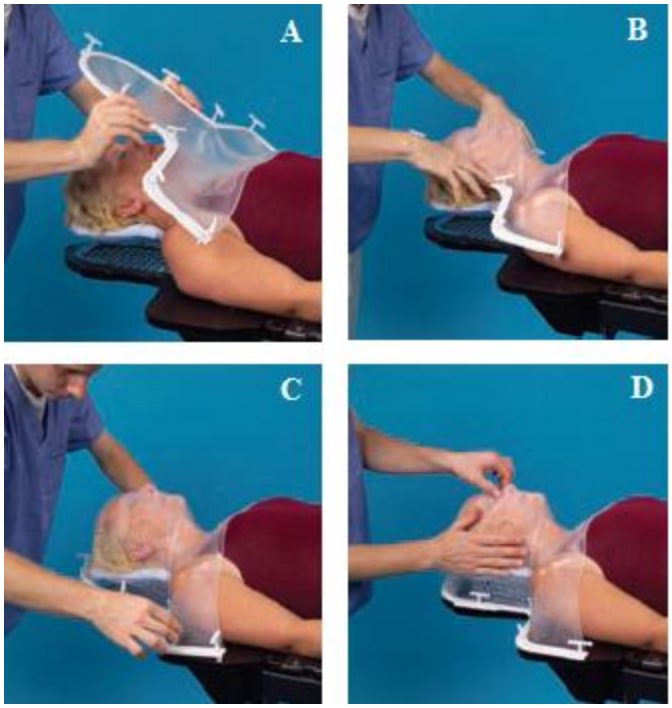


Figure 1-1 Thermoplastic mask impression. A. Thermoplastic mask being lowered over the patient's head-and-neck; B. Stretching of mask over patient's head, neck and shoulders, C. Locking mask onto couch; D. Final impression of patient anatomy. Image from Civco Medical Solution [22].

1.1.2 Treatment planning and delivery techniques

Following immobilisation, the pCT is acquired, which is a snapshot of the patient position in the treatment mask, and is used as the reference scan throughout radiation treatment. During planning, target volumes and OARs are outlined on the pCT, and used to assist plan generation. On treatment, the radiation therapist (RT) team attempt to reproduce the patient position as it was captured on the pCT, by repositioning the patient in the head support and refitting the treatment mask, for each treatment fraction.

In recent years, radiation therapy techniques for HNC have rapidly become more advanced, with treatments such as intensity-modulated radiation therapy (IMRT) and volumetric modulated arc therapy (VMAT). IMRT achieves dose delivery more conformal to the target volume compared to three dimensional conformal radiation therapy (3D-CRT), allowing for tumour dose escalation and increased sparing of normal tissue [6,23]. VMAT is a more recently developed treatment technique and achieves very similar dose distributions compared to IMRT with the treatment delivery generally much faster [24–26]. Faster treatment delivery can potentially reduce motion during treatment (intra-fraction motion), which may occur within the treatment mask. These conformal treatment techniques require accurate patient positioning to avoid geometric misses, which may result in tumour recurrence and/or unnecessary toxicity due to exposure of OARs to excessive dose [7]. Both VMAT and 3D-CRT treatment techniques are used at WBCC (Figure 1-2), depending on treatment indication and extent of disease, with the same immobilisation equipment and positioning techniques used for both treatments.

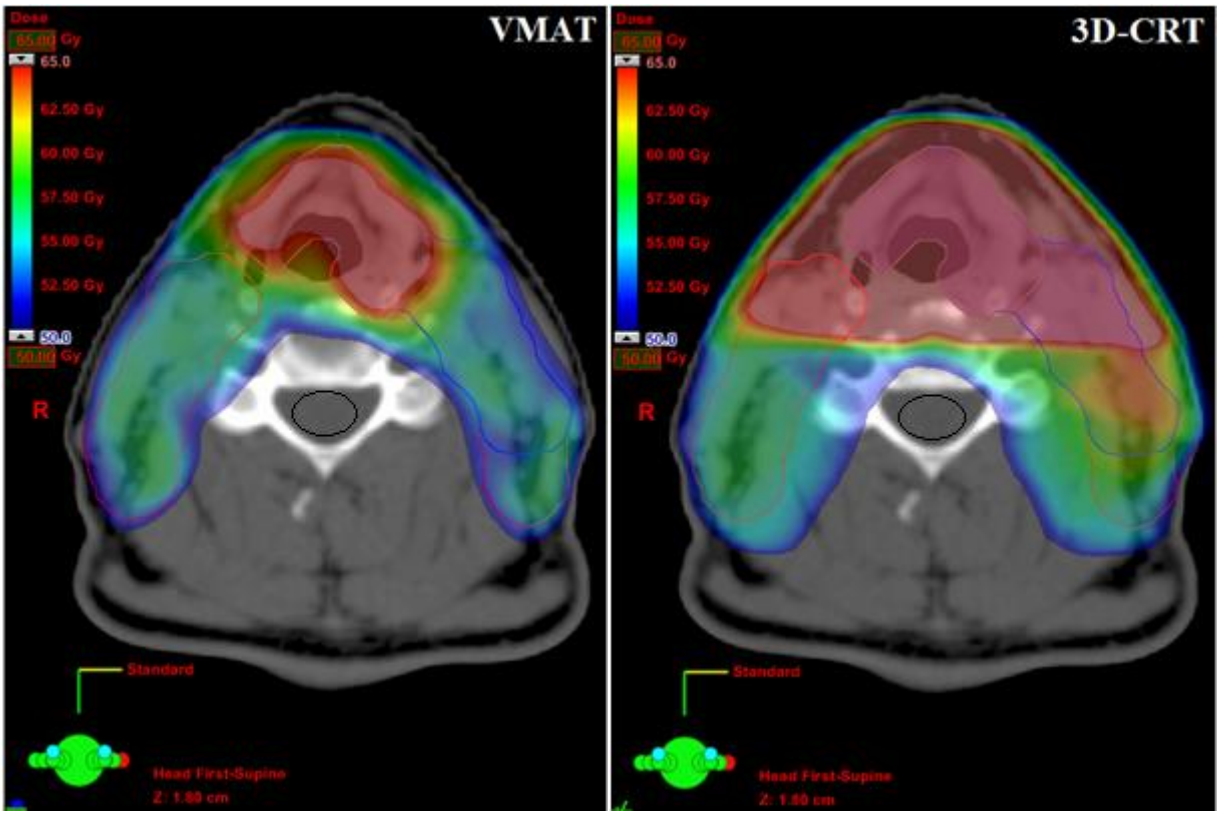


Figure 1-2 Comparison of VMAT and 3D-CRT treatment plans at WBCC. pCT at level of C4/C5 junction, target volumes outlined in red, blue and brown contours, dose displayed from 50Gy (blue) to 65Gy (red). Note increased dose conformity to target volume particularly around the high (red) dose volume and reduced dose around spinal cord (black contour) with VMAT.

1.1.3 Image Guided Radiation Therapy

Sufficient positioning accuracy is typically verified using image guided radiation therapy (IGRT) [27]. The term IGRT refers to imaging processes in both treatment planning and treatment delivery areas of radiation therapy [28]. For this study, IGRT includes the use of imaging techniques at the treatment machine and the assessment of patient position compared to the pCT.

At the WBCC, IGRT for HNC is primarily performed using two-dimensional kilo-voltage (2D-kV) imaging, with CBCT used as a secondary imaging tool. Both techniques are acquired using a kilo-voltage X-ray source and flat panel detector, which are mounted on the gantry of the linear accelerator perpendicular to the treatment beam [27,29] (Figure 1-3). 2D-kV images are captured as paired static X-ray images at orthogonal angles, 90º apart [30]. CBCT is acquired by rotating the gantry around the patient, capturing multiple image projections every 1-2º of rotation, which are then reconstructed to give a volumetric image dataset [29].

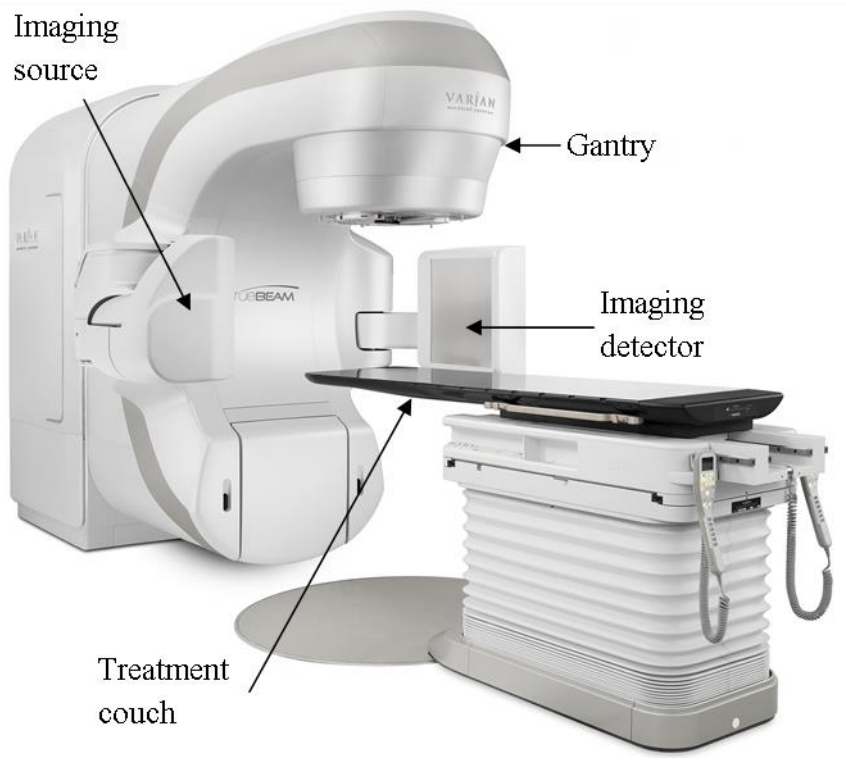


Figure 1-3 Varian TrueBeam linear accelerator. Radiation treatment is delivered from the head of the gantry, with imaging arms attached perpendicular to the treatment beam. Imaging arms identified as kilo-voltage X-ray imaging source and flat panel imaging detector. Treatment couch where patient is positioned. Image courtesy of Varian Medical Systems, Inc [31]. All rights reserved.

Using IGRT, images are acquired with the patient in the treatment position. Positioning errors may be measured at the machine console and if necessary corrected immediately prior to treatment delivery [27]. At the WBCC, translational errors are assessed and automatically corrected by remotely adjusting the treatment couch from the machine console. We do not currently have the technology required to correct pitch and roll, using a so-called 6D-treatment couch [32], while remote motion of couch rotation from the treatment console is currently not included in our IGRT protocol [33]. As a result, the correction of translational errors handles only part of the geometric uncertainties observed in repositioning patients for treatment. Residual errors due to rotations and the relative motion of different anatomical sub structures remain [11,34].

1.2 Management of uncertainties

Safety margins are used to reduce the impact of uncertainties in radiation therapy, and are applied to the clinical target volume (CTV) during the planning process [34]. The CTV contains the gross tumour volume (GTV) where present, plus any microscopic disease that needs irradiating [35,36]. A safety margin is then applied to obtain a planning target volume (PTV), which is a geometric concept used to ensure the CTV receives adequate dose when uncertainties are present (Figure 1-4). Several formulas for calculating appropriate PTV margins exist, each of which requires consideration of various systematic and random uncertainties across the entire radiation treatment process, to ensure accurate and safe treatment [34].

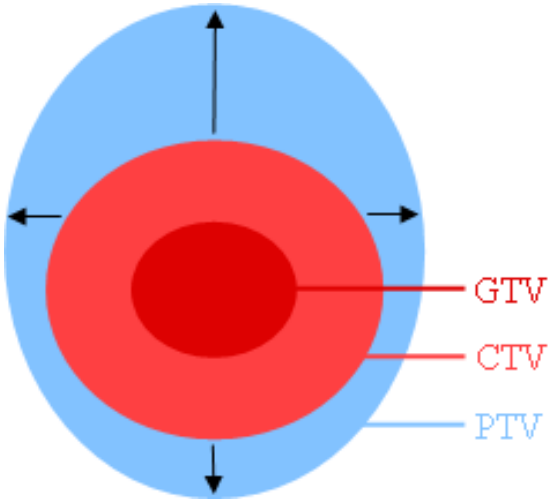


Figure 1-4 Basic representation of GTV, CTV and PTV. Gross tumour volume (GTV) in dark red, clinical target volume (CTV) in light red, planning target volume (PTV) in blue. Adapted from International Commission on Radiation Units and Measurements [35].

Setup uncertainties are commonly divided into systematic dispersion and random errors [34]. Systematic dispersion is the variation of the per patient average setup error [37], and is often confusingly referred to as systematic patient setup error. Systematic dispersion errors are typically caused by (but are not necessarily limited to) treatment preparation errors [34]. Random errors represent the fraction-to-fraction variation of the patient setup error [34,37]. Not included in this categorisation are systematic departmental setup errors such as incorrect protocols or faulty equipment, as these do not represent a variation or uncertainty. Therefore, these deviations should be minimised by means other than a statistical approach or PTV margin. Furthermore, it should be noted that some overlap between all of the above categories may exist due to time trends, which are generally not captured by the descriptive statistics used to determine the various setup error contributions [38]. For instance, these time trends can be caused by changes in patient anatomy or degradation of immobilisation equipment over the treatment course, or slow degradation of treatment couch during its lifetime. In addition, patient deformation due to, for example organ motion, weight loss and tumour shrinkage, or poor setup reproducibility, is generally not included in PTV margin calculations [11–15,34,39].

1.3 Deformation

Deformations and anatomical changes are common in patients with HNC [11,40]. Deformation is a current focus in the assessment of patient positioning for HNC [11–16], particularly as these uncertainties are generally not accounted for using translational adjustments or PTV margins [28]. As discussed in this study, deformation relates to the relative movement of different anatomical sub regions, which may be due to changes in posture or flexion in the head-and-neck region, weight loss or tumour shrinkage [15].

Several studies have reviewed deformation during HNC treatments, using different 3D imaging modalities [11–16]. The benefit of 3D information for reviewing deformation is that anatomical and positional changes that cannot be accurately assessed with 2D-kV imaging may be reviewed [11]. Using 3D data, multiple rigid registrations can be performed and the movement of a specific region of interest (ROI) relative to a reference ROI can be calculated. As a result, a number of studies reported deformation within the head-and-neck region exceeding the applied PTV margins, despite the use of immobilisation equipment [11–16]. Clinically, deformation may result in reduction of dose to the target or increased irradiation of OARs, leading to possible tumour recurrence or unnecessary toxicities.

1.4 Status of patient positioning at WBCC

At the time of implementation of CBCT at WBCC, imaging showed that patient positioning for HNC treatments was not optimal and was thought to require improvement. In particular, differences in posture or neck flex were regularly observed during IGRT. As demonstrated by Figure 1-5, registration of the pCT and CBCT on C1-C3 shows a considerable amount of deformation throughout the scan volume, with mismatches observed in the spine, mandible, occipital bone and larynx.

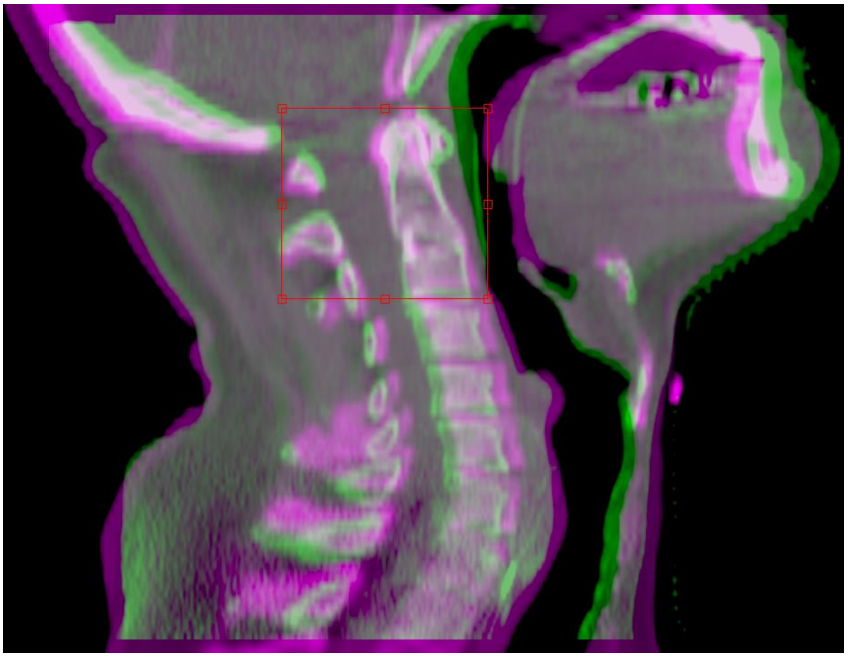


Figure 1-5 Example of HNC patient positioning during 2011/2012 period at WBCC. pCT (purple) and CBCT (green) registered on C1-C3. Superimposed image highlights areas where a good match (white) or mismatch between the reference and treatment images is obtained (either green or purple).

The observation that setup accuracy needed to improve led to the establishment of a multidisciplinary team (MDT) focused on patient positioning, with representation from the radiation oncologists, medical physicists and radiation therapists. The MDT undertook a systematic review of positioning policies and procedures, with the aim of identifying possible areas for improvement. From this review, the MDT formulated five main recommendations, which subsequently instigated the following changes;

1. A significant overhaul of IGRT protocols, including a shift to online imaging corrections and increased utilisation of CBCT, as well as the definition of specific instructions for image matching and review.

2. Instigation of weekly multidisciplinary image review meetings.

3. Increased preparation documentation and photographs required at the time of mask making for HNC patients, in order to provide more detailed information for setup at the treatment machines.

4. A review of the types of head supports being used in the immobilisation of HNC patients.

5. Retraining of radiation therapists in immobilisation equipment.

These changes eventually lead to improved patient positioning being observed during weekly image review meetings. An example of patient positioning in 2013/2014 is demonstrated by Figure 1-6, where registration of the pCT and CBCT on C1-C3 shows considerably less deformation, compared to Figure 1-5. While these improvements were encouraging, trends in setup accuracy were not quantified in a formal way.

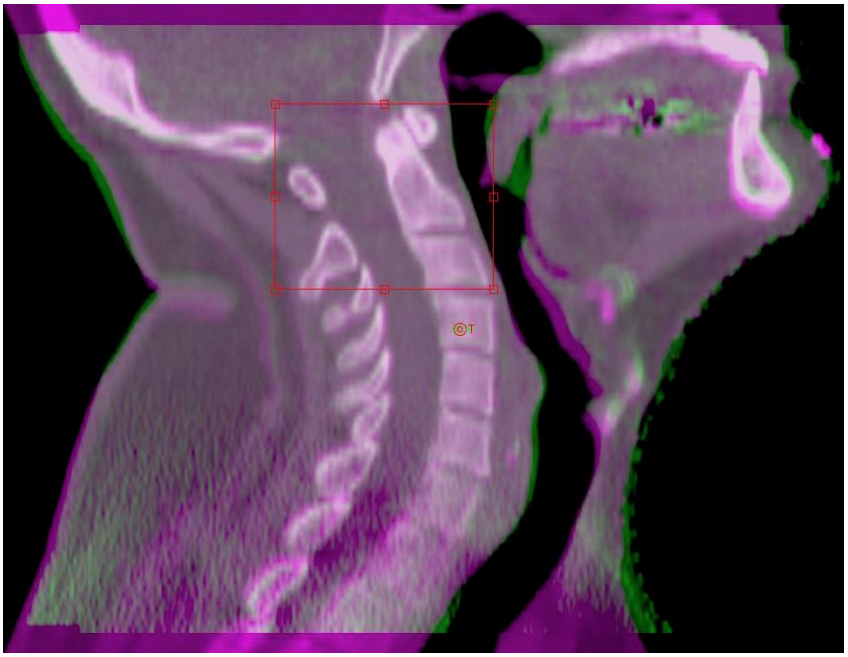


Figure 1-6 Example of HNC patient positioning during 2013/2014 period at WBCC. pCT (purple) and CBCT (green) registered on C1-C3. Superimposed image highlights areas where a good match (white) or mismatch between the reference and treatment images is obtained (either green or purple).

1.5 Statistical Process Control

In this thesis, Statistical Process Control (SPC) charts are a primary method for the analysis of the results, and are therefore separately introduced in this section. SPC refers to a collection of quality management tools, widely used in manufacturing processes since the 1950's [41]. One of the most prominent parts of SPC is the control chart, first developed by Walter A. Shewhart of Bell Telephone Laboratories in the 1920's [41–43]. Control charts operate based on Shewhart's theory of variability, which assumes that natural variation will exist in any process, even in well controlled processes [41,42]. This variation is therefore referred to as controlled variation, while uncontrolled variation relates to exceptional deviations [43]. Control charts use process limits to differentiate between these two types of variation [41–43]. Figure 1-7 demonstrates a generic control chart, with the key quality measure plotted on the *y*-axis and time or patient number on the *x*-axis. Process limits are calculated based on a reference group, which represents a period of process stability [42] and is plotted on the example control chart in red. Process limits are calculated as a centre line (CL), representing the average value of the data for the reference group, as well as an upper process limit (UPL) and lower process limit (LPL) [41–43]. Upper and lower process limits are usually designed using three sigma ($\pm 3\sigma$) either side of the CL, where sigma represents the dispersion statistic calculated for the data [41]. The selection of three sigma limits is often described by the empirical rule; where given a homogenous, normally distributed dataset, 99.7% of data will fall within three sigma units [44]. Observations outside these limits are deemed worthwhile investigating [43].

Once process limits have been calculated, charts may be plotted in what is typically referred to as the observation period [45]. Data is commonly plotted prospectively, using a scatter plot with straight line connectors, which allows trends in process stability to be more easily observed [41]. Plotting data in real time allows for prompt investigation of exceptional variation [41–43,45]. Advantages of SPC charts include a consistent and effective approach in process control. This is achieved through pre-defined rules of interpretation, which may vary depending on the process or type of data, as well as the type of control chart being used [41,42].

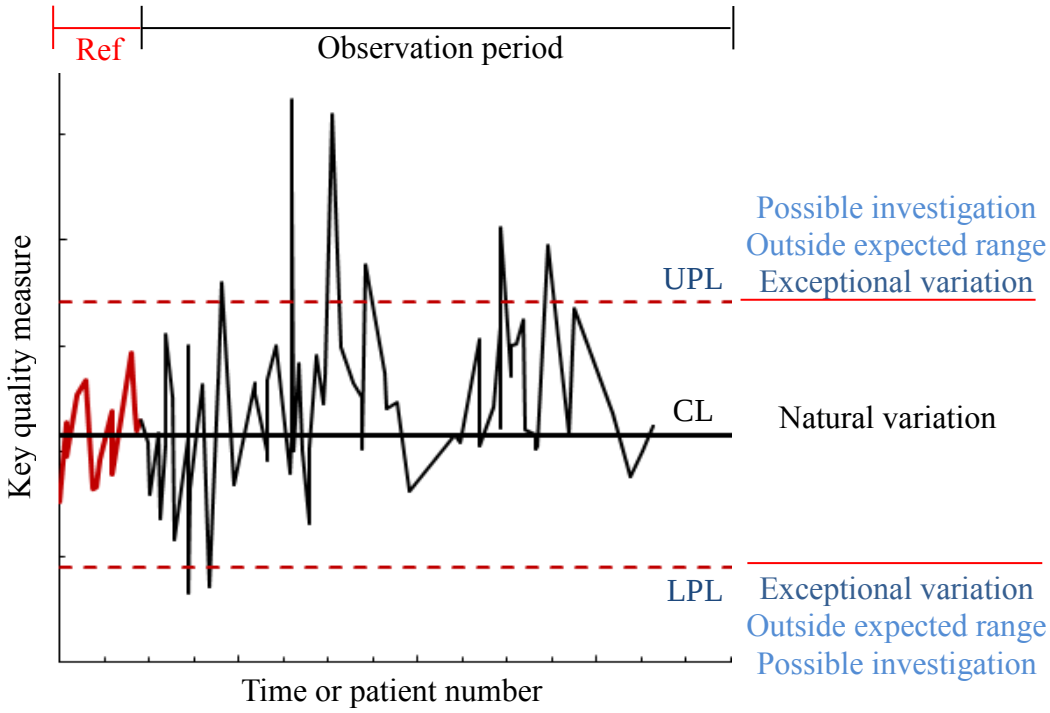


Figure 1-7 Example of generic SPC control chart. Adapted from Pawlicki *et al.* [45].

A wide range of SPC charts are available to monitor process stability, with the type of control chart based on the kind of data being analysed [41,43]. Control charts for continuous data may be based on subgroups or individuals. For example, mean and range charts are calculated using rational subgroups, with subgroup mean or range

plotted as a single data point on the relevant control chart [42]. However, when subgroups cannot be defined in a meaningful way, individual value (*x)* and moving range (*mR)* charts are typically used [41–43]. Commonly, Exponentially Weighted Moving Average (*EWMA*) charts are used in combination with *x*- and *mR*-SPC charts. As the name implies, these charts display a weighted average of historical data points from the *x*-SPC and are used to highlight small shifts of the process mean [41,46].

The role of SPC as a tool in radiation therapy quality assurance (QA) was previously discussed by Pawlicki *et al.* [38] in 2005. At this time, a limited number of references on this topic were available from the literature [38,47,48]. Since then, there have been several publications on SPC for linear accelerator QA [38,49] and IMRT/VMAT QA [50–56]. Despite a thorough literature review, only one paper discussing the use of SPC in patient positioning was found, in which Ung & Wee [57] used control charts to identify positioning errors in prostate patients.

1.6 Aim and objectives

As discussed in Section 1.4, a multidisciplinary team was set up and made a number of process changes, after which improved positioning accuracy was observed. However, this improvement in setup accuracy was not quantified, nor was a specific cause of change identified.

Aim

The overarching aim of this project was to retrospectively quantify the improvement of patient setup accuracy in radiation treatments for head-and-neck cancer at the WBCC. This aim was achieved by investigating the following objectives:

Objectives

1. Determine setup accuracy, specifically as the level of deformation during HNC treatment.
2. Assess trends in patient positioning over time.
3. Investigate possible causes of change in setup accuracy.
4. Benchmark setup accuracy in terms of deformation with published literature.
5. Discuss different measures for the management of patient positioning in HNC treatments in the future.

2 MATERIALS & METHODS

Local ethics approval was considered but was not required because all patient information and imaging was collected retrospectively from existing patient records, without collecting additional data. The study was registered with the hospital research department as a clinical audit.

2.1 Patient cohort

This study analysed setup accuracy using CBCT scans of 96 patients. All patients were treated for HNC using external beam radiation therapy at the WBCC. Patients were retrospectively selected from a cohort of 131 consecutive HNC patients, who had been imaged using CBCT. All patients were treated between May 2011 and March 2014, under four different imaging and positioning protocols (Section 2.3). Patients were selected irrespective of age, gender, tumour diagnosis and treatment technique. There were two main inclusion criteria to ensure sufficient data to estimate the average deformation for each patient, along with a minimum of three regions of interest on which to calculate deformation (Section 2.4**)**:

1. Availability of at least three weekly CBCTs or daily CBCT over a minimum of 10 treatments.
2. Minimum CBCT field of view encompassing C1 vertebra superiorly and/or C5 vertebra inferiorly. Anterior-posterior and lateral scan borders were determined by the field of view, which was sufficient to include relevant anatomy and therefore specific limits were not applied in these directions.

Overall, 23 potential patients were excluded due to insufficient number of images, while 11 patients were excluded due to inadequate scan lengths. One further patient

was excluded due to poor image quality. Five patients required a repeat pCT and/or mask during treatment, after which a new treatment plan was created for the remaining fractions. These new plans were treated as a separate entity within the study database, with separate datasets created for each new pCT. Therefore, a total of 101 patient datasets were available for analysis across 96 patients.

2.2 Patient positioning

For the purpose of this study, the date of pCT acquisition and mask making was used as the reference date for analysis. For all treatments, the isocentre was used as the reference point for positioning corrections.

All patients were immobilised using a Kevlar Reloadable Head and Shoulder S-Frame Mask, 2.4mm as shown in Figure 2-1 (Q-Fix, Avondale PA)

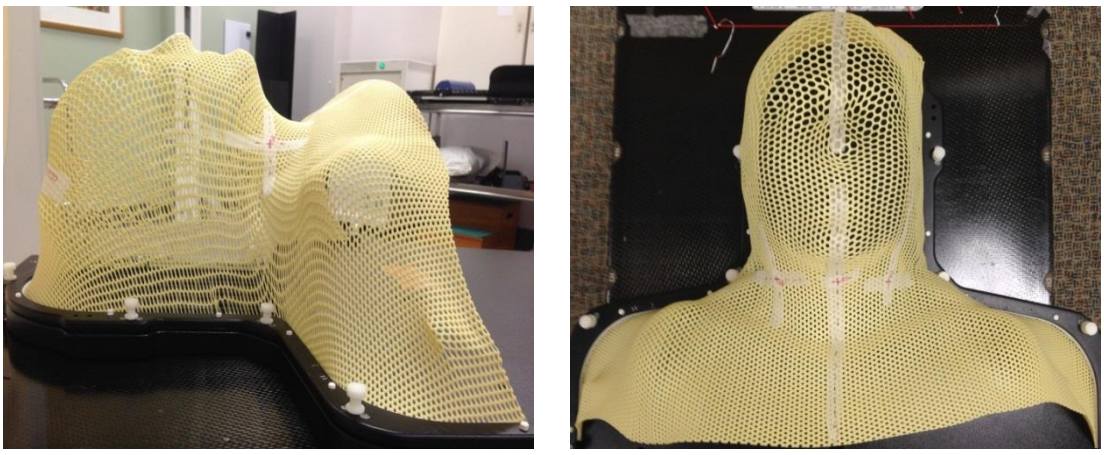


Figure 2-1 Top and side views of Kevlar Reloadable Head and Shoulder S-Frame Mask 2.4mm, without patient. Not to scale.

Various head support products were used for patient immobilisation during the study period. These included both standard shaped head rests and custom made head supports as listed;

- Q-Fix Standard Silverman B Head Rest (Q-Fix, Avondale, PA)
- Q-Fix Standard Silverman F Head Rest (Q-Fix, Avondale, PA)
- Civco Custom Vac-Lok Cushion Type-S, 61x65cm, 8 litre fill, nylon with indexing (Civco Medical Solutions, Kalona, IA).

- Klarity Custom Head and Shoulder Support Vacuum Bag, 62x67cm, 9 litres fill, nylon with in-house indexing (Klarity Medical Products LLC., Newark, OH).
- Klarity Custom Head and Shoulder Support Vacuum Bag, 62x67cm, 7 litres fill, nylon with indexing (Klarity Medical Products LLC., Newark, OH).

Each custom head support is a vacuum cushion containing polystyrene beads, which forms a rigid mould when the air is removed (vacuumed) [58]. These head supports are all re-usable once the patient has completed treatment.

Table 2-1 Main head support type, reference label and incidence

Head Support Type	**Reference Label**	**Incidence (*n*)**
Civco Vac-Lok Cushion Type-S	HS1	42
Klarity Custom Head & Should Support Vacuum Bag (9 litre)	HS2	26
Civco Vac-Lok Cushion Type-S + thermoplastic neck roll	HS3	14
Klarity Custom Head & Should Support Vacuum Bag (7 litre)	HS4	6

For this study, only three patients used the standard Silverman B or F head rests while one patient used the Silverman F with a Civco Vac-Lok Cushion. In order to improve patient set-up, in-house developed support materials were also used for some patients. Two types of neck roll were developed in-house, one using thermoplastic material and another made from polystyrene. The purpose of these neck rolls was to provide a more standard neck shape with vacuum bags, where the polystyrene fill was not sufficient to provide adequate neck support. Neck rolls were used for 17 patients, 14 of which were thermoplastic neck rolls in conjunction with Civco Vac-Lok cushions

(HS3). A polystyrene neck roll combined with a Civco Vac-Lok cushion was used for two patients, and a thermoplastic neck roll was used in conjunction with a Klarity 9 litre vacuum bag only for a single patient. Table 2-1and Figure 2-2 outline the main head supports used in this study.

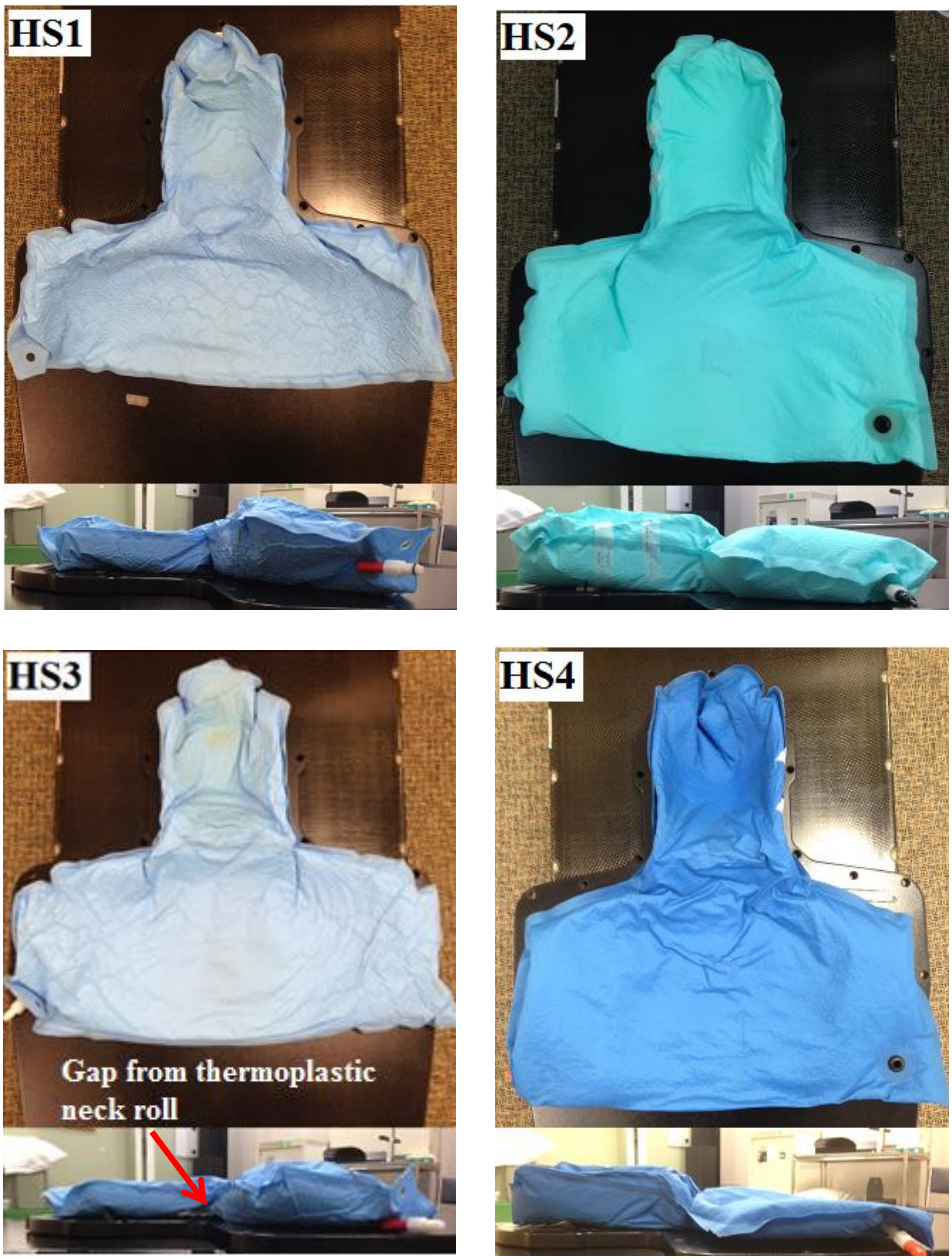


Figure 2-2 Top and side views of the four main head support types. Not to scale. See Table 2-1 for more information on each head support type. Note the difference between HS1 and HS3 as the gap under the neck where the thermoplastic neck roll is placed (indicated by red arrow).

Mouth pieces, which are identified as bite blocks (n = 5) or tongue depressors (n = 18), were used to remove the palate from the irradiated volume and/or to stablise the positioning of the tongue (Figure 2-3). Bite blocks were created using a plastic syringe plus adapt-it thermoplastic pellets (Q-Fix, Avondale, PA). Tongue depressors were made using a flat plate with 2/3 20mL syringe plus adapt-it thermoplastic pellets (Q-Fix, Avondale, PA).

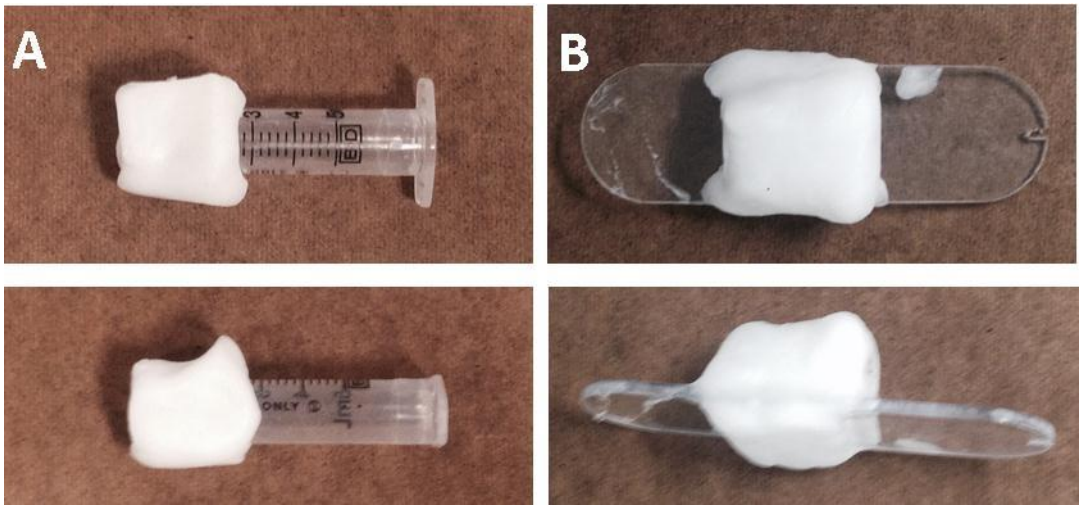


Figure 2-3 Top and side view of bite block (A) and tongue depressor (B). Not to scale. Adapt-it thermoplastic pellets (white) are moulded around each mouth piece and soften when heated to get impression of teeth/gums.

Shoulder retractors were used during mask making for all patients. Thirty-seven patients also used shoulder retractors during treatment but this was abandoned for later treatments after changes made by the MDT. A knee rest was used for all patients to improve comfort, with two patients requiring an additional pillow under the knees.

An anterior reference mark, inferior to the mask was used to aid in reproducing patient position at mask making/acquisition of the pCT. This was either a tattoo (n = 84), nevus (n = 2) or nipple line (n = 3). Seven patients had no anterior reference, because the treatment volumes were located superiorly, not extending into the lower

neck. Initially, the distance between the reference marks on the mask and the anterior reference (n = 36) were used for initial patient setup. After changes in protocol for HNC treatment preparation made by the MDT, this distance was measured from the head support indexing the treatment couch (n = 60).

Bolus was used for 11 patients to provide sufficient dose delivery for superficial target volumes, and was placed either inside the mask (n = 10) or outside the mask (n = 1) as was appropriate.

2.3 Treatment planning and delivery

This section provides background information on treatment planning and delivery approach. This information did not specifically impact the analysis of the results but has been included to provide an overview of these aspects of HNC treatment at the WBCC.

pCT scans were acquired on a Philips Brilliance Big Bore CT Scanner (Koninklijke Philips *N*.V., Eindhoven, The Netherlands). Scans were reconstructed in 3 mm slice thickness, with a 1x1 mm^2 in-plane spatial resolution.

VMAT (n = 57) and 3D-CRT (n = 39) were the treatment techniques used during this study. All treatment plans were generated using Eclipse Treatment Planning System (Varian Medical Systems Inc., Palo Alto, CA), with treatment delivery performed using a Varian Clinac Ex equipped with On-Board Imaging or a Varian Truebeam accelerator (Varian Medical Systems Inc., Palo Alto, CA).

Four image guidance protocols were applied throughout the study period, with changes in on-treatment imaging protocols due to the introduction of VMAT and changes initiated by the MDT for patient positioning. A summary of each protocol and the incidence of application are provided in Table 2-2.

Table 2-2 Clinical image guided radiation therapy protocols and incidence

IGRT Protocol	Modality	Frequency	Online match	*n*
IGRT Protocol 1	2D-kV	#1-2	Yes – shift >0.3cm	19
	CBCT	#3, weekly	Yes – shift >0.3cm	
IGRT Protocol 2	2D-kV	Daily	Yes – shift >0.0cm	24
	CBCT	Weekly	Yes – shift >0.0cm	
IGRT Protocol 3	2D-kV	Daily	Yes – shift >0.0cm	50
	CBCT	Weekly	No, only verification	
IGRT Protocol 4	CBCT	Daily	Yes – shift >0.0cm	12

Eight patients required different imaging approaches due to setup difficulties during treatment, which were managed through increased imaging frequency. Effectively, this resulted in the application of a combination of two of the four protocols for these patients during their course of treatment, as Protocols 2 and 4 (n = 2), 3 and 4 (n = 5), 1 and 3 (n = 1).

For all protocols, rigid registrations of the pCT and 2D-kV or CBCT images were performed using bony anatomy. 2D-kV image registrations were done using manual matching processes, while registrations were typically performed using automatic registration tools for CBCT imaging.

All imaging was performed prior to treatment delivery. Where both modalities were performed on the same treatment fraction as in, for example, IGRT protocol 3 or 4, 2D-kV images were acquired first for online position correction, followed by CBCT for weekly verification of the correctness of positioning based on 2D-kV imaging.

2.4 Multiple rigid registration protocol

For this study, image registrations were performed using Varian Offline Review software (Varian Medical Systems Inc., Palo Alto, CA). All registrations were done by the researcher as the sole observer. An intra- and inter-observer audit of 15 patients (2 fractions per patient) was performed to confirm that these registrations were consistent and accurate.

Quantification of deformation in head-and-neck setups was achieved by performing multiple rigid registrations on each imaged fraction (*f*). Registrations were done using eight separate ROIs; C1-C3, C3-C5, C5-C7, C7-caudal, mandible, occipital bone, suprasternal notch (SSN) and larynx (Figure 2-4). This selection of ROIs reproduced those used by van Kranen *et al.* [11], with the only difference being the inclusion of C7 in the inferior spine ROI. These match structures encompassed those regions typically included in the irradiated volume for many head-and-neck cancers [11], and were consistent with other studies, allowing comparisons to be made with previous work [12–14].

In order to efficiently register each sub region in an highly automated way, the ROI and Intensity Range tools available in Offline Review were used [30]. For each registration, the ROI box was defined on the pCT to cover the relevant anatomy without additional margin (Figure 2-4). Each ROI was matched one at a time, over the entire patient dataset, to ensure consistent ROI definition for each individual patient. The Bones Intensity Range with Hounsfield Unit (HU) limits of 200 to 1700 was used for all automatic registrations excluding larynx. This guaranteed that only those voxels within the specified HU range were used by the 3D Match algorithm [30].

Considering the low density cartilage in the larynx, the Intensity Range was not restricted to a specific HU range while matching this ROI.

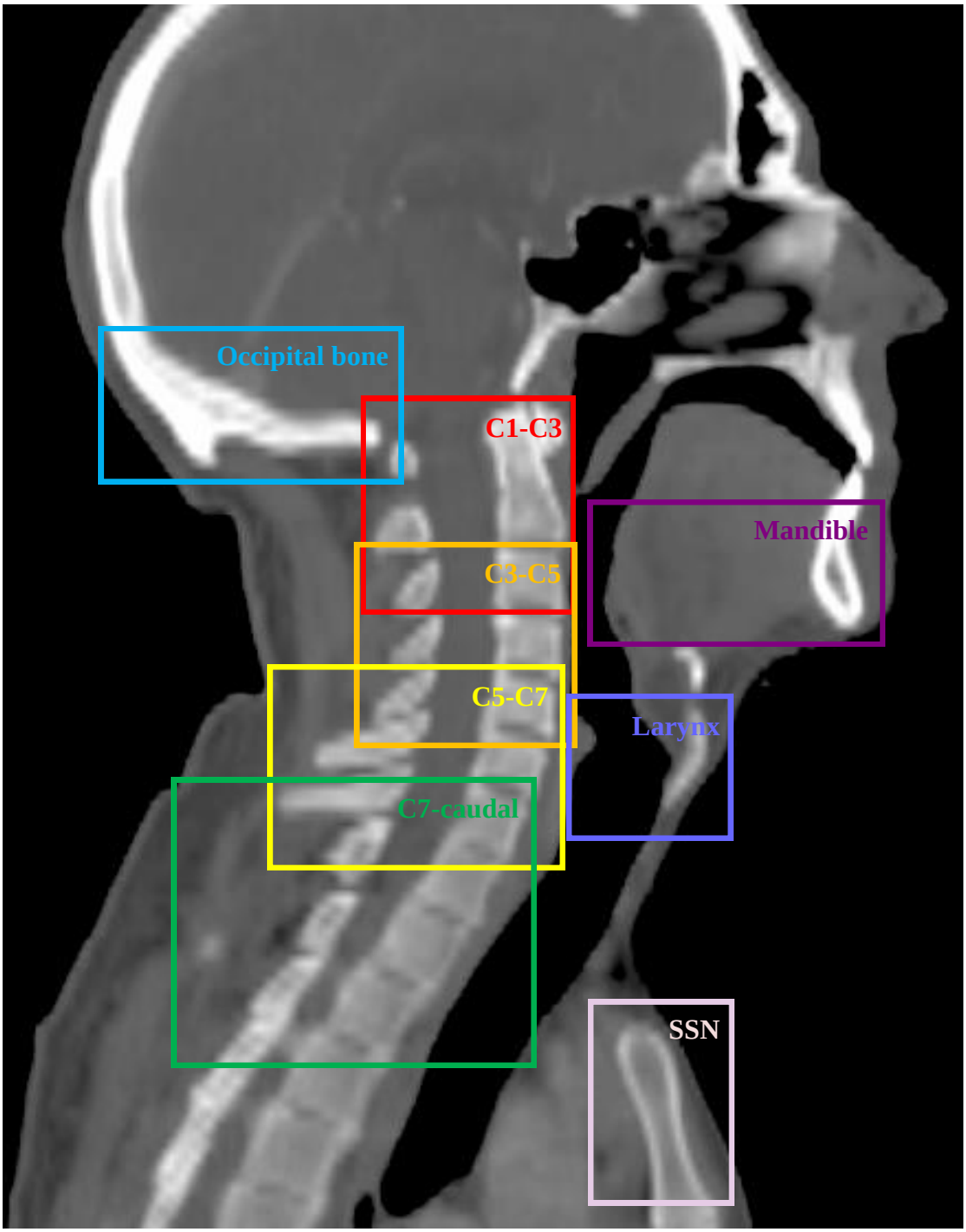


Figure 2-4 Definition of the eight regions of interest used in the multiple rigid registration protocol

Automatic registrations were then completed using the default match setting of the Offline Review software. This setting applied a three step automatic match configuration. In each step, a Downhill Simplex optimiser and Mutual Information Similarity Measure was used, with increasing internal resolution from 8 mm to 4 mm to 2 mm [30]. All automatic matches were visually inspected and manually adjusted if required.

Three degrees of freedom were included in each registration; Vertical (Vrt), Longitudinal (Lng) and Lateral (Lat). Rotational errors were not included.

In total, 740 CBCT images were analysed in the study protocol over 101 patient datasets. This resulted in 7.3 CBCT images on average, and a range of 3-30 images per dataset. All results were recorded in a database, which was created using Excel (Microsoft v 2010 Redmond Campus, Redmond, WA).

2.5 Setup uncertainties

2.5.1 Calculation of deformation

Deformation of an ROI at a particular fraction was defined as the geometrical shift of that ROI with respect to the pCT scan and relative to that of a reference ROI. Deformation was therefore calculated for each ROI, translational axis (*k*) and fraction (*f*) by subtracting the match results of the reference ROI ($E_{f,k}^{ref}$) from the individual ROI ($E_{f,k}^{ROI}$), as follows:

$$D_{f,k}^{ROI} = E_{f,k}^{ROI} - E_{f,k}^{ref} \quad \textbf{Eq 1}$$

where *k* represents either the vertical, longitudinal, or lateral axis. In addition, the resultant 3D vector length (v_f^{ROI}) of these three orthogonal components were calculated for each ROI and fraction (*f*) using [37]:

$$v_f^{ROI} = \sqrt{\left[D_{f,vrt}^{ROI}\right]^2 + \left[D_{f,lng}^{ROI}\right]^2 + \left[D_{f,lat}^{ROI}\right]^2} \quad \textbf{Eq 2}$$

C1-C3 was selected as the reference ROI. As described by van Kranen *et al.* [11], C1-C3 is often central to both the target and organ at risk volumes, and is also supported by the head rest, giving the expectation of reproducible positioning. C1-C3 is also commonly used as a reference structure at the WBCC when clinically matching head-and-neck images. It allowed not only comparison between ROIs across datasets within this study, but also with the results of previous work [11–14].

2.5.2 Patient statistics

Calculation of patient statistics for each translational axis and 3D vector were performed using the methods described by Remeijer *et al.* [37] for conventional set-up errors. The patient systematic deformation (m_p^{ROI}) and patient random deformation (σ_p^{ROI}) were calculated per ROI, using all imaged fractions per patient (F_p^{ROI}):

$$m_p^{ROI} = \sum_{f=1}^{F_p^{ROI}} \frac{x_p^{ROI}}{F_p^{ROI}}$$

Eq 3

$$\sigma_p^{ROI} = \sqrt{\frac{1}{F_p^{ROI} - 1} \sum_{f=1}^{F_p^{ROI}} (x_p^{ROI} - m_p^{ROI})^2}$$

Eq 4

where x_p^{ROI} can either be the deformation per axis per ROI ($D_{f,k}^{ROI}$), or the 3D deformation vector for each ROI (v_f^{ROI}).

The overall 3D deformation was also calculated for each dataset by averaging the 3D vectors across all available ROIs per dataset, where N_p^{ROI} represents the number of ROIs registered per patient;

$$v_p^m = \sum_{p=1}^{ROI_p} \frac{m_p^{ROI}}{N_p^{ROI}}$$

Eq 5

$$v_p^\sigma = \sum_{p=1}^{ROI_p} \frac{\sigma_p^{ROI}}{N_p^{ROI}}$$

Eq 6

2.5.3 Group statistics

Calculation of group statistics was also performed using methods described by Remeijer *et al.* [37] for conventional set-up errors.

The overall mean of patient means (M_{ROI}) was calculated for each ROI, using the total number of imaged fractions (N_{img}^{ROI}) for all patient datasets (P_{ROI}):

$$N_{img}^{ROI} = \sum_{p=1}^{P_{ROI}} F_p^{ROI}$$

Eq 7

$$M_{ROI} = \frac{1}{N_{img}^{ROI}} \sum_{p=1}^{P_{ROI}} \sum_{f=1}^{F_p^{ROI}} x_p^{ROI}$$

Eq 8

The variation of the patient systematic deformation was calculated to give the group systematic dispersion per ROI, Σ_{ROI};

$$\Sigma_{ROI} = \sqrt{\frac{P_{ROI}}{N_{img}^{ROI}(P_{ROI}-1)} \sum_{p=1}^{P_{ROI}} F_p^{ROI}(m_p^{ROI} - M_{ROI})^2}$$

Eq 9

The group random deformation per ROI, σ_{ROI}, was calculated by the root mean square of the patient standard deviations;

$$\sigma_{ROI} = \sqrt{\frac{1}{N_{img}^{ROI} - P_{ROI}} \sum_{p=1}^{P_{ROI}} \sum_{f=1}^{F_p^{ROI}} (x_p^{ROI} - m_p^{ROI})^2}$$

Eq 10

2.6 Normal probability test

A normal probability test was conducted for patient systematic (v_p^m) and random (v_p^σ) overall deformation. Histograms were used to provide a visual inspection of the study distributions in comparison to a bell curve for a normal distribution [44]. The data was also graphed using quantile-quantile normal probability plots. In this test, the study data (observed quantiles) were plotted against expected quantiles from a normal distribution and interpreted against a fit line; where departures from the fit line indicated a non-normal distribution [59]. As both of these tests were visually interpreted, subjectivity in the results may occur. Therefore, a Shapiro-Wilk test was also performed in MATLAB (The Mathworks, Natick, MA, USA) to give a more objective measure of normality. The Shapiro-Wilk tests the normality of data against the test statistic W, under the null hypothesis that $W = 1$ for normally distributed data [60]. The threshold for significance in the Shapiro-Wilk test was applied as $p < 0.05$.

2.7 Statistical Process Control

2.7.1 Control chart selection

As SPC charts were used in this study to investigate a change in setup accuracy over several years including many patients, overall *per patient* deviations were selected as the focus for SPC review. Specifically, both the systematic deformation and the variation in patient deformation over the treatment were investigated (Eq 3-Eq 6). Considering that the data of each patient would therefore be represented by a single point, *x*- and *mR* –SPC charts are presented rather than SPC charts monitoring the behaviour of subgroups. The *EWMA*-SPC chart was also included in the control chart selection, as is commonly used with *x*- and *mR*-SPC charts [41]. This chart type is sensitive in detecting small shifts in process mean [41,46], which based on the aim of this study, was considered useful to identify changes in patient positioning over several years.

2.7.2 Initial reference group selection

As there is no quantitative comparison between groups in SPC like in commonly applied statistical analysis, a formal sample size calculation was not applicable. In SPC, the reference group represents a group of observations on which to calculate the control limits to be applied for late observations. These control limits only serve to monitor the (future) stability of the process and indicate the need for further investigations of discrepancies. If process results do not become available at a high frequency (like in this study), it is common to start monitoring based on a pragmatically chosen reference dataset in which the 'process' results seem to be stable. The first results from the observation period are then used to judge the

appropriateness of this division. A reference group of 15 individual datasets was deemed sufficient for SPC which were selected using a grading system, as described below

Starting from the final dataset, patient images were visually graded by the researcher and an independent reviewer. A score of one to five was given to each patient, with a score of one representing minor visual setup deformations and a score of five representing major visual setup deformations. From this, the average visual deformation score was reviewed in each grouping of 15 consecutive patients. This initial review of the data indicated that the positioning results for the 15 most recent patients, treated at the end of 2013 and start of 2014, were better compared to the previous patients. Therefore, the entire cohort was divided into two subgroups labelled as the study group and reference group, respectively. It should be noted that this selection method needed verification. Therefore, during the quantitative analysis of the results, further assessment was made whether this initial division of the data was meaningful, and the reference group actually represented a period with improved and consistent patient positioning results.

2.7.3 Time scale calculation

The standard approach of SPC is to monitor processes by plotting data as it becomes available [41,42]. However, to retrospectively analyse the historical data in this study, a reversed time base was used. Therefore, using the ref_date_i for each patient dataset p, the number of days before the reference date of the most recent patient dataset (ref_date_{i96}) was used as the SPC time base, and was calculated using:

$$Time\ to\ last\ reference\ data\ point_i = ref\_date_{i96} - ref\_date_i \quad \textbf{Eq 11}$$

2.7.4 Control chart calculation

The patient systematic (m_p^{ROI}) and random (σ_p^{ROI}) 3D deformation vectors per ROI, as well as the overall 3D deformation vectors (v_p^m, v_p^σ) were used as the input data to generate the *x*-, *mR*- and *EWMA*-SPC charts.

Individual value and Moving Range

x-SPC charts plot one-at-a-time values, while the *mR*-SPC charts the difference between subsequent individual measurements [42,43]. As indicate above, patient systematic, random and overall deformation vectors were used as the individual values (*x*) for this analysis. The *mR*-SPC charts were generated by calculating the difference between adjacent individual values *x* using [41]:

$$mR_i = x_i - x_{i-1} \quad \textbf{Eq 12}$$

Process limits for both the *x*-SPC and *mR*-SPC charts were calculated based on the reference group, using the following equations [41].

For the *x*-SPC chart;

$$UPL = \overline{x} + 3\,\frac{\overline{mR}}{d_2} \quad \textbf{Eq 13}$$

$$CL = \overline{x} \quad \textbf{Eq 14}$$

$$LPL = \overline{x} - 3\,\frac{\overline{mR}}{d_2} \quad \textbf{Eq 15}$$

For the moving range chart;

$$UPL = \left(1 + 3\,\frac{d_3}{d_2}\right)\overline{mR}$$

Eq 16

$$CL = \overline{mR}$$

Eq 17

$$LPL = \left(1 - 3\,\frac{d_3}{d_2}\right)\overline{mR}$$

Eq 18

Where $\overline{x}$ equals the mean individual value and $\overline{mR}$ the mean moving range for the reference group. Bias correction factors $d_2 = 1.128$ and $d_3 = 0.8525$ for subgroup size $n = 2$ were used to calculate control charts, which is a standard approach for subgroup size $n = 1$ patient [43].

Exponentially Weighted Moving Average

The *EWMA* value is defined as [41,46];

$$EWMA_i = \lambda x_i + (1 - \lambda)EWMA_{i-1}$$

Eq 19

Corresponding process limits were then calculated as [41,46];

$$UPL = \overline{EWMA} + L\sigma\sqrt{\frac{\lambda}{(2\ -\ \lambda)}\,[1 - (1 -\ \lambda)^{2i}]}$$

Eq 20

$$CL = \overline{EWMA}$$

Eq 21

$$LPL = \overline{EWMA} - L\sigma\sqrt{\frac{\lambda}{(2\ -\ \lambda)}\ [1 - (1 -\ \lambda)^{2i}]}$$

Eq 22

Where;

- $EWMA_0$ is taken equal to the CL value for *x*-SPC chart, $\overline{x}$
- *x* is the individual value
- *i* is the patient number
- $\overline{EWMA}$ is the mean *EWMA* for the reference group
- σ is the standard deviation of the individual values for the reference group

λ is a constant that determines the depth of memory of the *EWMA*, where a value of 1 includes only the current value while smaller values of λ give more weight to the inclusion of previous data points [41]. Note that usually this implies older data but for this retrospective analysis, these are the more recent data points. *L* determines the width of the control limits [41,46]. Common values for λ and *L* are 0.2 and 3, respectively [41]. However, as the overall systematic and random patient deformations in this study were not normally distributed (Figure 3-2), λ and *L* were set to λ = 0.05 and *L* = 2.492. Borror *et al.* [46] have shown that, even for strongly asymmetric and skewed data distributions, these λ and *L* values yield virtually the same type I and type II error probabilities as the more common λ and *L* values for normally distributed data.

2.7.5 Rules of interpretation for SPC

The Western Electric Company classified rules for detecting an out-of-control signal in SPC based on three zones [43]. Figure 2-5 displays each zone, while Figure 2-6 demonstrates each of the zone rules.

The zone rules B-D are less commonly used due to the increased risk of returning a false positive [43]. Furthermore, as *EWMA*-SPC charts filter out the specific trends per definition due to the averaging, only one rule of interpretation was applied in this study, that is any value outside the process limits needs to be investigated [46].

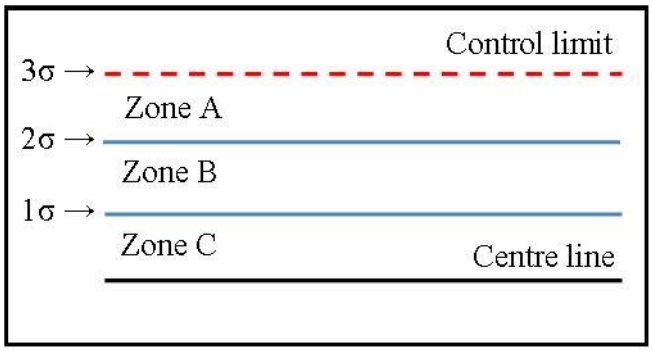


Figure 2-5 Western Electric Zones for detecting out-of-control signals. Zone A located between 3σ (control limit) and 2σ; Zone B located between 2σ and 1σ; Zone C located between 1σ and the CL. Adapted from Wheeler & Chambers [43]

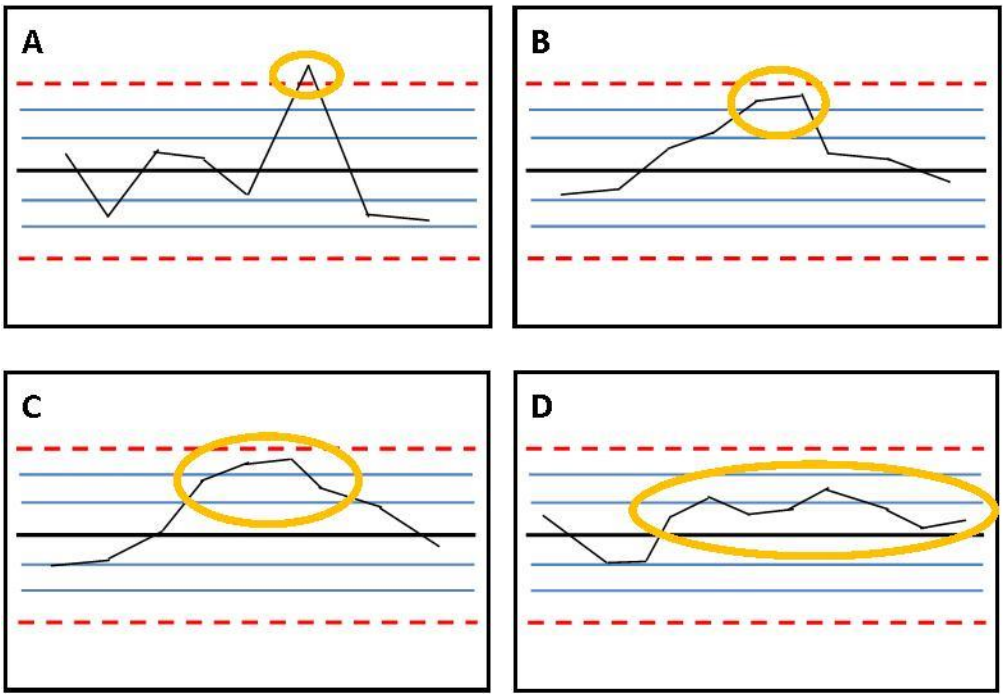


Figure 2-6 Western Electric Zone Rules for detecting an out-of-control signal in SPC analysis. A: Test 1 – any point outside the control limits; B: Test 2 – two out of three successive points fall in Zone A or beyond; C: Test 3 – four out of five successive points fall in Zone C or beyond; D: Test 4 – eight consecutive points on one side of the CL. Investigation points indicated by orange circles. Adapted from Wheeler & Chambers [43]

2.8 Investigation of potential causes of change

A systematic review of departmental process changes that occurred during the study period was performed to identify potential causes of change in setup accuracy. This review was primarily based on previous MDT meeting minutes and edits to departmental protocols, as well as a review of patient treatment information, with three main focus points:

A. Intentional process changes implemented by the MDT for patient positioning.
    - Instigation of weekly image review meetings.
    - Increased documentation and photographs required at the time of mask making and pCT, in order to provide more detailed information for setup at treatment. Included in this change were adjustments in the measurement of the anterior reference setup mark, as well as removal of the use of shoulder retractors at treatment (Section 2.2).
    - Review of IGRT protocols, which resulted in a shift to an online imaging protocol and increased utilisation of CBCT (IGRT Protocol 3, IGRT Protocol 4 as discussed in Section 2.3), as well as defining site specific matching instructions and standardised image review techniques.
    - Review of the head supports being used for HNC treatment. This resulted in the introduction of new head supports, as well as some additions/alterations to the existing head supports as discussed in Section 2.2.

- Re-training for mask making, where tutorials on fitting head supports and mask making were provided to all RTs over a roll out period of approximately 4 months.

B. Wider departmental process changes potentially affecting the positioning accuracy during HNC treatments.
    - Introduction of VMAT for HNC treatment
    - Increased patient care to prevent weight loss

C. Additional factors, not specifically associated with process changes in patient positioning.
    - Chemotherapy – as concurrent cisplatin, nil or other chemotherapy pre radiation therapy, including R/CHOP and RBVD.
    - Surgery – as major, minor or nil surgery pre radiation therapy.
    - Treatment indication – as larynx, oral cavity, pharynx, skin or other.
    - Primary disease site – as upper, middle or lower head-and-neck.
    - Total dose – as D <=54Gy, 54 < D <= 66Gy and D > 66Gy.
    - Number of treatment fractions – as F <= 20, 20 < F <=30 and F > 30.
    - Instructions for head position – as neutral, extended or chin down.
    - Other immobilisation using mouth piece – as bite block, tongue depressor or nil.

The factors grouped under point C represent potential variables that may affect patient positioning, where no specific or intended process change had occurred. These factors were investigated to highlight unintended process changes that may have impacted patient positioning. It should be noted there is likely some overlap between those variables listed under point C, where for example individual treatment indications may show different treatment sites, chemotherapy or surgical approach and total

treatment dose and fractionation. This could not be accounted for, due to the retrospective nature of this study and difficulties in selecting defined groups based on clinical information.

Changes or upgrades to equipment for both pCT and CBCT imaging were not investigated as potential cause for the observed changes in patient positioning. This exclusion was based on the registration protocol employed for this study, where bony anatomy was used as the match structure (Section 2.4). Typically, limitations in registration of pCT/CBCT imaging relate to soft tissue matching due to poor contrast, while bony anatomy is generally visible across a range of image quality levels [61]. Furthermore, all acquisition parameters for pCT and CBCT are determined using fixed protocols, and a tight quality assurance programme is routinely applied for both modalities at WBCC to warrant the consistency of imaging quality. It was therefore concluded that image quality was not a limiting factor for the accuracy of the image registration applied in this study, and equipment changes / upgrades were excluded from further investigations.

Mostly, those factors listed as potential causes of changes are categorical variables, which were reviewed using time plots as discussed in Section 2.8.2. The only numerical variable of weight loss was reviewed separately, as discussed in Section 2.8.1.

2.8.1 Weight loss

A number of changes were made within the department relating to patient care plans and the level of input from nursing, dietician and speech language therapists during the study period. For that purpose, the total weight loss per patient (WL_p) was used as the endpoint of these changes, and was expressed as the percentage change in total

body weight at end of treatment (w_p^{EOT}) relative to the total body weight at pCT (w_p^{pCT}):

$$WL_p = \frac{w_p^{EOT} - w_p^{pCT}}{w_p^{pCT}} \times 100\%$$

Eq 23

SPC was used to review time trends in weight loss and linear regression was applied to correlate weight loss and changes in patient positioning. SPC charts were calculated using the same methods outlined in Section 2.7, including the same reverse time base and reference group.

2.8.2 Time plots for categorical variables

The remaining categorical variables were reviewed using time plots, to verify whether changes in positioning accuracy coincided with process changes or altered additional factors. Time plots were created using the same time scale as the SPC charts (Section 2.7.3). This initial crude analysis was used to manage the workload of this study and exclude factors from further analysis that did not show changes over time. Factors which could not be excluded in this way were then analysed further by testing the statistical significance of correlations.

2.8.3 Kruskal-Wallis ANOVA

Kruskal-Wallis was used to test the statistical significance between categorical variables and overall patient systematic and random deformation. Kruskal-Wallis is a non-parametric equivalent to analysis of variance (ANOVA) for parametric variables [62]. It is used where there are more than two independent subgroups, the assumptions of normality are not met with the measurement variable and the

subgroups have an unequal variance [62,63]. Each measurement variable is substituted for a rank and the distributions of ranks are then compared to the distribution of ranks of the whole population [62]. The test statistic H was calculated using MATLAB (The Mathworks, Natick, MA, USA), with a p-value <0.05 as a threshold for significance. Post-hoc comparisons were made using a Dunn-Sidak correction for multiple corrections [64].

2.9 Margin calculation

Using the group setup uncertainties, the contribution of deformation to the overall PTV margins was calculated by applying the margin recipe described by van Herk *et al.* [65] for rigid body setup inaccuracies;

$$m = 2.5\,\Sigma + 0.7\sigma$$

Eq 24

For application in this study, Σ represents the group systematic dispersion per ROI, Σ_{ROI}, and σ the group random deformation per ROI, σ_{ROI}. As the analysis results are based on deformation relative to C1-C3, these PTV margin calculations only apply when patients are positioned using C1-C3 as match structure. Therefore, the results from these calculations should not be used as indication of the required PTV margin for setup errors but are only intended to give an indication of the contribution of deformation to the overall PTV margin.

3 RESULTS

3.1 Multiple rigid registration protocol

As this was a retrospective study based on available clinical data, a large variation in the number of patient datasets (P_{ROI}) and images (N_{img}^{ROI}) per ROI was observed (Table 3-1). This variance was due to the variation in treatment indications and the finite scan length of the CBCT, which is typically centred on the isocentre [30]. Based on the low patient and image numbers for the suprasternal notch (SSN), this ROI was excluded from analysis.

Table 3-1 Number of patient datasets (P_{ROI}) and images (N_{img}^{ROI}) per ROI

ROI	P_{ROI}	N_{img}^{ROI}
C1-C3	101	740
C3-C5	101	740
C5-C7	92	645
C7-caudal	74	524
Mandible	97	697
SSN	9	43
Occipital bone	85	613
Larynx	97	699

Figure 3-1 visually demonstrates the definition of various ROIs in the multiple rigid registration protocol employed in this study, where each image represents an individual ROI match. As shown by Figure 3-1, deformation was observed for this particular patient with all registrations, irrespective of the selected ROI match. While this represents one of 740 images, the results presented in Figure 3-1 were common throughout the study cohort, with varying levels of deformation. In order to quantify the observed deformation, the displacement of individual ROIs relative to C1-C3 was calculated.

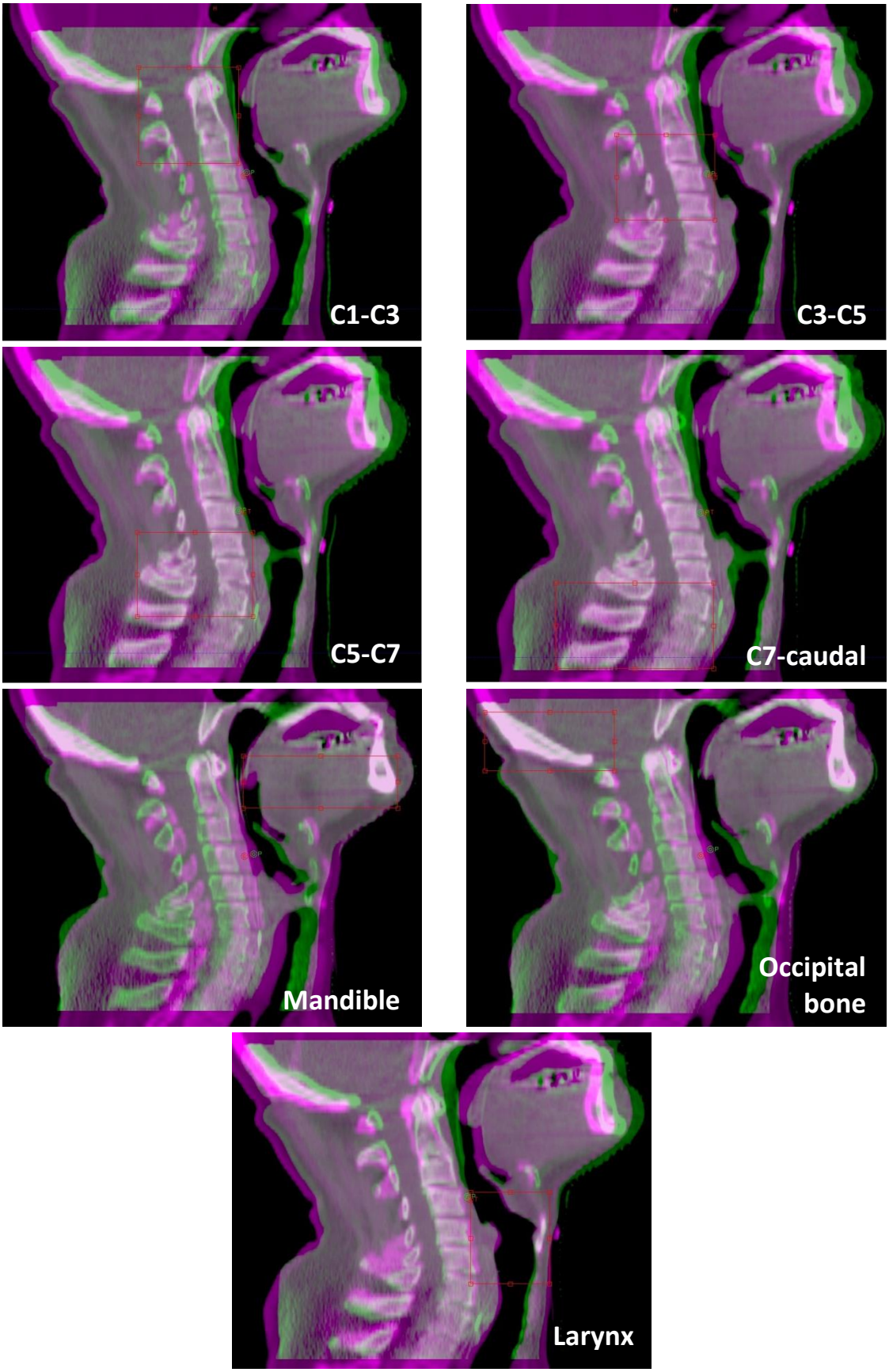


Figure 3-1 Definition of various ROIs in the multiple rigid registration protocol. The colour blend superimposes the pCT (purple) and CBCT (green) images and highlights regions where a good match (white) or a mismatch between the reference and treatment images is obtained (either green or purple).

3.2 Setup uncertainties

Deformation was calculated for each translational axis and associated 3D vectors using Eq 1 and Eq 2. The deformation statistics per patient m_p^{ROI} and σ_p^{ROI} were then calculated using Eq 3 and Eq 4. The overall systematic deformation (v_p^m) and random deformation (v_p^σ) were also calculated per patient, using Eq 5 and Eq 6. These overall deformation numbers do not present a clinically relevant quantity, but were merely a quality measure that provided a way to analyse the large dataset using SPC (Section 3.4).

Subsequently, the group mean, $M_{ROI,}$ group dispersion of the systematic deformation, Σ_{ROI}, and the group random deformation, σ_{ROI}, were calculated using equations 7-10 (Table 3-2). For readability purposes, the group dispersion of the systematic deformation will be referred to as the systematic dispersion in the remainder of this thesis.

Patients whose data formed the basis of this study were divided into three separate cohorts. The entire study cohort (WBCC overall; n = 101), patients who started treatment in 2011/2012 where positioning was observed to require improvement (WBCC 2011/2012; n = 41) and patients who were selected as the reference group for SPC analysis, as discussed in Section 2.7.2 (WBCC reference group; n = 15).

Table 3-2 Group mean, systematic dispersion and group random deformation relative to C1-C3 or similar structure*

M_{ROI} [mm]	C3-C5				C5-C7				C7-caudal				Mandible				Occipital Bone				Larynx			
	Vrt	**Lng**	**Lat**	**3D**	**Vrt**	**Lng**	**Lat**	**3D**	**Vrt**	**Lng**	**Lat**	**3D**	**Vrt**	**Lng**	**Lat**	**3D**	**Vrt**	**Lng**	**Lat**	**3D**	**Vrt**	**Lng**	**Lat**	**3D**
WBCC overall	-0.3	-0.1	0.0	1.6	-0.2	-0.1	-0.2	2.3	0.4	0.1	-0.3	2.9	1.0	-0.7	-0.2	2.9	-0.4	0.3	0.2	2.2	0.0	0.4	0.1	3.2
WBCC 2011/2012	-0.5	-0.2	0.1	1.6	-0.3	0.0	0.1	2.4	0.5	0.3	0.0	3.1	1.6	-0.6	-0.4	3.3	0.2	0.0	0.0	2.2	0.1	0.9	0.1	3.2
WBCC reference group	0.1	0.0	0.0	1.2	0.7	0.1	-0.1	1.9	2.0	0.7	-0.5	3.4	0.4	0.1	0.1	2.3	-0.2	0.3	0.2	2.0	0.6	0.2	0.2	3.0
Polat *et al.*, [12] 2007	-	-	-	3.0	-	-	-	3.0	-	-	-	-	-	-	-	3.2	-	-	-	2.7	-	-	-	-
van Kranen *et al.*, [11] 2009	0.1	0.1	0.1	-	0.2	-0.2	0.4	-	0.8	0.5	0.9	-	0.3	1.3	0.0	-	0.3	0.6	0.4	-	0.2	-0.3	0.5	-
Giske *et al.*, [13] 2011	-	-	-	-	1.2	0.3	0.1	-	2.0	0.3	0.0	-	0.3	-0.6	0.0	-	-	-	-	-	1.7	-0.5	0.2	-
Graff *et al.*, [14] 2013	-	-	-	-	-	-	-	-	-	-	-	2.7	-	-	-	4.0	-	-	-	-	-	-	-	-

Σ_{ROI} [mm]	C3-C5				C5-C7				C7-caudal				Mandible				Occipital Bone				Larynx			
	Vrt	**Lng**	**Lat**	**3D**	**Vrt**	**Lng**	**Lat**	**3D**	**Vrt**	**Lng**	**Lat**	**3D**	**Vrt**	**Lng**	**Lat**	**3D**	**Vrt**	**Lng**	**Lat**	**3D**	**Vrt**	**Lng**	**Lat**	**3D**
WBCC overall	1.0	0.3	0.6	0.6	1.4	0.7	1.0	0.8	1.8	0.9	1.3	1.1	1.5	1.8	0.7	1.6	1.3	0.8	0.7	0.7	1.3	3.8	1.0	2.3
WBCC 2011/2012	0.6	0.2	0.3	0.3	0.8	0.5	0.6	0.3	1.1	0.7	0.9	0.6	1.0	0.9	0.4	0.7	0.8	0.4	0.4	0.1	0.7	2.1	0.6	1.2
WBCC reference group	0.2	0.1	0.2	0.2	0.4	0.2	0.3	0.2	1.0	0.4	0.6	0.5	0.4	0.6	0.3	0.3	0.2	0.3	0.2	0.1	0.4	0.6	0.3	0.6
Polat *et al.*, [12] 2007	-	-	-	2.1	-	-	-	2.1	-	-	-	-	-	-	-	2.0	-	-	-	1.8	-	-	-	-
van Kranen *et al.*, [11] 2009	0.5	0.4	0.6	-	1.2	1.4	1.6	-	3.2	2.2	3.3	-	1.2	2.5	1.2	-	0.6	2.0	1.0	-	1.1	3.8	1.5	-
Giske *et al.*, [13] 2011	-	-	-	-	2.2	0.7	1.6	-	2.5	1.6	3.5	-	1.3	1.6	0.6	-	-	-	-	-	2.1	1.9	1.1	-
Graff *et al.*, [14] 2013	-	-	-	-	-	-	-	-	-	-	-	1.2	-	-	-	2.7	-	-	-	-	-	-	-	-

σ_{ROI} [mm]	C3-C5				C5-C7				C7-caudal				Mandible				Occipital Bone				Larynx			
	Vrt	**Lng**	**Lat**	**3D**	**Vrt**	**Lng**	**Lat**	**3D**	**Vrt**	**Lng**	**Lat**	**3D**	**Vrt**	**Lng**	**Lat**	**3D**	**Vrt**	**Lng**	**Lat**	**3D**	**Vrt**	**Lng**	**Lat**	**3D**
WBCC overall	1.1	0.6	0.8	0.9	1.3	0.8	1.0	1.0	1.4	1.1	1.5	1.2	1.3	1.5	1.0	1.3	1.3	0.9	1.0	1.1	1.1	2.9	1.1	2.3
WBCC 2011/2012	0.6	0.4	0.4	0.5	0.8	0.5	0.6	0.6	0.9	0.7	1.0	0.8	0.8	1.0	0.6	0.8	0.9	0.5	0.6	0.7	0.7	1.5	0.7	1.1
WBCC reference group	0.4	0.2	0.3	0.3	0.5	0.3	0.4	0.3	0.5	0.4	0.7	0.5	0.4	0.4	0.3	0.4	0.4	0.2	0.3	0.3	0.4	1.3	0.5	1.0
Polat *et al.*, [12] 2007	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
van Kranen *et al.*, [11] 2009	0.7	0.5	0.8	-	1.5	1.7	1.4	-	3.3	2.2	2.8	-	1.2	2.3	1.3	-	0.7	1.6	1.1	-	1.5	2.5	1.5	-
Giske *et al.*, [13] 2011	-	-	-	-	2.3	0.7	1.1	-	2.7	1.4	3.2	-	1.1	1.4	0.6	-	-	-	-	-	1.7	1.6	0.9	-
Graff *et al.*, [14] 2013	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

***** The ROIs used in this study match those used by van Kranen *et al.* [11], however other studies had used varying sub regions, which were compared to the ROIs that most closely relate in this study. Polat *et al.* [12] matched on C4-C6 and the whole skull, which were compared to the results for C3-C5/C5-C7 and the occipital bone respectively. Giske *et al.* [13] and Graff *et al.* [14] both used C1-C2 as the reference structure for the calculation of deformation, as well as included smaller ROIs compared to this study. The results Giske *et al.* [13] presented for C6 and T2 were compared to C5-C7 and C7-caudal respectively. Graff *et al.* [14] included C7-T1, which was also compared to C7-caudal. All other structures included in Table 3-2 match the ROI listed in the column header.

The results in Table 3-2 showed that the group mean deformation for all WBCC cohorts was smaller than 1 mm along nearly all Cartesian axes, indicating that there is no significant systematic departmental error. Furthermore, the level of deformation was smaller for vertebrae close to C1-C3, and increased for vertebrae located further away. The larynx and mandible are relatively mobile structures, the motion of which could probably not be fully prevented. Consequently, these ROIs displayed the largest group deformation statistics, in particular the systematic dispersion of the larynx in the lng direction.

The initial patient cohort (2011/2012) demonstrated higher deformation values for almost all ROIs, in comparison with the reference group. This is consistent with the improvements in setup accuracy observed during weekly image review meetings, however, time trends during this period of change are generally not captured by descriptive statistics [38]. Therefore, a review of time trends was done using SPC (Section 3.4).

In comparison to previous literature, the results of this study were comparable if not better for most ROIs. Further discussion around benchmarking and comparisons with the literature is detailed in Section 4.1.2.

3.3 Normal probability test

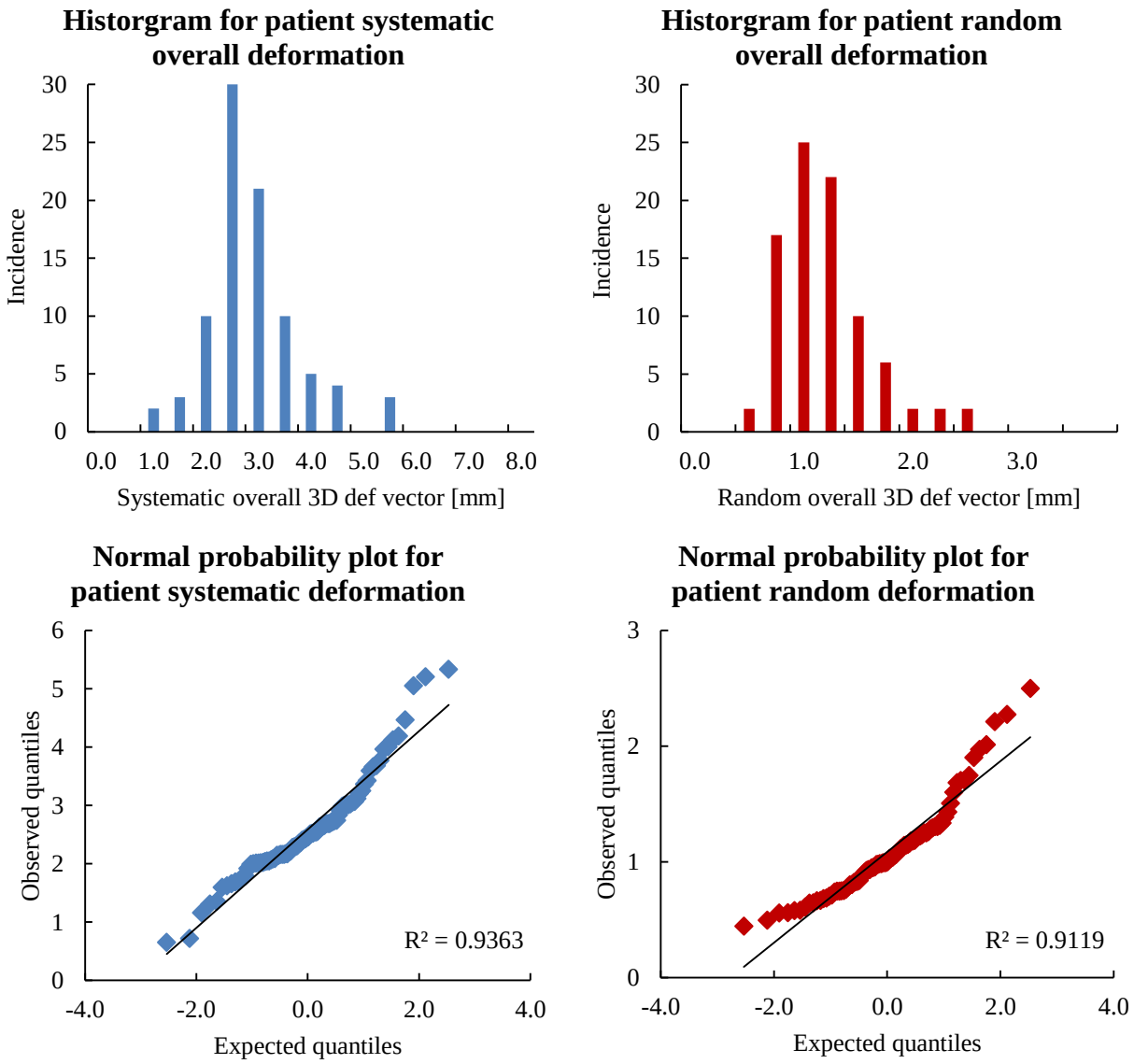


Figure 3-2 Histograms and normal probability plots for systematic and random overall deformation 3D vectors.

Histograms (Figure 3-2) for patient systematic and random overall deformation both demonstrated right skewed distributions. Additionally, several departures from the fit line in the normal probability plots were observed (Figure 3-2), indicating departures from a normal distribution within this dataset. Along with histograms and normal probability plots, a Shapiro-Wilk test [60] was performed. The results for patient

systematic (W = 0.9347, p = 0.0005) and random (W = 0.9112, p = 0.00005) deformation both significantly rejected the null hypothesis, confirming the data were not normally distributed.

3.4 Statistical Process Control

SPC charts were generated using the 3D overall deformation vectors relative to C1-C3 as the key quality measure for primary analysis, as well as the 3D deformation vectors for each ROI. Separate SPC charts for systematic and random deformation were created.

As per the generic SPC chart presented in Section 1.5, the reference group of 15 patients is plotted in red on all SPC charts, with the study group shown in black. The centre line is charted in black, with upper and lower process limits as dotted red lines.

As discussed in Section 2.7.3, a reverse time base was applied on the *x*-axis to maintain the common layout of the SPC charts, with the reference group at the left-hand side and the observation period at the right-hand side while performing a retrospective study. Thus, the time axis of all SPC charts in this thesis run from right to left, where t=0 represents the last patient included in the reference group who started treatment in 2014, and t= -1000 days relates to data from 2011.

3.4.1 Individual value and moving range charts

Figure 3-3 and Figure 3-4 show the individual value (*x*-SPC) and moving range (*mR*-SPC) charts for overall deformation. *x*-SPC and *mR*-SPC charts for each ROI are available in Appendix A and Appendix B, respectively. The *x*-SPC and *mR*-SPC charts for both patient systematic and random overall deformation demonstrated a decrease in deformation over a long time with multiple data points outside the control limits during the study, indicating several exceptional deviations.

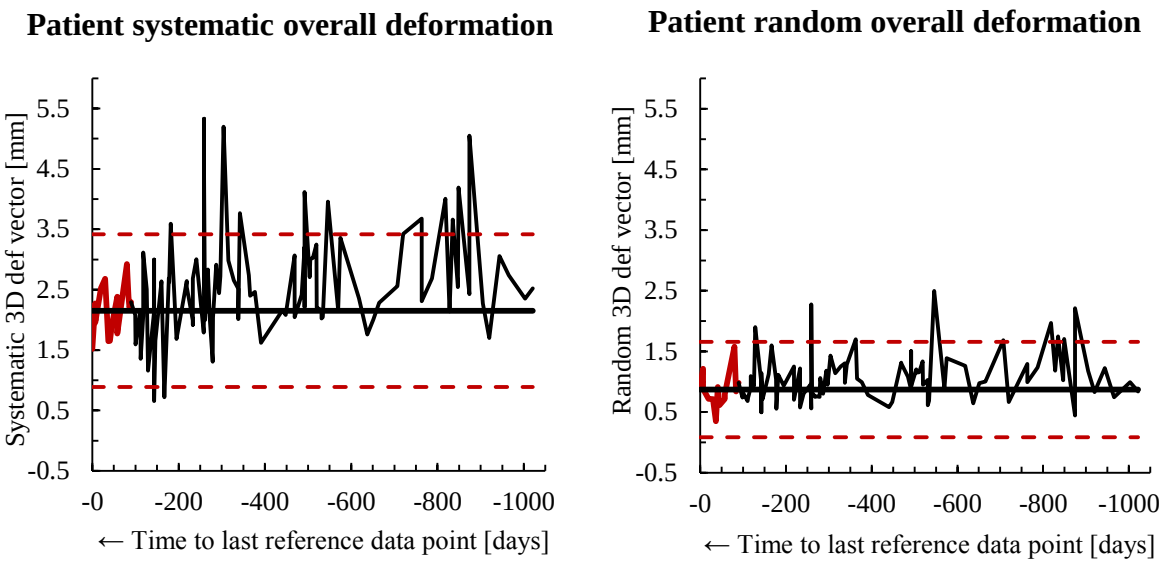


Figure 3-3 *x*-SPC charts for patient systematic and random overall deformation relative to C1-C3. Reference group plotted in red, study group shown in black. Centre line charted in black; upper and lower process limits shown as dotted red lines.

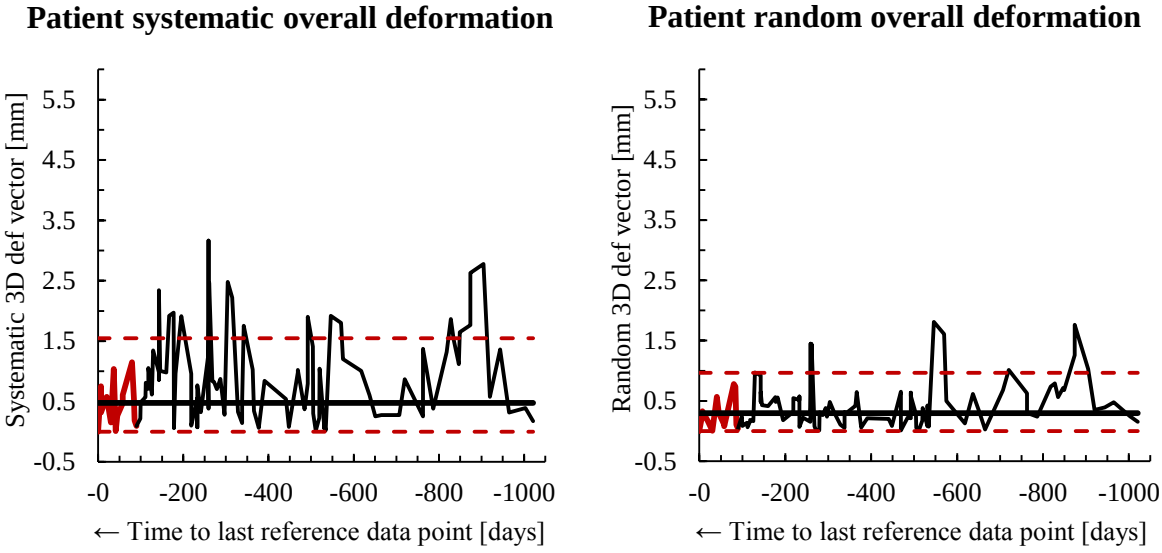


Figure 3-4 *mR*-SPC charts for patient systematic and random overall deformation relative to C1-C3. See Figure 3-3 for explanation of the charts.

The interpretation of these results must take into account the known limitations of *x*- and *mR*-SPC charts [46]. The process limits for these SPC charts are based on the assumption of normally distributed data, which patient systematic and random overall deformation did not meet, as demonstrated by Figure 3-2. Furthermore, this retrospective study aimed to highlight persistent changes in patient positioning over a long time period, which *x*- and *mR*-SPC charts are not very sensitive in identifying [41,42]. These factors indicated that *x*- and *mR*-SPC charts were not ideally suited for analysing the data in this study. In contrast, *EWMA*-SPC charts are better suited to detect small shifts in the average process results and do not require normally distributed data [46]. The application of *EWMA*-SPC charts seemed more appropriate for this dataset, and these charts were therefore used as the primary tool to retrospectively detect specific improvements in patient positioning over time.

3.4.2 Exponentially Weighted Moving Average charts

Figure 3-5 shows the *EWMA*-SPC charts for patient overall systematic and random deformations.

These charts showed an improvement in setup accuracy, with a reduction in the systematic overall deformation per patient from 2.8 mm ± 0.1 mm (1 S.D) in 2011 to 0.9 mm ± 0.0 mm (1 S.D) in 2014. A similar reduction was shown in random overall deformation per patient, from 1.2 mm ± 0.0 mm (1 S.D) in 2011 to 0.9 mm ± 0.0 mm (1 S.D) in 2013/2014. Closer inspection showed that there seemed to be two distinct time periods where a considerable improvement occurred, at t = -250 to -300 days and around t = -800 days.

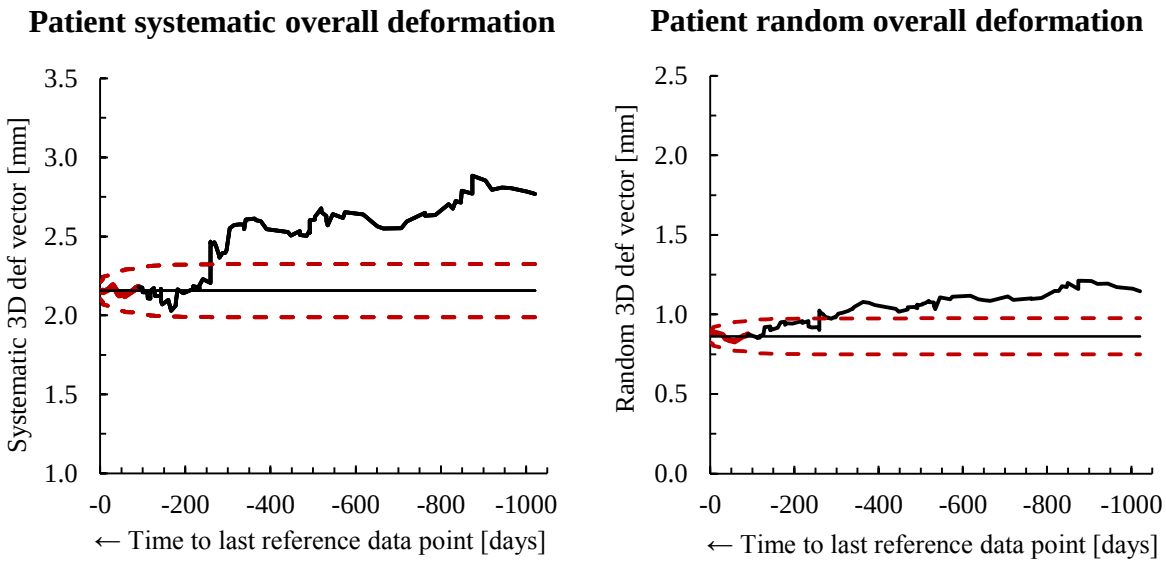


Figure 3-5 *EWMA*-SPC charts for patient systematic and random overall deformation relative to C1-C3. See Figure 3-3 for explanation of the charts. Note that the same scale is used for both panels but a different offset was applied.

The change at t = -250 to -300 days presented a shift outside the control limits on both the patient systematic and random overall deformation *EWMA*-SPC charts. Therefore, the patient positioning results before that point in time could be considered to be significantly different, compared to the reference group. The second shift observed around t = -800 days could not be clearly interpreted based on the original control limits. For this reason, separate SPC charts using the time period between t = -300 to -800 as the reference group were created. This alternate reference group was selected as it demonstrated a plateau in the results, indicating relatively consistent positioning accuracy during this time. The trends in Figure 3-6 show that a smaller process change seemed to have occured around t = -800 days but the results before t = -800 days did not fall outside the process limits. It could not be proven that this shift in setup accuracy was a significant change and was therefore not further investigated.

No other shifts around the control limits were observed in the results, making t = -250 to -300 days the focus point of further analysis and the investigation into potential causes of change, as described in Section 3.5.

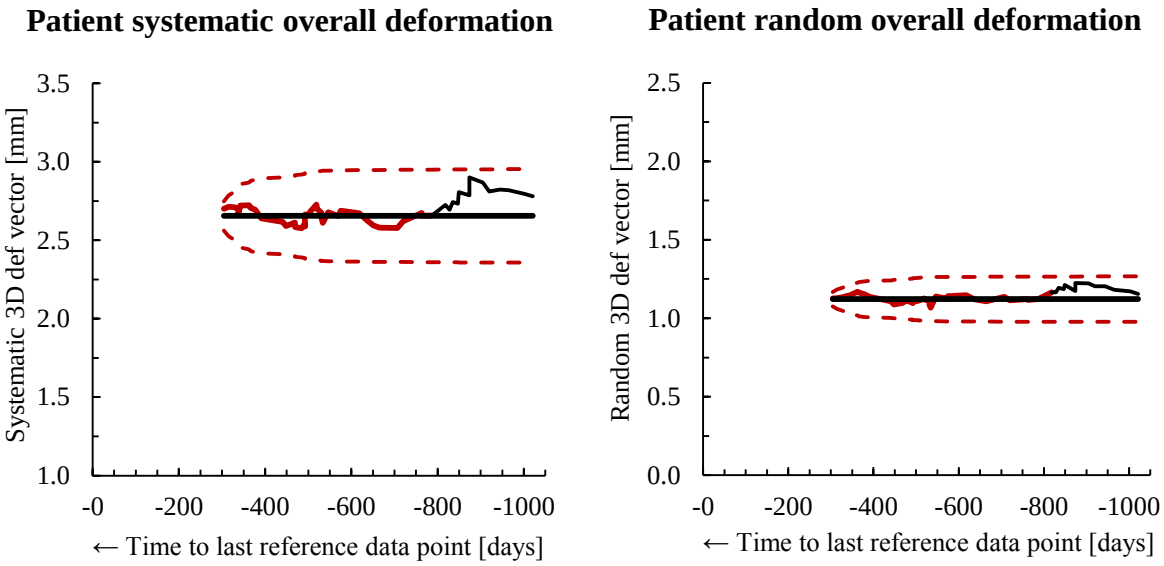


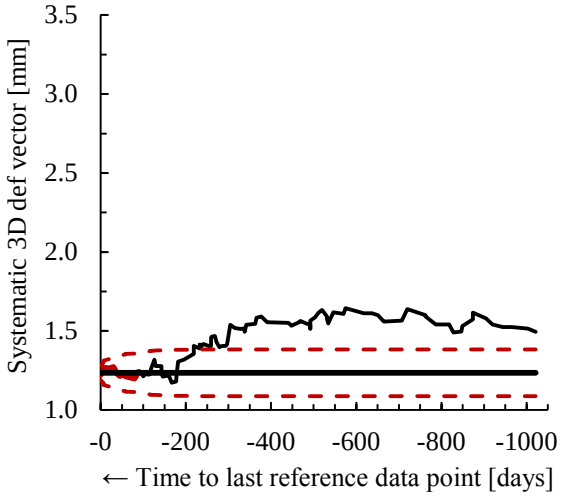
Figure 3-6 *EWMA*-SPC charts for patient systematic and random overall deformation relative to C1-C3, using patients treated before t = -300 days only. See Figure 3-3 for explanation of the charts. Note that the same scale is used for both panels but a different offset was applied.

While overall deformation was used as the primary analysis parameter, *EWMA*-SPC charts for individual ROIs were also reviewed, in Figure 3-7. This provided an indication as to which ROIs demonstrated the most improved setup accuracy, relative to C1-C3.

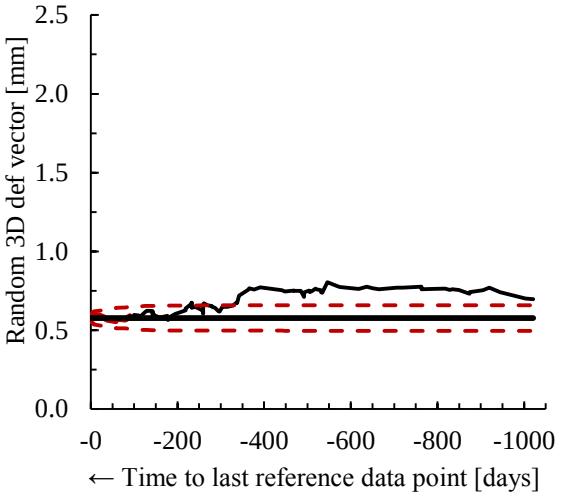
As shown in these results moderate positioning improvements were seen in C3-C5 and C5-C7, for both patient systematic and random deformation. C7-caudal showed no significant improvement in systematic deformation and only moderate improvement in random deformation, while the occipital bone showed no significant

improvement throughout the study period. The greatest reduction in deformation relative to C1-C3 was seen in the systematic deformation for the mandible and larynx, with both of these structures also showing moderate improvement in random deformation. Although not all ROIs show an abrupt shift of the data to fall outside the process limits, the changes that were observed were mostly seen around the same time as the shift outside of the control limits observed on the overall deformation *EWMA*-SPC charts. This further supports the use of t = -250 to -300 days as the analysis reference point, considering both overall and individual ROI deformation SPC charts.

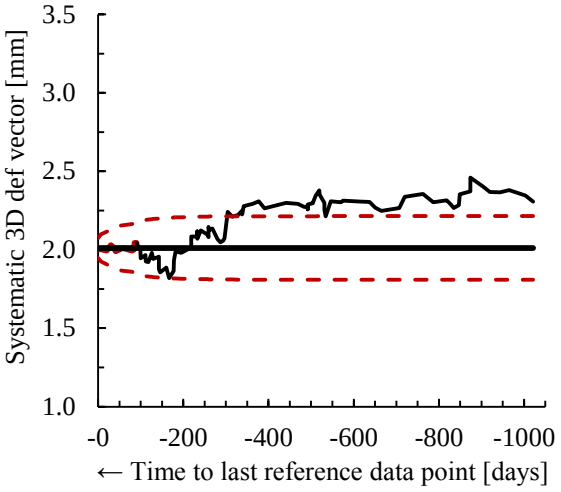
Patient systematic C3-C5 deformation



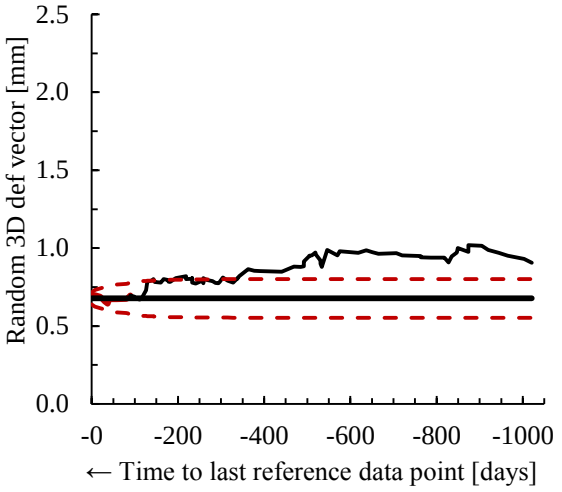
Patient random C3-C5 deformation



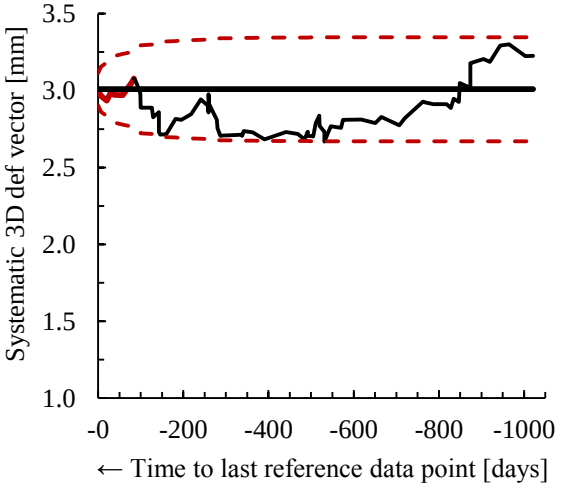
Patient systematic C5-C7 deformation



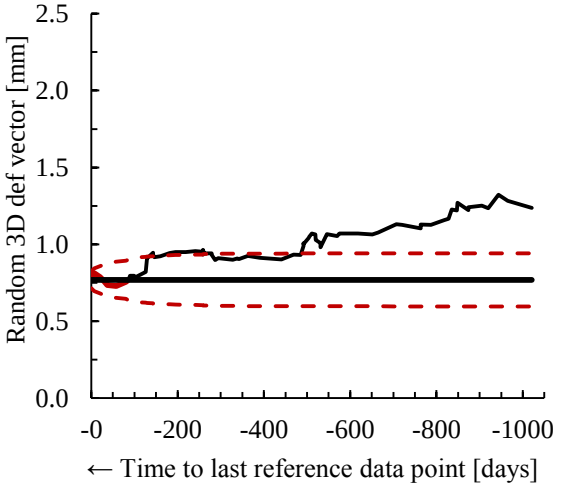
Patient random C5-C7 deformation



Patient systematic C7-caudal deformation



Patient random C7-caudal deformation



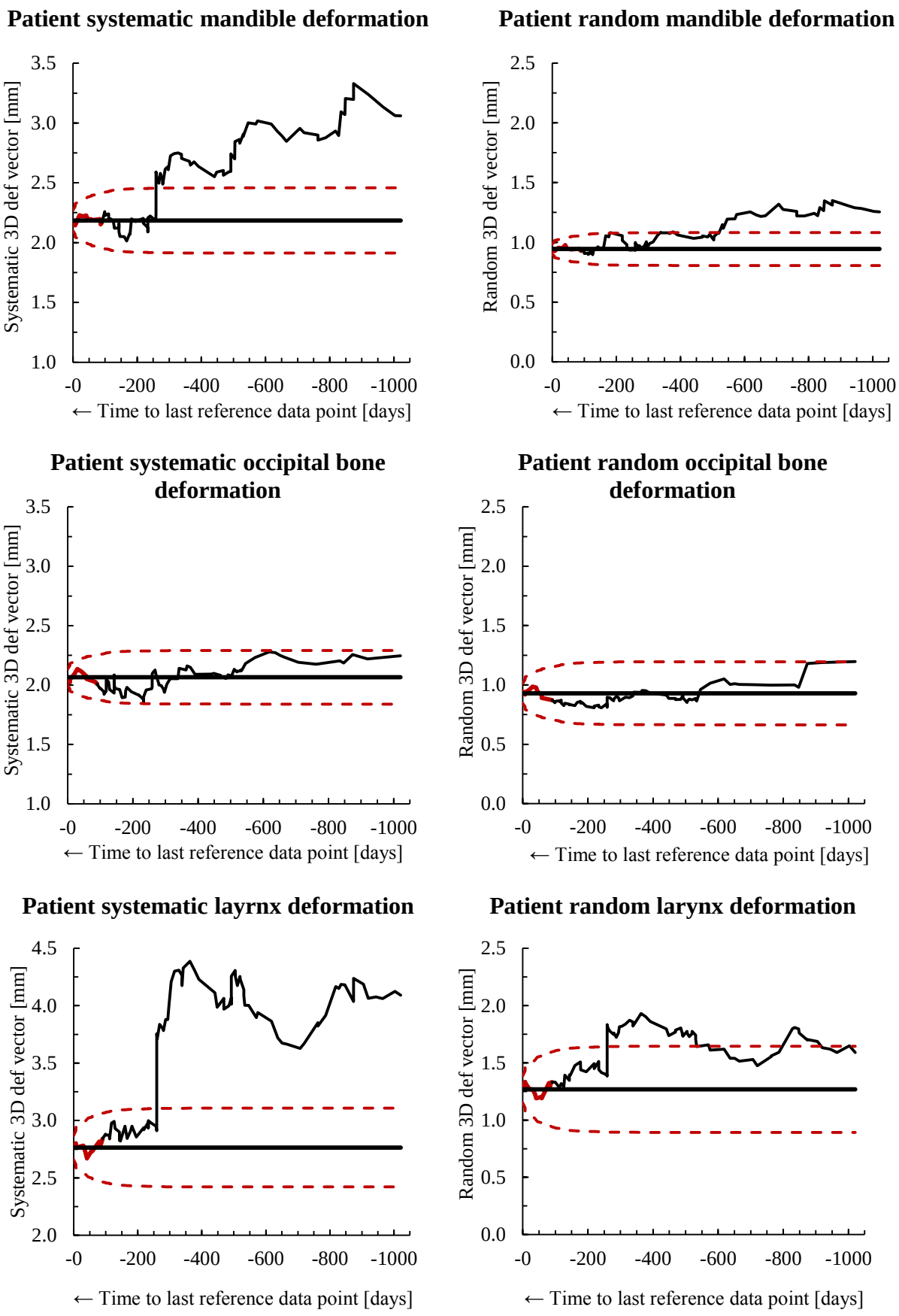


Figure 3-7 *EWMA*-SPC charts for patient systematic and random deformation per ROI relative to C1-C3. See Figure 3-3 for explanation of the charts. Note that the same scale is used for both panels but a different offset was applied.

3.5 Investigation of potential causes of change

As demonstrated by the *EWMA*-SPC charts in the previous section, the only shift in patient systematic and random overall deformation, which could be interpreted as a significant improvement in patient positioning, occurred around t = -250 to -300 days before the last reference data point. This time period therefore became the focus of further analysis.

3.5.1 Weight loss

Weight loss was reviewed as a potential cause of change in setup accuracy, based on departmental process changes around patient care to prevent weight loss. Weight loss was reviewed over the study period using SPC, with charts generated using the same methods as for setup accuracy. Again, the *EWMA*-SPC chart was used as the primary analysis tool, as displayed in Figure 3-8. *x*- and *mR*-SPC for weight loss are available in Appendix C.

Remarkably, this *EWMA*-SPC chart showed a reduction in total weight loss during the study period similar to that observed for setup accuracy. However, the reduction in weight loss occurred around t = -400 days, which was slightly earlier than the improvements in setup accuracy (Figure 3-5). The change in weight loss was also relatively small, from 5.1% ± 0.2% (1 S.D) in the period between t = -400 to -600 days to 3.6% ± 0.1% (1 S.D) in 2014. Furthermore, the weight loss measured represented total body weight and did not specifically relate to the head-and-neck region. Therefore, in order to further quantify the relationship between weight loss and setup accuracy, these two variables were compared using linear regression, as displayed in Figure 3-9.

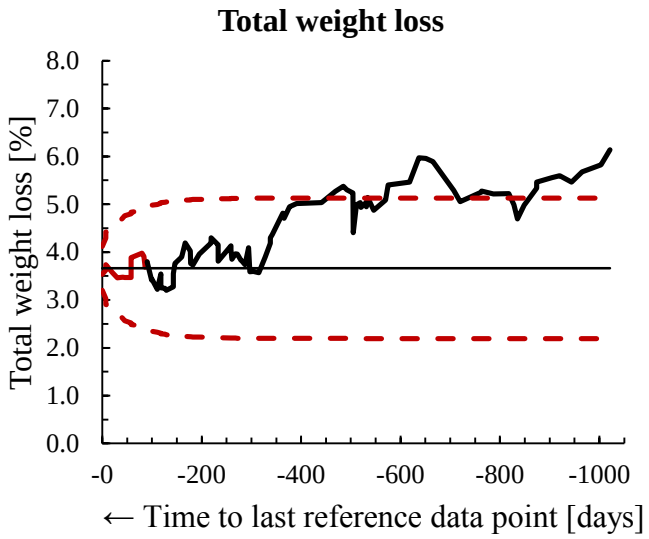


Figure 3-8 *EWMA*-SPC charts for total weight loss. See Figure 3-3 for explanation of chart.

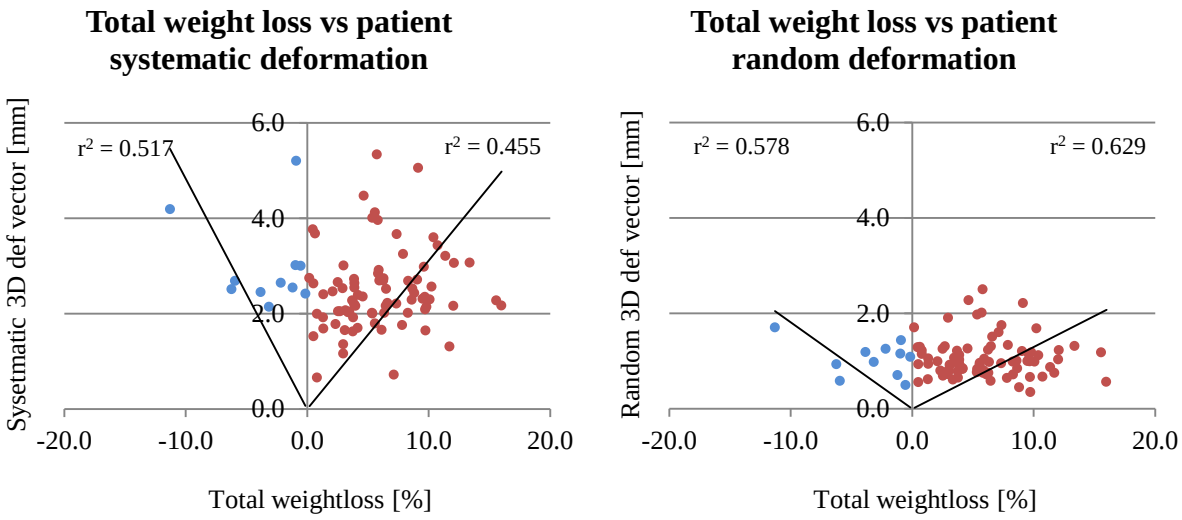


Figure 3-9 Linear regression showing total weight loss versus overall deformation. Weight loss (red dots) and weight gain (blue dots) plotted separately within the same linear regression chart, to demonstrate the individual relationship of these types of weight change on setup accuracy.

The scatter plots attempt to fit a theoretical model to assess the impact of weight loss on setup accuracy. This model excludes changes in setup accuracy without weight loss and therefore, the fitted linear regression line was forced to pass through the origin to avoid changes in setup accuracy from other factors being attributed to weight loss. Both the plots for systematic and random overall deformation showed that there was no significant correlation between setup accuracy and weight loss. This is further supported by the square of the correlation coefficient, *r*, which was calculated to assess the goodness of fit of the linear regression line. r^2 represents the amount of variation in deformation that can be described by its linear relationship with total weight loss, where r^2 = 1 shows a perfect correlation, and r^2 = 0 shows no linear relationship [44]. Based on the results, no significant correlation between change in body weight and setup accuracy could be demonstrated within this study cohort.

3.5.2 Time plots for categorical variables

Time plots were created to highlight potential factors affecting patient positioning, including intentional patient positioning process changes, wider departmental changes and additional factors that may have impacted setup accuracy. These plots chart the incidence of different factors over time, using the same reverse time scale as applied in SPC analysis.

In reviewing all available time plots, charts for IGRT protocol changes, treatment indication, head supports and retraining for immobilisation equipment all showed changes around the analysis reference point of t = -250 to -300 days (Figure 3-10). Other time plots did not show changes coinciding with the reduction in deformation

observed in SPC analysis and were therefore excluded from further analysis (Appendix D).

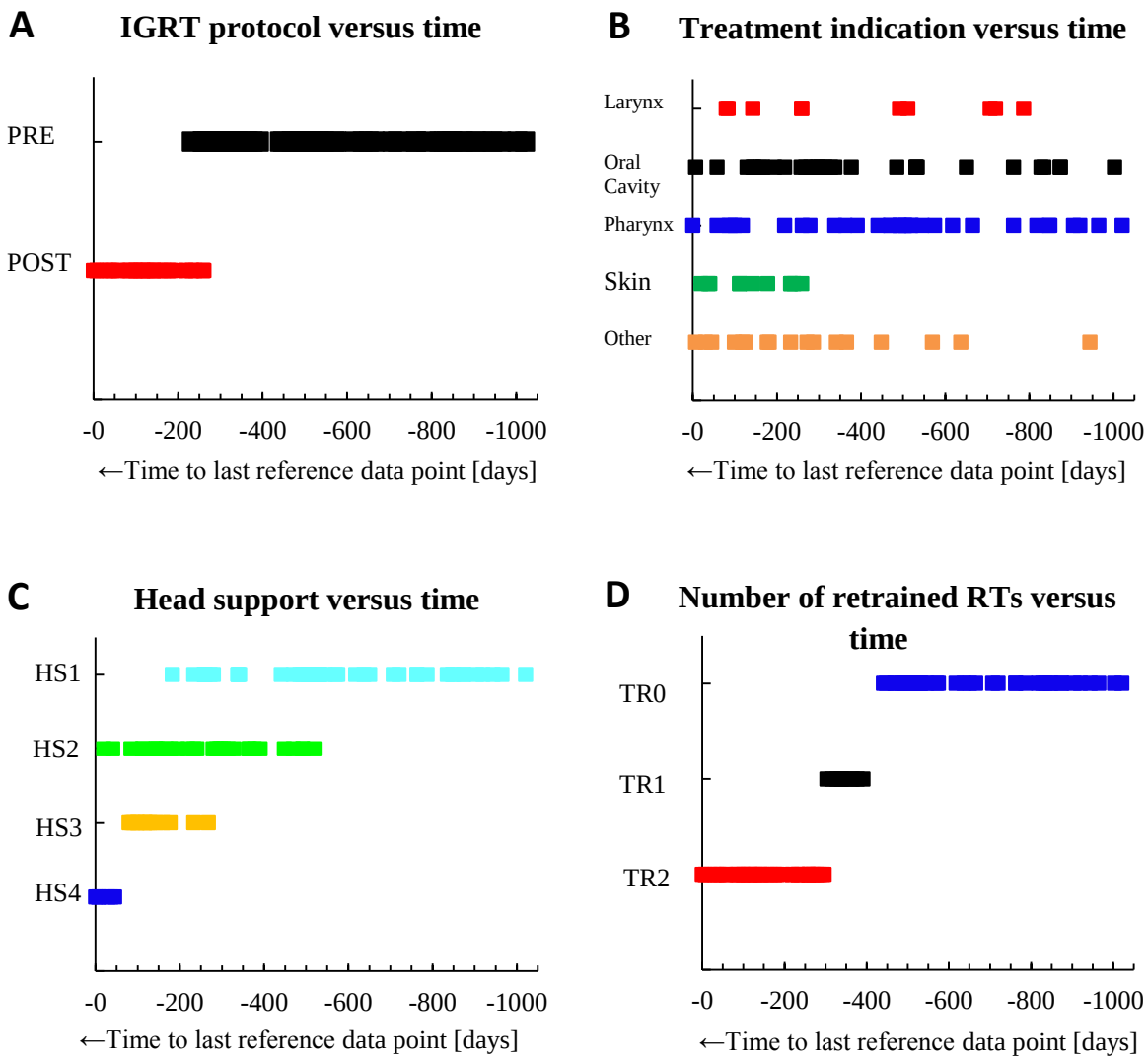


Figure 3-10 Time plots A-D for variables demonstrating changes coinciding with changes in setup accuracy. The *x*-axis represents the same reverse time base used in SPC analysis; the *y*-axis represents the different subgroups for each variable. Each individual point on the time plot represents a single incidence for the relevant subgroup variable, plotted against the reference date for that patient dataset. See text for additional information on each variable and subgrouping.

Changes in IGRT protocol

Figure 3-10-A demonstrates the incidence of patients treated prior to (Pre) and after (Post) the change in IGRT protocol. A change in incidence occurs at approximately t = -250 to -300 days, which coincides with the observed improvement in patient positioning. This change included a shift to online imaging for all patients, as well as standardised image review techniques and patient specific matching instructions, which resulted in a different prioritisation of match structures for clinical positioning corrections. In addition, the application of CBCT was extended to all HNC treatments where previously it was limited to specific treatment techniques. The new IGRT protocol therefore includes the change observed in treatment indication, where skin cancer patients were added to the study cohort, based on the increased utilisation of CBCT across all treatment indications and the selection criteria applied to this study, as patients imaged with CBCT (Section 2.1). However, different prioritisation of match structures and increased imaging frequency do not affect the initial patient setup in any way and can consequently not explain the observed change in deformation. Therefore, this change in IGRT protocol was ruled out as a possible cause of the observed change in positioning accuracy and was not continued into further analysis.

Changes in treatment indication

Changes in treatment indication were demonstrated in Figure 3-10-B. As shown in this time plot, the incidence of larynx, oral cavity, pharynx and other treatment indications was reasonably constant throughout the study period. A change was observed around t = -250 to -300 days, with the addition of skin cancer diagnoses as a

treatment indication. As this constituted an unintended change that may affect setup accuracy, it was included for further analysis.

Changes in head supports

The plot for head support versus time (Figure 3-10-C) demonstrated that four main head supports were used during the study period, indicated as HS1, HS2, HS3 and HS4 (see Section 2.2 for details of the differences between these head rests). Head supports were phased in over time, as demonstrated by the overlapping incidence for each type. HS1 was the only type phased out within the study period, and was used from the start of the study up until t = -180 days. HS2 was introduced around t= -500 days, while the introduction of HS3 seemed to coincide with the improvement observed in setup accuracy at t = -250 to -300 days. The last head support introduced during the study period was HS4, which was used within the reference group for SPC chart calculation. These observations suggest a potential correlation between the observed change in setup accuracy and type of head support within this study cohort.

Retraining for immobilisation equipment

As demonstrated by Figure 3-10-D, data could be divided into three separate patient cohorts in terms of level of retraining; where immobilisation was performed with no retrained RTs (TR0), with a minimum of one retrained RT (TR1) or with a minimum of two retrained RTs (TR2). The results show that also the completion of retraining for immobilisation equipment seemed to coincide with the observed changes in setup accuracy. In particular, the plots for the TR2 group started around t= -250 to -300 days, which suggests a correlation between this intentional process change and the observed improvements in setup accuracy.

3.5.3 Kruskal-Wallis ANOVA

Kruskal-Wallis ANOVA (Section 2.8.3) was used to test and further quantify the statistical significance of the correlation between the observed improvements in setup accuracy and each of the identified potential causes of change; treatment indication, head support and retraining. Initial analysis was performed on each variable separately, with the results displayed below. Histograms were created to demonstrate the distribution and subgroup size for each variable, while each mean rank graph shows the position of the mean rank and confidence interval for each group within the study population.

Treatment indication

Kruskal-Wallis analysis for treatment indication was limited to those patients starting treatment after June 2013, which was the point where skin cancer treatments were included in the study cohort. All patients treated prior to the observed change in treatment indication were therefore excluded, in order to prevent a so called 'selection bias'. As shown in Figure 3-11, the histograms for patient systematic and random overall deformation demonstrate comparable deformation values for each treatment indication, with larynx and oral cavity showing the greatest variation. These results were similarly reflected in the graphs for mean rank and confidence interval, which show non-significant differences in the mean rank position for patient systematic (p = 0.575) and random (p = 0.207) overall deformation. Therefore, this unintentional change could be ruled out as the cause of the observed change in setup accuracy around t = -250 to -300 days.

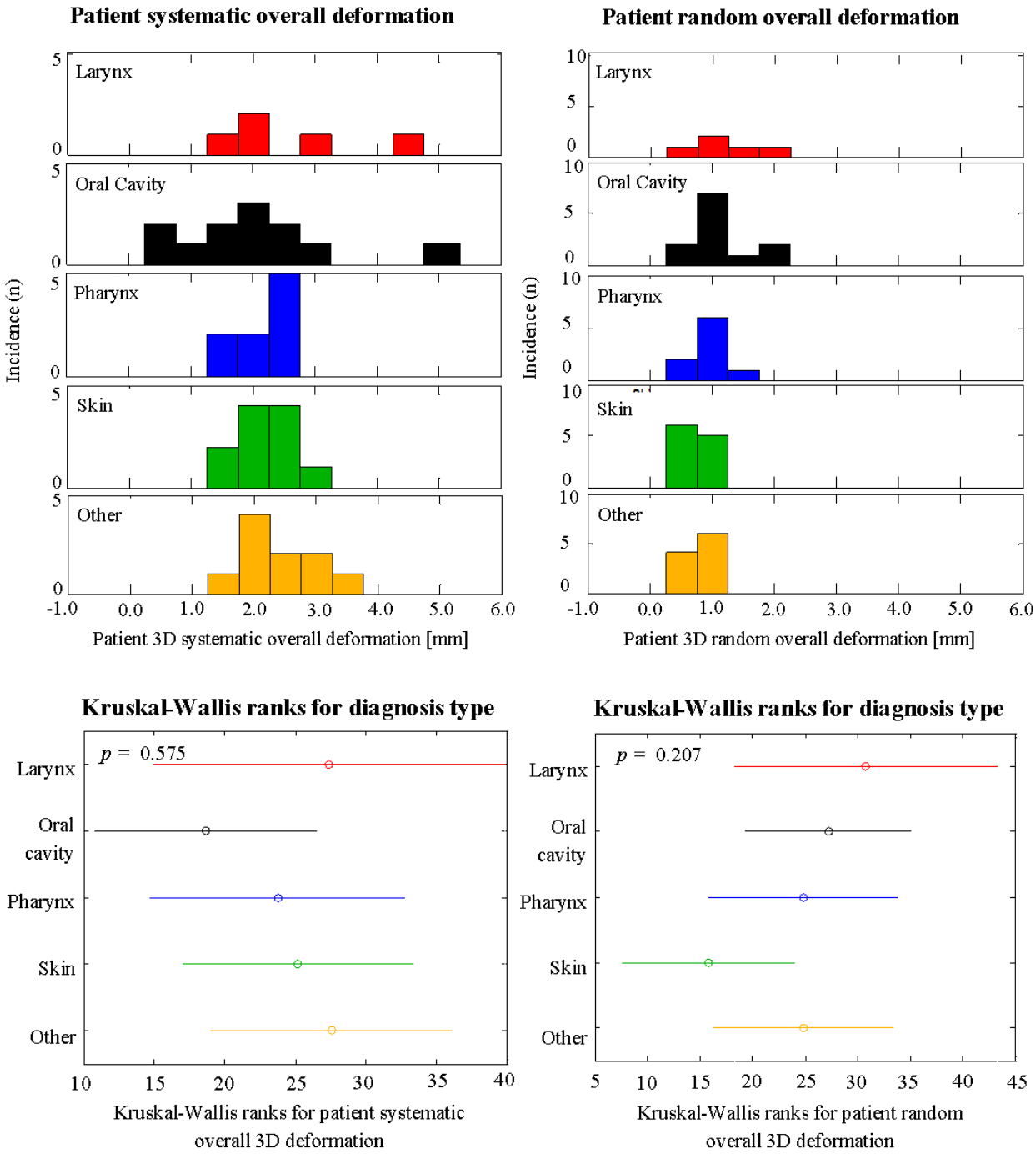


Figure 3-11 Kruskal-Wallis ANOVA for treatment indication. Analysis based on patient datasets from June 2013-March 2014 only. Histograms show the distribution and incidence of patient systematic and random overall deformation for each treatment indication. Graphs for mean rank and confidence interval demonstrate the position of each subgroup for patient systematic and random overall deformation.

Head supports

The analysis results for head supports are displayed in Figure 3-12. Histograms demonstrate large variance in subgroup size, with much higher incidence in HS1/HS2 cohorts and lower numbers for HS3/HS4. The distribution of each histogram suggests lower overall deformation for HS3 and HS4, however these differences were difficult to discern due to the variation in subgroup size. The position of the mean rank and confidence intervals more clearly demonstrated differences between head supports, with lower mean ranks observed for both HS3 and HS4. These differences were shown as significant for patient systematic overall deformation ($p = 0.003$) but non-significant for patient random overall deformation ($p = 0.145$).

Despite the significant *p*-value, overlap of the confidence intervals was observed for patient systematic overall deformation, indicating that these results may only be of borderline significance. HS4 had the smallest subgroup ($n = 6$), which met the minimum requirements to perform a Kruskal-Wallis analysis [63], but resulted in a wide confidence interval for this group, extending into negative values. Therefore, Kruskal-Wallis analysis suggests newer head supports had a positive impact on setup accuracy, however further analysis on HS3 and HS4 with larger subgroup sizes in the future would be required to further validate this conclusion.

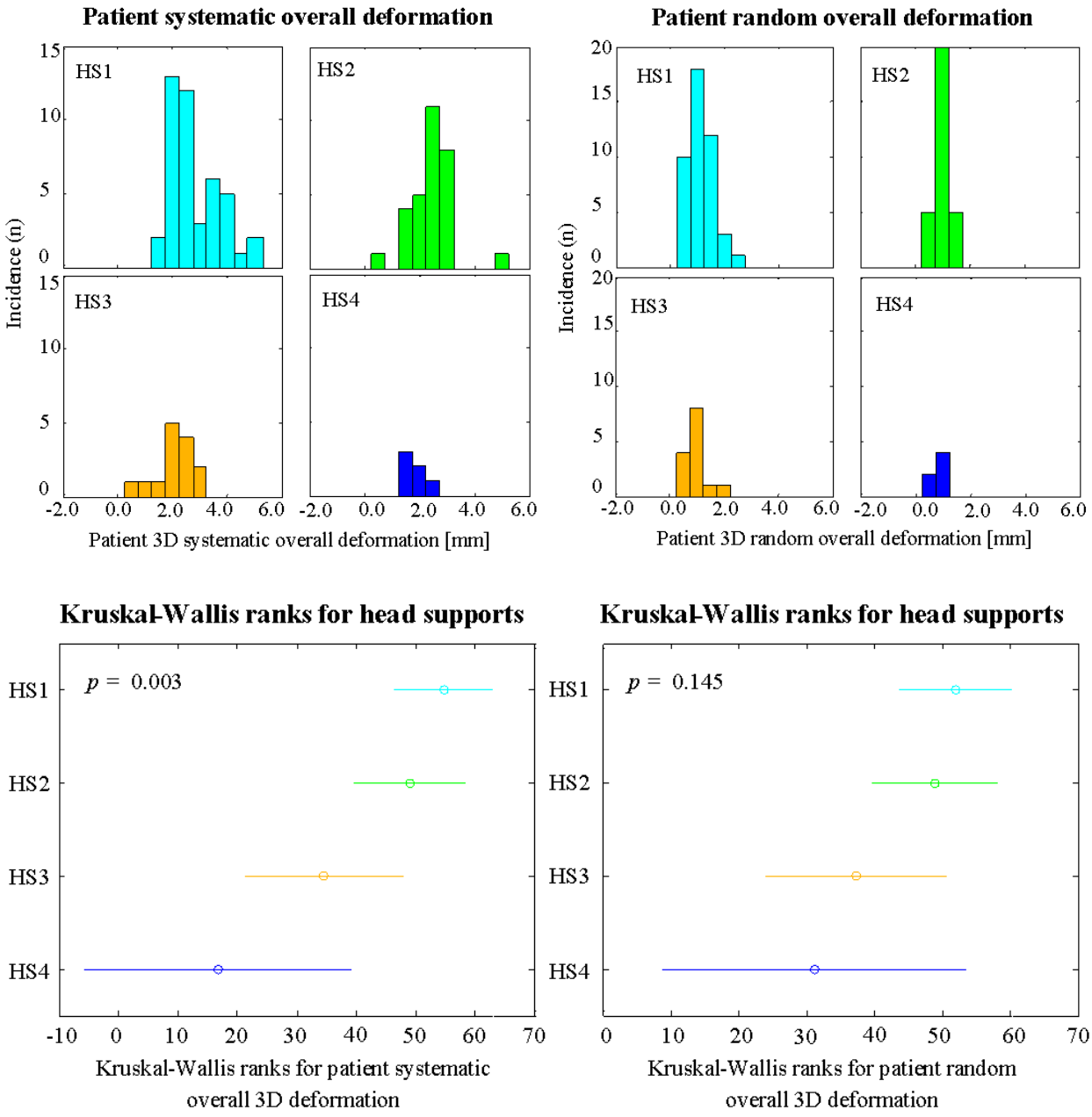


Figure 3-12 Kruskal-Wallis ANOVA for head supports. Histograms show the distribution and incidence of patient systematic and random overall deformation for each head support type. Graphs for mean rank and confidence interval demonstrate the position of each subgroup for patient systematic and random overall deformation.

Retraining for immobilisation equipment

The third potential cause that was analysed using Kruskal-Wallis ANOVA was retraining of RTs in the production of immobilisation equipment, including both custom head supports and thermoplastic masks (Figure 3-13).

Again, histograms demonstrated a considerable variation in the subgroup sizes, with the TR1 cohort showing the lowest patient numbers (n = 10). Similar to the analysis of head supports, the small subgroup size resulted in a wide confidence interval for TR1, leaving the impact of a minimum of one retrained RT involved in immobilisation uncertain. However, in both the histogram distributions and mean ranks, differences between the TR0 and TR2 cohorts were observed. In both Kruskal-Wallis rank charts, substantially lower mean ranks were shown for TR2, with significant difference between the groups found for both patient systematic (p = 0.003) and random (p = 0.004) overall deformation. Furthermore, the confidence intervals for TR0 and TR2 were both relatively small, with no observed overlap between these subgroups, indicating retraining had a substantial impact on patient positioning accuracy.

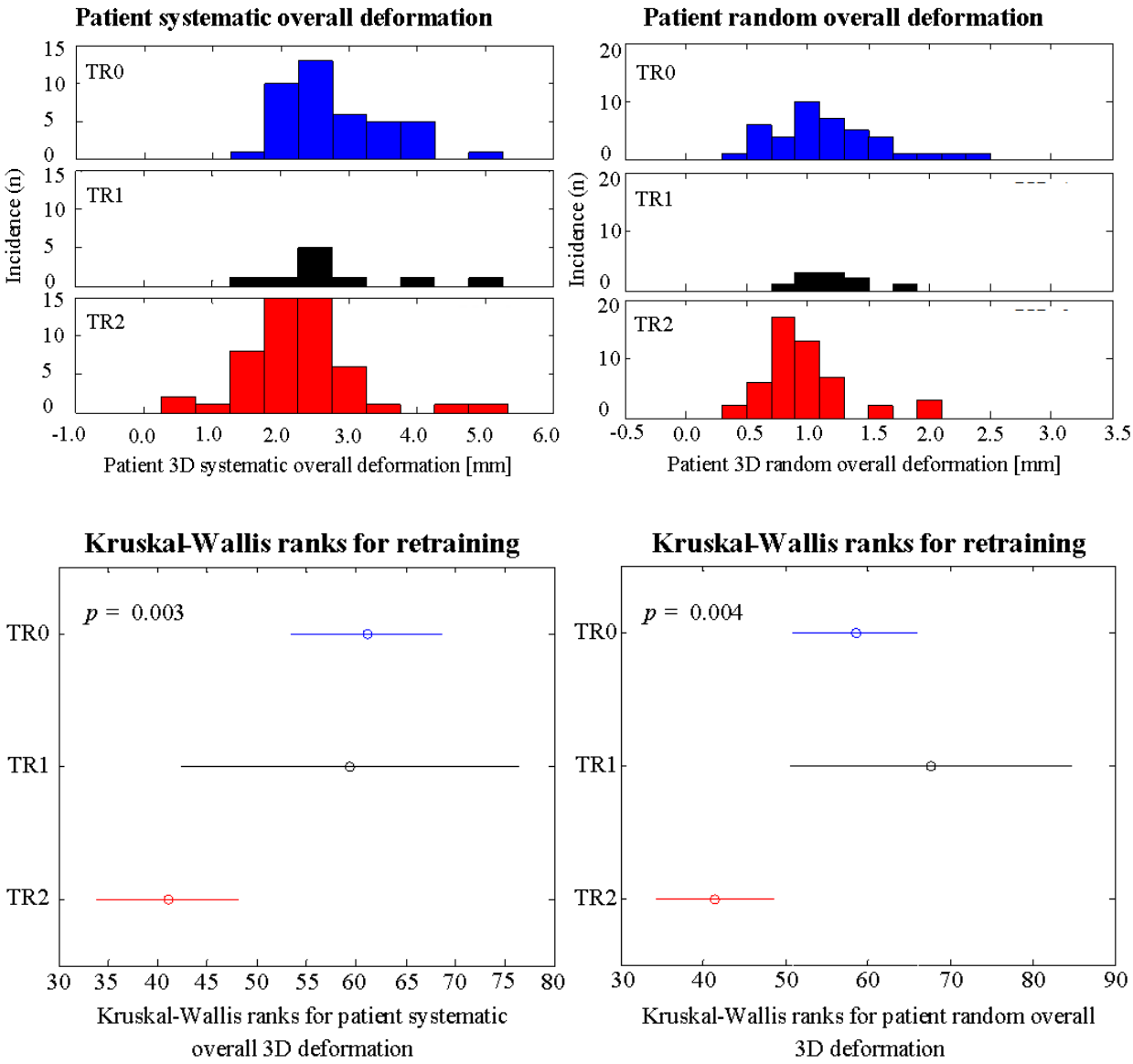


Figure 3-13 Kruskal-Wallis ANOVA for retraining of RTs for immobilisation equipment. Histograms show the distribution and incidence of patient systematic and random overall deformation for each cohort of retrained staff. Graphs for mean rank and confidence interval demonstrate the position of each subgroup for patient systematic and random overall deformation.

Head supports and retraining

Retraining provided instruction on how to fit/make custom head supports, as well as how to make a treatment mask. Therefore, the correlation between head support and retraining with setup accuracy was reviewed using Kruskal-Wallis. Patient datasets were stratified by retraining group (TR0/TR2 only) and head support type. From this analysis, differences in patient systematic overall deformation were indicated as significant ($p = 0.004$), but non-significant for random deformation ($p = 0.305$).

As demonstrated by the histograms (Figure 3-14, Figure 3-15), the size of each subgroup is very unbalanced. Only HS1 and HS2 were used within the cohort of patients were retraining of RTs had not yet occurred (TR0). Furthermore, the subgroup size for HS2 TR0 was small, with $n = 7$. The impact of the smaller subgroup size again resulted in a large confidence interval for HS2 TR0, in both patient systematic and random overall deformation charts. Non-significant differences between HS1 and HS2 within the TR0 cohort were observed in the results, as overlapping mean ranks and confidence intervals. By repeating the comparisons with each individual head support type between the two retraining cohorts, HS1 shows no significant differences between TR0 and TR2, while HS2 shows some improvement in systematic deformation only. This suggests that retraining in immobilisation equipment may result in improved setup accuracy, dependent on the quality of the immobilisation materials being used. However, this is again of borderline significance, as indicated by the overlapping confidence intervals.

In focusing on the differences between the TR0 and TR2 cohorts, lower mean ranks for systematic deformation in all TR2 groups was observed, compared to TR0. Within the TR2 cohort, improved systematic deformation with head support type was also

demonstrated, which was consistent with the results of the previous analysis based on head supports alone (Figure 3-12). However, due to inhomogeneous subgroup sizes, definitive conclusions as to which factor had the greatest impact could not be made.

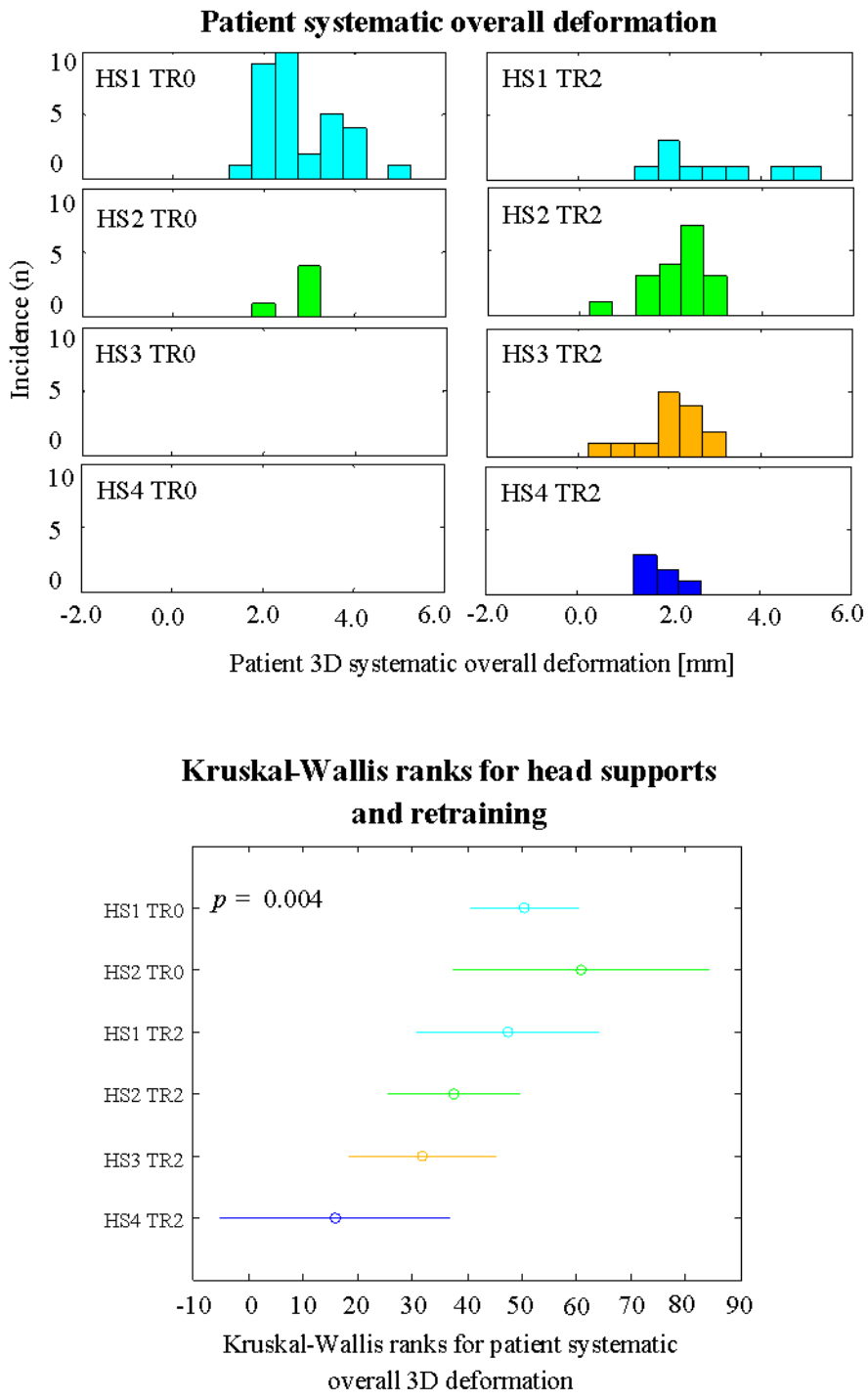


Figure 3-14 Kruskal-Wallis ANOVA for head supports and retraining of RTs for immobilisation equipment, for patient systematic overall deformation. Histograms showing the distribution and incidence of patient systematic overall deformation for TR0 and TR2 cohorts of retrained staff per head support type. Graphs for mean rank and confidence interval demonstrate the position of each subgroup for patient systematic overall deformation.

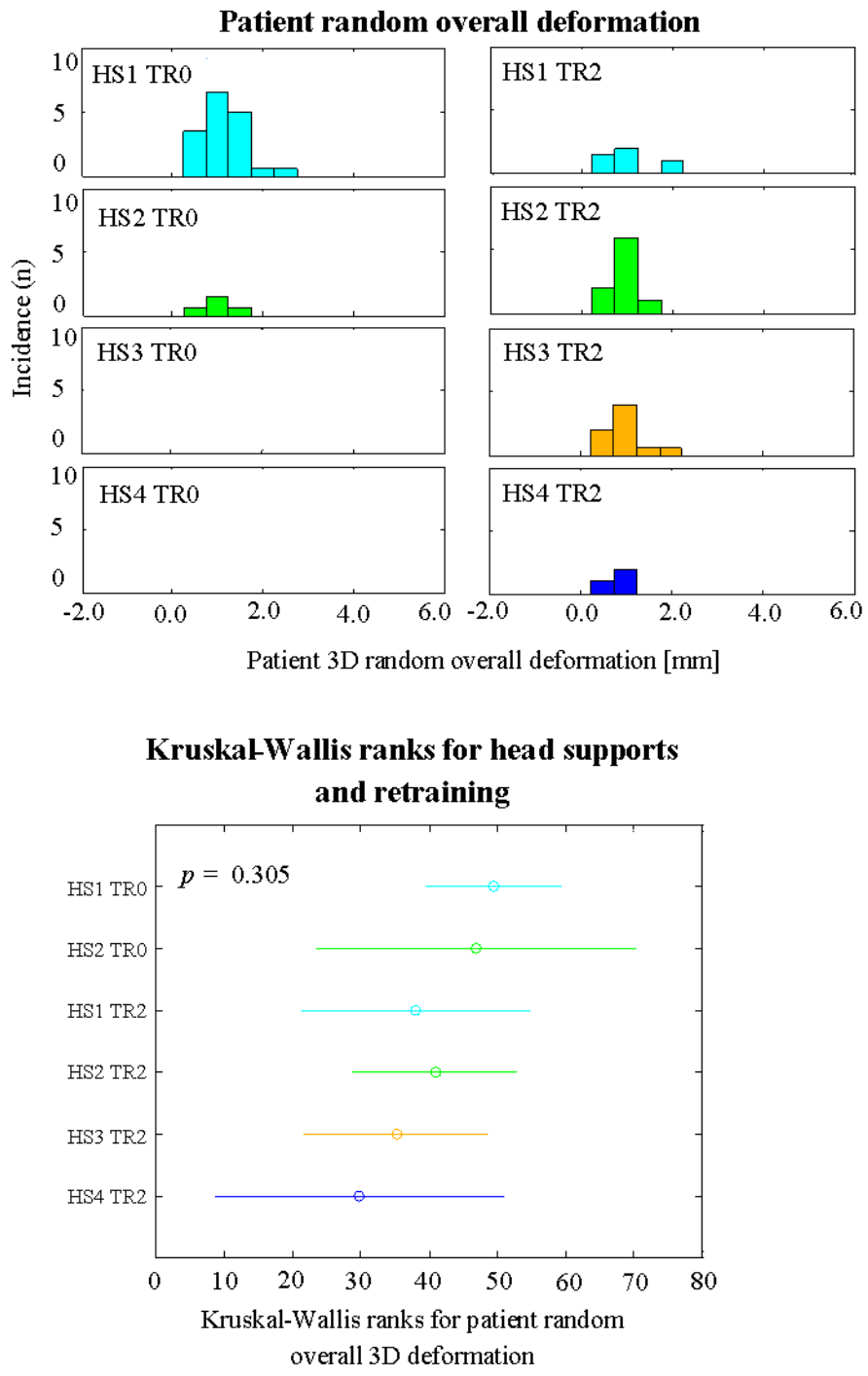


Figure 3-15 Kruskal-Wallis ANOVA for head supports and retraining of RTs for immobilisation equipment, for patient random overall deformation. Histograms showing the distribution and incidence of patient random overall deformation for TR0 and TR2 cohorts of retrained staff per head support type. Graphs for mean rank and confidence interval demonstrate the position of each subgroup for patient random overall deformation.

3.6 Margin calculation

Table 3-3 displays the contribution of deformation to anisotropic PTV margins, calculated for each translational axis per ROI. As per the results for setup uncertainties, margins were calculated for three WBCC cohorts (described in Section **Error! Reference source not found.**). These calculations assumed patient ositioning to C1-C3, as explained in Section 2.9. These values do not take into account other uncertainties in the radiation therapy process and must not be applied as PTV margins. They merely demonstrate the impact of deformation on PTV margins. The results show that the contribution of deformation alone is typically less than the overall PTV margin applied in current clinical practice of 5 mm. This margin is however exceeded for C7-caudal (vrt) and mandible (lng) for the overall cohort, while the lng margin calculated for the larynx far exceeds current PTV margins for both the overall and 2011/2012 cohorts. Further comparisons between the 2011/2012 and reference group are discussed in Section 4.1.5.

Table 3-3 Contribution of deformation to PTV margin per ROI, per translational axis

ROI	PTV margin [mm]								
	WBCC overall			WBCC 2011/2012			WBCC reference group		
	Vrt	Lng	Lat	Vrt	Lng	Lat	Vrt	Lng	Lat
C3-C5	3.5	1.4	2.1	1.9	0.8	1.3	1.0	0.6	0.7
C5-C7	3.7	2.3	3.5	2.8	1.9	1.9	1.6	0.7	1.3
C7-caudal	5.7	3.3	3.6	3.6	2.5	3.2	2.9	1.3	2.2
Mandible	3.9	5.8	2.7	3.3	3.2	1.7	1.5	1.8	1.0
Occipital bone	3.4	2.6	2.7	2.9	1.4	1.7	1.0	0.9	0.7
Larynx	4.0	12.0	3.5	2.5	6.6	2.2	1.3	2.7	1.1

4 DISCUSSION

The overarching aim of this project was to retrospectively quantify patient setup accuracy in radiation treatment for HNC. This study was initiated following a number of changes to patient positioning within the department, which led to improvements in setup accuracy being observed during weekly image review meetings. Based on previous literature, deformation was used as the measure of setup accuracy, which was assessed using CBCT scans of 96 HNC patients treated at WBCC.

4.1 Interpretation of results

4.1.1 Multiple rigid registration protocol

The multiple rigid registration protocol employed in this study provided the means to quantify deformation. It successfully demonstrated the movement of seven anatomical sub-regions within the head-and-neck region, as shown in Figure 3-1. This approach to measuring setup accuracy was consistent with previous literature, where multiple rigid registration protocols have been widely applied to assess deformation in the head-and-neck [11–16].

The selection of ROIs used in this protocol represents those regions typically included in the irradiated volume for many head-and-neck cancers [11]. The base of the skull, clivus, sphenoid sinus and/or maxillary sinus were not included, similar to most of the above mentioned deformation studies. However, these structures could in principle have provided additional deformation information relevant for nasopharyngeal or paranasal sinus cancer treatments, where target volumes frequently abut or overlap with the brainstem or orbits [3,4]. No major deformations in these regions were

observed clinically in our department. Similarly, both Giske *et al.* [13] and Graff *et al.* [14] reported only moderate systematic deformation for these structures, ranging between 0.2 – 1.5 mm per translational axis [13] and 1.3 – 1.5 mm for 3D vectors [14].

4.1.2 Setup uncertainties

Deformation and resultant setup uncertainties were calculated relative to C1-C3. The selection of this reference structure provided a clinically relevant measure of deformation for this patient cohort. Since the introduction of the new IGRT protocol and the requirement for patient specific matching instructions (Section 3.5.2), C1-C3 or similar (C1-C2, C2-C3) was used clinically as the registration ROI for approximately 73% of patients at WBCC. Additionally, C1-C3 has also been described in previous literature as a stable ROI for repositioning due its location in the head support [11], while also acting as the pivot point for the head-and-neck region [14,15]. Furthermore, this reference ROI allows for comparison of results with previous literature, where the same or similar reference structures were used to calculate deformation [11–14].

In this study, setup accuracy was reported in a manner consistent with literature on calculating setup uncertainties [34,37], as well as previous studies on patient deformation during head-and-neck treatments [11–14]. Based on the group statistics for this patient cohort, the relative movement of different anatomical regions within the head-and-neck was clearly demonstrated for each translational axis and associated 3D vectors. Similar to van Kranen *et al.* [11] these results show that ROIs closest to C1-C3 have the lowest deformation, which then increases with increasing distance. Interestingly, the results from the current study demonstrate similar if not lower

deformation values compared to previous literature, in particular after improvements in patient setup were implemented by the MDT.

In order to benchmark the setup accuracy reported at the WBCC, results from previous literature where deformation was reviewed using similar methods were included in Table 3-2. As discussed, patient results from this study were divided into three separate groups, which allowed an assessment of initial positioning accuracy (WBCC 2011/2012) against the positioning accuracy of patients that were treated at the end of 2013 and early 2014 and were used as the reference group for SPC analysis (WBCC reference group).

In reviewing these results, the 3D vectors for mean and systematic dispersion in each patient cohort (overall, 2011/2012 and reference group) were comparable or lower than the results presented by Polat *et al.* [12] and Graff *et al.* [14], for each related structure. Similarly, in comparison to van Kranen *et al.* [11] and Giske *et al.* [13], the group deformation statistics reported for the WBCC were equivalent if not lower for most structures and axes where comparisons were made. Similar to Giske *et al.* [13] large deformation in the lng axis for the mandible was not observed, compared to that reported by van Kranen *et al.* [11]. However, the large systematic dispersion of the larynx reported by van Kranen *et al.* [11] was seen in the overall and 2011/2012 patient cohorts, while lower systematic dispersion in the larynx was seen in the reference group, similar to that reported by Giske *et al.* [13]. These previous studies demonstrated large systematic dispersion in the vrt axis for C7-caudal [11,13], with Giske *et al.* [13] also showing similar results for C5-C7. These larger uncertainties at the inferior spine ROIs were not reflected in the WBCC cohorts

Possible causes for the differences in setup accuracy between the current study and previous literature are difficult to identify, as previous work tended to focus on the impact of deformation and not specifically the possible causes [11–14]. The main differences that may be identified relate to the immobilisation equipment used. In terms of head supports, Polat *et al.* [12] used standard head supports, which may explain the larger deformation, compared to the WBCC. However, van Kranen *et al.* [11] who also showed larger deformation for most ROIs, used custom head supports for their patient cohort. Giske *et al.* [13] used a vacuum mould under the shoulders and thorax but no neck rest, which may explain the increased level of deformation in the vrt axis for C5-C7 and C7-caudal. Graff *et al.* [14] did not specify the type of immobilisation used, either in terms of head support or treatment mask. The differences observed in immobilisation equipment was also reflected by the types of treatment mask used, where Polat *et al.* [12] used three point (head only) thermoplastic masks, van Kranen *et al.* [11] used five point (head, neck and shoulder) thermoplastic masks, and Giske *et al.* [13] used three point scotch cast masks. As previously presented in Section 1.1.1, comparable repositioning accuracy with different treatment mask types and materials have been reported [8–10]. Consequently, the deformation reported for each department does not show a direct relation with the immobilisation equipment used. Immobilisation equipment can therefore only partially explain the differences observed in the comparison of the WBCC with previous literature. It does however clearly demonstrate that differences in deformation can occur and it is important to quantify deformation in individual departments, in order to accurately assess treatment accuracy.

Comparisons between the 2011/2012 cohort and the reference group show that the initial positioning accuracy of the earlier cohort was closest to the results presented in previous literature. However, by focusing on patient positioning issues as a MDT, there have been significant reductions in the level of deformation in HNC treatment setups. This reduction is indicated by the group statistics for the reference group treated at the end of 2013 and early 2014. However, as previously discussed, time trends are not adequately captured by descriptive statistics, therefore SPC charts were used to give more detailed information about these trends.

4.1.3 Statistical Process Control

Within this retrospective study, SPC charts effectively demonstrated time trends in setup accuracy for HNC treatments. The use of overall deformation as the key quality measure provided a suitable assessment of the deformation for six individual ROIs, relative to C1-C3 (Figure 3-1). Furthermore, the setup accuracy observed in the reference group was confirmed to be a stable period of patient positioning, compared to the study period. A group of 15 patients was originally selected for the reference group using a subjective grading system (Section 2.7.2). The final results of the SPC analysis did however confirm that this selection was also representative for a (slightly) larger group of patients. Therefore, it was appropriate to use this period as a reference group to calculate process limits for SPC.

Analysis of setup accuracy using SPC was limited to *EWMA*-SPC charts only, due to the effectiveness of this chart type in detecting small shifts in process stability [41,42,46]. *EWMA*-SPC was also selected based on the non-normality of the dataset (Figure 3-2), and the limitations of *x*-SPC and *mR*-SPC under these conditions [46]. The normality of data is often discussed in relation to SPC. Some groups state that the underlying distribution of data should be checked prior to generating control charts, and where data is found to be non-normal, transformation to an approximately normal distribution should performed [41,42]. Conversely, Wheeler & Chambers [43] demonstrated that while the design of the control chart is based on the normal probability model, these tools are still robust enough to detect exceptional variation for non-normal data. As demonstrated by Borror *et al.* [46] an appropriately designed *EWMA*-SPC is robust to the violation of normality constrains, while the performance of the *x*-SPC charts is degraded.

As discussed in Section 2.7.5, only one rule of interpretation is typically applied to *EWMA*-SPC. This rule states that any value outside the process limits likely indicates a process change and therefore needs to be investigated. This rule provided a simple interpretation of the results, and only one significant shift outside the control limits observed around t = -250 to -300 days before the last reference data point on both patient systematic and random *EWMA*-SPC charts for overall deformation (Figure 3-5). This shift then became the focus of the investigation of potential causes of change of setup accuracy.

4.1.4 Investigation of potential causes of change

Several factors were considered in the investigation of potential causes of change (Section 2.8). Time plots efficiently identified four potential variables that demonstrated changes coinciding with the improvements observed in setup accuracy (Figure 3-10). Weight loss was reviewed independently as this was only numerical variable. Time plots were limited to intentional process changes, specific to patient positioning for HNC treatment and wider departmental changes, as well as other treatment related factors representing unintended changes that may have impacted setup accuracy. Not included in this analysis were factors relating to medical and family histories, as well as social circumstances, such as home support, alcohol intake and smoking history. These factors often affect how well patients tolerate HNC treatment [3], which may in turn impact on setup accuracy [15]. These factors were excluded from analysis, because consistent information could not be collected for all patients due to the retrospective nature of this study and therefore meaningful comparisons could not be made.

Weight loss

Weight loss was calculated as the percentage difference between initial weight at pCT and weight at end of treatment. Percentage weight loss was selected for analysis over absolute weight loss, in order to allow simplified comparisons between patients. As indicated in the results, SPC showed reduction in total weight loss during the study period (Figure 3-8). However, a significant correlation between weight change and setup accuracy was not observed within this study cohort (Figure 3-9). This outcome is consistent with the results presented by Qi *et al.* [66] who found no statistically significant differences in setup accuracy for HNC treatment based on weight loss, both absolute and percentage. Similarly, van Kranen *et al.* [11] reviewed setup accuracy per patient and reported no significant time trends in patients with weight loss.

Treatment indication

Kruskal-Wallis ANOVA for treatment indication (Figure 3-11) demonstrated non-significant correlations with both patient systematic (p = 0.575) and random (p = 0.207) overall deformation. Based on the results, this unintentional process change was ruled out as the cause of change for improved setup accuracy within this patient cohort. It may be hypothesised that no significant correlation in this analysis was observed due to the relatively high level of grouping of treatment indications that were used for this study. As previously indicated in Section 2.8, overlap between those categorical variables listed as additional factors for analysis was expected. Within each treatment indication, different treatment sites, chemotherapy or surgical approach and total treatment dose and fractionation may be included. For example, diagnoses of nasopharynx, oropharynx and hypopharynx were all included within the

same treatment indication, pharynx. However, the primary site of these indications will vary, with nasopharynx grouped in the upper head-and-neck, oropharynx in the mid and hypopharynx in the lower. Similarly, categories of skin and other cover a broad range of primary tumour locations and treatment approaches. It is therefore expected that, if other factors do affect setup accuracy, these effects likely get blurred and may even cancel each other out completely. This most likely resulted in the finding of a non-significant correlation between treatment indication and setup accuracy.

Head supports and retraining of radiation therapists

Head supports and retraining of RTs for immobilisation equipment demonstrated the most significant correlation with improvements in setup accuracy. Analysis of head supports was performed on the four main types used within the department, which were all custom made for each patient (Section 2.2). Previous literature has shown that setup accuracy is improved with the use of custom head supports, compared to standard head rests [20,21]. However, there is a paucity of research on the differences among the various commercially available custom head supports. As demonstrated in this study, significant differences in patient systematic overall deformation (p = 0.003) between the four custom head supports used at the WBCC was found.

HS1 demonstrated the lowest ranks for overall deformation and was used for most of the study period. During this time, noticeable degradation of the polystyrene fill was observed within the department, even though no formal QA was in place to determine the suitability of continued use of these head supports. HS2 was introduced in an attempt to phase out HS1, however these vacuum bags were delivered without indexing. Therefore, an in-house solution was developed to index these head supports

on the pCT and treatment couches as a temporary solution. Furthermore, the 9-litre fill used in HS2 seemed to be too bulky, which gave highly elevated head-and-neck positions and resulted in increased stretching of the treatment mask. These issues cumulated into moderate positioning accuracy for HS2 within this patient cohort, as demonstrated by Figure 3-12.

Due to the issues experienced with HS2, a thermoplastic neck roll was used with HS1 (grouped as HS3) to provide a more standard neck shape where degraded bag fill did not provide adequate support. The introduction of this additional neck support coincided with the improvements in setup accuracy, as indicated by the time plot for head supports (Figure 3-12). HS3 demonstrated lower mean ranks in Kruskal-Wallis ANOVA compared to HS1 and HS2, suggesting the addition of this neck support was effective in improving setup accuracy. However, with the eventual introduction of HS4, further improvements in setup accuracy were seen. HS4 was developed by members of the MDT, in conjunction with the manufacturer following WBCC experiences with HS2. A lower 7-litre fill was requested, along with in-built indexing. The improvements resulted in better repositioning of patients with lower deformation ranks. However, the subgroup population for HS4 was relatively small (n = 6), which resulted in a wide confidence interval in the results. Therefore, Kruskal-Wallis analysis suggests HS4 provides the most accurate repositioning for HNC treatments, however additional data is required to further validate this conclusion.

Retraining of RTs demonstrated the most significant improvement in setup accuracy. The differences observed between retraining cohorts were statistically significant for both systematic (p = 0.003) and random (p = 0.004) deformation. The tutorials for immobilisation equipment were focused on both the impression process for the

thermoplastic mask, as well as instruction on fitting custom head supports. Therefore, convolution between head support and retraining may have occurred. This was reviewed using repeat Kruskal-Wallis ANOVA with subgroups stratified by head support and training cohort. This analysis further supported the results observed in the independent analysis of each variable. The TR0 cohort showed lower setup accuracy compared to TR2, regardless of head support type. Furthermore, similar improvements in setup accuracy within the TR2 cohort were seen for the four head support types. Only HS1 and HS2 were used within both the TR0 and TR2 cohorts. As shown in the results, no significant improvement was seen in setup accuracy for HS1, with only mild improvements seen for HS2. Overall, this suggests that retraining appears to improve setup accuracy, dependent on the quality of the immobilisation materials being used.

4.1.5 Margin Calculations

PTV margin calculations are derived using dose coverage probabilities and were originally based on rigid body motion for a single target volume [65]. However, they can be expanded to include organ motion [67]. Deformation could be included in PTV margin calculations [39], by using different PTV margins for each distinct ROI. It should be kept in mind that this would presently be very impractical, because there are no tools available to facilitate this in our current treatment planning system. Furthermore, as patients are positioned according to a specific protocol, deformation data reflecting the mobility of anatomical sub regions with respect to that particular setup protocol is required. Thus, from the data in the present study, only the (theoretical) contribution of deformation to the overall PTV margin under the

assumption that patients are setup using IGRT registrations to C1-C3 can be calculated.

As previously discussed, approximately 73% of patients are positioned using C1-C3 or similar match structure at WBCC, making the results in Table 3.3 relevant to current practice. The appropriateness of PTV margins applied at WBCC may therefore be reviewed in relation to the contribution of deformation. In this analysis, the contribution of other uncertainties was assumed to give a 4 mm PTV margin, which is not unrealistic considering the PTV margins of 3-5 mm that are reported in literature, which do not include deformation [25,68,69]. By subtracting this margin contribution from the clinically applied PTV margins for HNC of 5 mm, the remaining available contribution would be 3 mm. This subtraction was performed by applying the method described by van Herk [34] to add several uncertainties for PTV margin calculations.

$$Residual\ margin = \sqrt{5^2 - 4^2} = 3\ mm$$

Based on the results displayed in Table 3.3, the 2011/2012 cohort shows five axes across three ROIs with the contribution of deformation exceeding the theoretical threshold of 3 mm. C7-caudal and the mandible each show axes with margin contributions narrowly over 3 mm, while the larynx demonstrated substantial contribution in lng direction. The results for the larynx must be interpreted with care, as the larynx itself is an avoidance structure (unless the treatment indication, where positioning is based on the larynx not C1-C3), although structures attached may be part of the target volume.

All structures within the reference group showed results less than 3 mm. From this interpretation, the improvements made in patient positioning are further demonstrated by the contribution of deformation relative to C1-C3 to the overall PTV margin. Continued monitoring of positioning results for patients treated after this study did highlight two exceptional cases where larger deformation was observed. These results could be explained by exceptional variation related to patient specific factors, which for one patient was due to the removal of all of their hair, but also emphasised the need for ongoing monitoring of setup accuracy even though our positioning results have improved (on average) at the WBCC.

4.2 Limitations

4.2.1 Observer variability and bias

Having used a single observer for the multiple rigid registration protocol, intra-observer variability and bias (inter-observer variability with one observer) was reviewed. Using 15 patients (2 fractions per patient), an independent audit and repeat registrations were performed by a second reviewer and the researcher respectively. These secondary registrations agreed with the setup uncertainties recorded for the study within 1 mm for all axes. This indicates that the variability between two observers was limited, as well as the variability in repeat registrations by the researcher. Variation was likely limited by the specificity of the multiple rigid registration protocol and the application of an auto registration algorithm.

4.2.2 Limited degrees of freedom

Translational errors were measured and reported in this study protocol, as vrt, lng and lat axes. The decision to limit the analysis to three degrees of freedom was primarily based on a main objective of this study, which was to determine the change in distance between two objects (deformation), not the change in mutual orientation (rotations) [11]. Due to the use of small ROIs in this analysis, the impact of orientation on target volume coverage was considered to be small and therefore less relevant for the analysis of deformation.

4.2.3 Retrospective analysis

Retrospective analysis limited the accuracy and efficiency with which potential causes of change were identified. As outlined in Section 2.8, identification of possible factors affecting setup accuracy was based on a thorough review of previous meeting minutes and changes to departmental protocols. This review covered a period of almost three years and identified major process changes, though may have missed small modifications in practice that were managed within the team, and were possibly not recorded in the documents identified for review. While it is unlikely small process changes would result in the significant improvements observed in setup accuracy, it is important to note that these potential changes were not accounted for. Furthermore, as those changes included in the analysis were based on retrospective clinical patient information, significant variation in subgroup populations for different factors was observed. The impact on the analysis for head supports and retraining was previously discussed in Section 3.5.3.

4.2.4 Kruskal-Wallis ANOVA

Kruskal-Wallis ANOVA was selected to test the statistical significance of possible causes of change in setup accuracy because this dataset was not normally distributed and the subgroups had an unequal variance (Figure 3-2). The limitation of this approach was the loss of information and power that occurs in substituting patient systematic and random overall deformation for ranks [63]. This means that while differences in head support and retraining were highlighted by the analysis the magnitude of these changes in terms of overall deformation is unknown.

Additionally, as Kruskal-Wallis ANOVA was performed using overall deformation only, correlations between the improvements observed in the setup accuracy of individual ROIs were not made. The improvements in repositioning of the spine, occipital bone and mandible may be logically described by changes in head supports and retraining for immobilisation equipment. Improved head supports would provide better support through the neck, resulting in less flexion, while improved mask impressions would reduce movement within the shell. However, the most substantial improvement was observed in the larynx, which is the most difficult to explain as immobilisation does not specifically restrict this motion [70,71]. Individual analysis of changes in setup accuracy against the setup uncertainties recorded for each ROI is required to further analyse these changes, but could not be carried out due to the time restrictions for this study.

4.3 Strategies for the management of patient positioning

In line with previous literature, this study demonstrates that the mobility of different anatomical sub regions within the head-and-neck can be considerable despite the use of immobilisation equipment. Over time, this department has been able to reduce the level of deformation observed during HNC treatments, by focusing on this issue as a MDT. However, additional and ongoing measures to maintain and further improve setup accuracy are required.

4.3.1 Quality assurance of immobilisation equipment

Head supports were identified as a cause of change in setup accuracy, with HS4 demonstrating higher repositioning accuracy for HNC treatments. Based on these results, the department is currently phasing out HS1 and HS3, and keeping HS2 only for those patients where an elevated head position is required due to an inability to lie flat. Additional supplies of HS4 have been ordered and this immobilisation device is being confirmed as the primary head support at WBCC.

As previously stated, the quality of some of the HS1 bags had degraded over time but a lack of QA prevented identification and removal of these unsuitable head rests. In order to avoid a future systematic degradation of the quality of head support and masks, a formal QA system for immobilisation equipment is suggested. This would at least require a coding system to track all head supports, which would allow the department to monitor the number of head supports in circulation at any time, as well as identify any faulty or degraded bags based on positioning accuracy. In addition to reviewing individual head supports, a regular audit of thermoplastic masks is recommended, for instance to highlight the need for retraining.

4.3.2 Retraining for immobilisation equipment

The other key factor identified as contributing to reduced deformation was retraining of radiation therapists in the use of immobilisation equipment. It is therefore suggested that regular tutorials on immobilisation equipment be held, with a minimum requirement that all RTs attend a refresher session at least once a year.

4.3.3 Statistical Process Control

This project has demonstrated a possible application of SPC in reviewing setup accuracy for HNC treatments. Continued use of this tool in a prospective setting has the potential to effectively monitor setup accuracy, identifying changes and improving processes. The comparison of control limits in SPC and process specifications to determine the process capability would further increase the efficacy of SPC.

Specifications are often set for processes or outputs, and in radiation therapy are usually based on clinical judgment, professional societies, best practice and/or clinical protocols [38,45]. In terms of setup accuracy, specifications are typically based on the positioning protocol being used, as either an offline or online protocol. Using the example of Pawlicki *et al.* [45], many quality assurance processes in radiation therapy are checked within specification limits of $\pm x\%$. Readings that fall within this range are not acted upon; meaning the maximum accuracy of this process can only be confirmed as $\pm x\%$ [45]. In reviewing processes using SPC, control limits are determined by the process itself and subsequently verified against the clinical specifications. When the process limits are well within the clinical specifications, exceptional deviations are less likely to violate clinical constrains, and eventually treatment improvements such as margin reduction may be justified, if the setup

accuracy is consistently improved. This requires a continuous cycle of improvement, where adjustments in process may be realised, and then thresholds recalculated [45].

While this study focused on departmental trends in patient deformation using SPC, this tool may also be applied at an individual patient level to effectively manage offline setup corrections. Ung & Wee [57] applied cumulative sum (CUSUM) control charts for this purpose during radiation treatment for prostate cancer. CUSUM is similar to the *EWMA*-SPC chart, but uses subgroups of data points [41,42]. The results presented by Ung & Wee [57] demonstrated that the CUSUM chart was useful in identifying systematic shifts in the setup accuracy, which reduced residual setup uncertainties exceeding the applied threshold by approximately half.

Initial tests in our department have shown that applying the *EWMA*-SPC chart during the treatment of individual patients is a feasible method for setup monitoring, where each treatment fraction represents an individual value in SPC. This would allow monitoring of patient specific deformation during treatment and may provide timely information where plan adaptation or new immobilisation equipment is required. Furthermore, this information would still be able to be applied to SPC charts reviewing departmental trends, meaning monitoring at both a patient and departmental level can be performed using the same data. However, the method of gathering information on deformation is a current drawback to this approach, as it is reliant on manual offline multi ROI matching that may be too time consuming to be practical in the clinic.

Relation between SPC charts and descriptive statistics

SPC not only helped in identifying the cause of the observed changes in patient deformation but also provided insight into how the group statistics M, Σ, and σ changed over time. Based on the SPC charts applied in this study, the CL of the patient systematic deformation chart represents the group mean of means (M) during the control period, while the variability of the data in that chart is related to the systematic dispersion (Σ). Similarly, the CL of the patient random deformation chart represents the group random deformation (σ) during the control period. When no changes occur and the process is stable, the conventional descriptive group statistics well describe all variation within the process. However, as soon as trends come in to play, SPC charts present an added value.

For instance, Figure 3-5 shows a slow downward trend in the patient systematic deformation during the study period, which eventually yields a lower mean of means in the control period. A reduction in the variation or ‘noise’ around the patient systematic deformation was also demonstrated, resulting in lower systematic dispersion, Σ. Similarly, the SPC charts for the patient random deformation show a downward trend eventually yielding a lower group random deformation. The 'noise' or variation of the patient random deformation represents the random dispersion, which is a statistic rarely calculated and not included in this analysis. This disucssion demonstrates the usefulness of SPC in reviewing setup accuracy, compared to providing only descriptive statistics.

4.3.4 Multiple rigid registration protocol

The ideal method for measuring deformation in the head-and-neck would be to employ an automatic multiple rigid registration tool. Following the study by van Kranen *et al.* [11], the same group continued this research line and developed a correction strategy to manage deformations in HNC treatments [72,73]. This approach consisted of rigid registrations looped over multiple ROIs, with optimal couch corrections derived using the setup uncertainties from each ROI [72]. However, this type of IGRT tool is not currently commercially available, meaning continued assessment of deformation at the WBCC must be performed offline by manually selecting different ROIs. While the residual setup errors in this subsequent study by van Kranen *et al.* [72] were comparable to the results for the reference group of the current study, it is preferable to improve setup accuracy by using improved immobilisation equipment, rather than using a complex correction protocol to minimise the impact of deformation. This further highlights the potential benefits of good quality management in this area.

4.3.5 Variable planning target volume margins

Another approach to managing deformation in the head-and-neck region would be to use anisotropic PTV margins. However, as previously indicated, current PTV margin recipes are based on rigid body motions and are not applicable to deformations [11,34]. Yang *et al.* [39] described a method for creating variable margins, based on the deformation results reported by van Kranen *et al.* [11]. These authors weighted the margin expansion based on the position of the target relative to various ROIs. Their study was based on a single patient and provided preliminary proof of principle, but failed to demonstrate substantial differences in the planned dose to target and organ

risk volumes with a variable margin [39]. Further validation of this type of approach is likely needed before clinical application, and similar as stated for the correction strategy discussed in the previous section, it is preferable to avoid deformations as much as possible by improving immobilisation equipment and the processes around it.

4.4 Summary and Conclusions

Accurate patient positioning is extremely important in radiation therapy for head-and-neck cancer. At the WBCC, the introduction of CBCT demonstrated that positioning accuracy required improvement. This led to the establishment of a multidisciplinary team focused on patient positioning, including representation from radiation oncologists, medical physicists and radiation therapists. A number of process changes around patient positioning were initiated by this group, which eventually led to improved setup accuracy being observed during weekly image review meetings.

The present study quantified setup accuracy at the WBCC using CBCT images from 96 patients treated for HNC. Setup accuracy was determined by calculating the deformation observed in the head-and-neck, as the movement of different anatomical sub regions relative to C1-C3. This demonstrated deformation exceeding the currently applied PTV margins, which may result in under dosing of the target volume or unnecessary irradiation of organs at risk.

Analysis of deformation using SPC clearly demonstrated time trends in setup accuracy. This tool was effective in identifying specific time points where improvements occurred, which could then be correlated with process changes made in the department. Through this, changes initiated by the MDT for patient positioning around head supports and retraining of radiation therapists in immobilisation equipment were identified as the main causes of reduced deformation in HNC treatments.

The overall conclusions for this study are:

- Deformation is commonly observed when repositioning patients during HNC treatment and may be measured by calculating the relative position of anatomical sub regions.

- SPC is an effective tool to monitor patient deformation over time.

- A multidisciplinary approach to patient positioning was successful in improving setup accuracy during head-and-neck radiation treatment.

- Improvement of immobilisation equipment and retraining were found to have contributed significantly to the reduction of patient deformation.

- Ongoing monitoring of positioning accuracy is required in order to maintain and/or improve current positioning accuracy.

References

1. National Head and Neck Cancer Tumour Standards Working Group. Standards of Service Provision for Head and Neck Cancer Patients in New Zealand-Provisional. Wellington, NZ: Ministry of Health; 2013.

2. Ministry of Health. Cancer: New registrations and death 2011. Wellington, NZ: Ministry of Health; 2014.

3. Argiris A, Karamouzis M V, Raben D, Ferris RL. Head and neck cancer. Lancet. 2008;371(9625):1695–709.

4. Marur S, Forastiere AA. Head and Neck Cancer: Changing Epidemiology, Diagnosis, and Treatment. Mayo Clin Proc. 2008;83(4):489–501.

5. Washington C, Leaver D, editors. Principles and practice of radiation therapy. 3 ed. St Louis, MO: Mosby Inc; 2010.

6. Lee N, Puri DR, Blanco AI, Chao KSC. Intensity-modulated radiation therapy in head and neck cancers: an update. Head Neck. 2007;29(4):387–400.

7. Verellen D, De Ridder M, Linthout N, Tournel K, Soete G, Storme G. Innovations in image-guided radiotherapy. Nat Rev Cancer. 2007 Dec;7(12):949–60.

8. Tsai J-S, Engler MJ, Ling MN, Wu JK, Kramer B, Dipetrillo T, et al. A non-invasive immobilization system and related quality assurance for dynamic intensity modulated radiation therapy of intracranial and head and neck disease. Int J Radiat Oncol Biol Phys. 1999;43(2):455–67.

9. Boda-Heggemann J, Walter C, Rahn A, Wertz H, Loeb I, Lohr F, et al. Repositioning accuracy of two different mask systems-3D revisited: comparison using true 3D/3D matching with cone-beam CT. Int J Radiat Oncol Biol Phys. 2006 Dec;66(5):1568–75.

10. Weltens C, Kesteloot K, Vandevelde G, Van den Bogaret W. Comparison of plastic and orfit masks for patient head fixation during radiotherapy: precision and costs. Int J Radiat Oncol Biol Phys. 1995;33(2):499–507.

11. Van Kranen S, van Beek S, Rasch C, van Herk M, Sonke J-J. Setup uncertainties of anatomical sub-regions in head-and-neck cancer patients after offline CBCT guidance. Int J Radiat Oncol Biol Phys. 2009;73(5):1566–73.

12. Polat B, Wilbert J, Baier K, Flentje M, Guckenberger M. Nonrigid patient setup errors in head-and-neck region. Strahlentherapie und Onkol. 2007 Sep;183(9):506–11.

13. Giske K, Stoiber EM, Schwarz M, Stoll A, Muenter 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(2):582–9.

14. 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. Elsevier Inc.; 2013 May 1;86(1):170–6.

15. 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 Apr 1;64(5):1559–69.

16. Ahn PH, Ahn AI, 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 Feb 1;73(2):626–33.

17. Q-Fix. How to fabricate Fibreplast Thermoplastic head & neck immobilization devices [pamphlet]. Avodale, PA: Q-Fix; No date.;

18. Gilbeau L, Octave-prignot M, Loncol T, Renard L, Scalliet P, Grégoire V. Comparison of setup accuracy of three different thermoplastic masks for the treatment of brain and head and neck tumors. Radiother Oncol. 2001;58(2):155–62.

19. Wellington Blood and Cancer Centre. Mould room - patient immobilisation aids [procedure document]. Wellington, NZ: Capital and Coast District Health Board; 2012.

20. Van Lin ENJT, van der Vight L, Huizenga H, Kaanders JHAM, Visser AG. Set-up improvement in head and neck radiotherapy using a 3D off-line EPID-based correction protocol and a customised head and neck support. Radiother Oncol. 2003;68:137–48.

21. Houweling AC, van der Meer S, van der Wal E, Terhaard CHJ, Raaijmakers CPJ. Improved immobilization using an individual head support in head-and-neck cancer patients. Radiother Oncol. Elsevier Ireland Ltd; 2010;96(1):100–3.

22. Civco Medical Solutions. Thermoplastic Tips [Internet]. 2007 [cited 2014 Sep 18]. Available from: http://www.civco.com/ro/resources/brochures /thermoplastic_tips.pdf

23. Puri DR, Chou W, Lee N. Intensity-Modulated Radiation Therapy in Head and Neck Cancers. Am J Clin Oncol. 2005 Aug;28(4):415–23.

24. Vanetti E, Clivio A, Nicolini G, Fogliata A, Ghosh-Laskar S, Agarwal JP, et al. Volumetric modulated arc radiotherapy for carcinomas of the oro-pharynx, hypo-pharynx and larynx: A treatment planning comparison with fixed field IMRT. Radiother Oncol. Elsevier Ireland Ltd; 2009;92(1):111–7.

25. Verbakel WFAR, Cuijpers JP, Hoffmans D, Bieker M, Slotman BJ, Senan S. Volumetric intensity-modulated arc therapy vs. conventional IMRT in head-and-neck cancer: a comparative planning and dosimetric study. Int J Radiat Oncol Biol Phys. 2009;74(1):252–9.

26. Holt A, Van Gestel D, Arends MP, Korevaar EW, Schuring D, Kunze-Busch MC, et al. Multi-institutional comparison of volumetric modulated arc therapy vs. intensity-modulated radiation therapy for head-and-neck cancer: a planning study. Radiat Oncol. Radiation Oncology; 2013 Jan;8(1):26.

27. Korreman S, Rasch C, McNair H, Verellen D, Oelfke U, Maingon P, et al. The European Society of Therapeutic Radiology and Oncology-European Institute of Radiotherapy (ESTRO-EIR) report on 3D CT-based in-room image guidance systems: a practical and technical review and guide. Radiother Oncol. Elsevier Ireland Ltd; 2010 Feb;94(2):129–44.

28. Van Herk M. Different Styles of Image-Guided Radiotherapy. Semin Radiat Oncol. 2007;17:258–67.

29. Jaffray DA, Siewerdsen JH, Wong JW, Martinez AA. Flat-panel cone-beam computed tomography for image-guided radiation therapy. Int J Radiat Oncol Biol Phys. 2002 Aug 1;53(5):1337–49.

30. Varian Medical Systems. On-Board Imager (OBI) Advanced Imaging Reference Guide. Palo Alto, CA: Varian Medical Systems, Inc.; 2012.

31. Varian Medical Systems. TrueBeam (TM) System Imager Arms Extended [Internet]. Palo Alto, CA: Varian Medical Systems, Inc.; [cited 2014 Dec 5]. Available from: http://newsroom.varian.com/index.php?s=31899&view=category&mode=gallery&cat=2470

32. Guckenberger M, Meyer J, Wilbert J, Baier K, Sauer O, Flentje M. Precision of image-guided radiotherapy (IGRT) in six degrees of freedom and limitations in clinical practice. Strahlentherapie und Onkol. 2007 Jun;183(6):307–13.

33. Wellington Blood and Cancer Centre. Verification - on treatment imaging procedure [procedure document]. Wellington, NZ: Capital and Coast District Health Board; 2013.

34. Van Herk M. Errors and Margins in Radiotherapy. Semin Radiat Oncol. 2004;14(1):52–64.

35. International Commission on Radiation Units and Measurements. ICRU Report 50: prescribing, recording and reporting photon beam therapy. Prescribing, Recording and Reporting Photon Beam Therapy, Report. Bethusda, MD: International Commission on Radiation Units and Measurements; 1993.

36. International Commission on Radiation Units and Measurements. ICRU Report 62: prescribing, recording and reporting photon beam therapy. Bethusda, MD: International Commission on Radiation Units and Measurements; 1999.

37. Remeijer P, Geerlof E, Ploeger L, Gilhuijs K, van Herk M, Lebesque J V. 3-D Portall image analysis in clinical practice: an evaluation of 2-D and 3-D analysis techniques as applied to 30 prostate cancer patients. Int J Radiat Oncol Biol Phys. 2000;46(5):1281–90.

38. Pawlicki T, Whitaker M, Boyer AL. Statistical process control for radiotherapy quality assurance. Med Phys. 2005;32:2777–86.

39. 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.

40. 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(4):960–70.

41. Montgomery DC. Introduction to Statistical Quality Control. Sixth. John Wiley & Sons, Inc; 2009.

42. Oakland JS. Statistical Process Control. Sixth. Butterwork-Heinemann; 2008.

43. Wheeler DJ, Chambers DS. Understanding Statistical Process Control. 3 ed. Knoxville, TN: SPC Press; 2010.

44. Petrie A, Sabin C. Medical Statistics at a Glance. Oxford, UK: Blackwell Sciene Ltd; 2000.

45. Pawlicki T, Chera B, Ning T, Marks LB. The Systematic Application of Quality Measures and Process Control in Clinical Radiation Oncology. Semin Radiat Oncol. Elsevier Inc.; 2012;22(1):70–6.

46. Borror CM, Montgomery DC, Runger GC. Robustness of the EWMA control chart to non-normality. J Qual Technol. 1999;31(3):309–16.

47. Holli KA, Laippala P, Ojala A, Pitkänen M. Quality control in health care: an experiment in radiotherapy planning for breast cancer patients after mastectomy. Int J Radiat Oncol Biol Phys. 1999;44(4):827–33.

48. Pitkänen MA, Holli KA, Ojala AT, Laippala P. Quality Assurance in Radiotherapy of Breast Cancer Variability in Planning Target Volume Delineation. Acta Oncol. 2001;40(1):50–5.

49. Sanghangthum T, Suriyapee S, Srisatit S, Pawlicki T. Retrospective analysis of linear accelerator output constancy checks using process control techniques. J Appl Clin Med Phys. 2013;14(1):147–60.

50. Breen SL, Moseley DJ, Zhang B, Sharpe MB. Statistical process control for IMRT dosimetric verification. Med Phys. 2008;35(10):4417–25.

51. Chung JB, Kim JS, Ha SW, Ye S-J. Statistical analysis of IMRT dosimetry quality assurance measurements for local delivery guideline. Radiat Oncol. BioMed Central Ltd; 2011 Jan;6(1):27–34.

52. Gérard K, Grandhaye J-P, Marchesi V, Kafrouni H, Husson F, Aletti P. A comprehensive analysis of the IMRT dose delivery process using statistical process control (SPC). Med Phys. 2009;36(4):1275–85.

53. Nordström F, af Wetterstedt S, Johnsson S, Ceberg C, Bäck SAJ. Control chart analysis of data from a multicenter monitor unit verification study. Radiother Oncol. 2012 Mar;102(3):364–70.

54. Palaniswaamy G, Scott Brame R, Yaddanapudi S, Rangaraj D, Mutic S. A statistical approach to IMRT patient-specific QA. Med Phys. 2012 Dec;39(12):7560–70.

55. Pawlicki T, Yoo S, Court LE, McMillan SK, Rice RK, Russell JD, et al. Process control analysis of IMRT QA: implications for clinical trials. Phys Med Biol. 2008 Sep 21;53(18):5193–205.

56. Sanghangthum T, Suriyapee S, Srisatit S, Pawlicki T. Statistical process control analysis for patient-specific IMRT and VMAT QA. J Radiat Res. 2013 May;54(3):546–52.

57. Ung NM, Wee L. Cumulative sum method in nonzero fixed action level setup correction strategy: an application of statistical process control for targeted prostate radiotherapy. Med Phys. 2012 May;39(5):2746–53.

58. Civco Medical Solutions. Vac-Lok ™ Cushions [Internet]. Civco; [cited 2014 Sep 18]. Available from: http://www.civco.com/ro/resources/brochures/vaclokcushionsguide.pdf

59. Thode HC. Testing for Normality. New York, NY: Hoboken: Marcel Dekker Inc; 2002.

60. Kaltenbach H-M. A Concise Guide to Statistics. Berlin, Heidelberg: Springer Berlin Heidelberg; 2012.

61. Kan MWK, Leung LHT, Wong W, Lam N. Radiation Dose From Cone Beam Computed Tomography for Image-Guided Radiation Therapy. Int J Radiat Oncol Biol Phys. 2008;70(1):272–9.

62. Hinton PR. Statistics Explained. Taylor and Francis; 2014.

63. Nussbaum EM. Categorical and Nonparametric Data Analysis: Choosing the Best Statistical Technique. Hoboken: Taylor and Francis; 2014.

64. Newman MC. Quantitative ecotoxicology. 2 ed. Boca Ranton, FL: CRC Press; 2013.

65. Van Herk M, Remeijer P, Rasch C, Lebesque J V. The probability of correct target dosage: dose-population histograms for deriving treatment margins in radiotherapy. Int J Radiat Oncol Biol Phys. 2000 Jul 1;47(4):1121–35.

66. Qi XS, Hu AY, Lee SP, Lee P, DeMarco J, Li XA, et al. Assessment of interfraction patient setup for head-and-neck cancer intensity modulated radiation therapy using multiple computed tomography-based image guidance. Int J Radiat Oncol Biol Phys. Elsevier Inc.; 2013 Jul 1;86(3):432–9.

67. Sonke J-J, Rossi M, Wolthaus J, van Herk M, Damen E, Belderbos J. Frameless stereotactic body radiotherapy for lung cancer using four-dimensional cone beam CT guidance. Int J Radiat Oncol Biol Phys. 2009 Jun 1;74(2):567–74.

68. 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 Apr;76(5):1353–9.

69. Feng FY, Kim HM, Lyden TH, Haxer MJ, Feng M, Worden FP, et al. Intensity-modulated radiotherapy of head and neck cancer aiming to reduce dysphagia: early dose-effect relationships for the swallowing structures. Int J Radiat Oncol Biol Phys. 2007 Aug 1;68(5):1289–98.

70. Van Asselen B, Raaijmakers CPJ, Lagendijk JJW, Terhaard CHJ. Intrafraction motions of the larynx during radiotherapy. Int J Radiat Oncol Biol Phys. 2003 Jun;56(2):384–90.

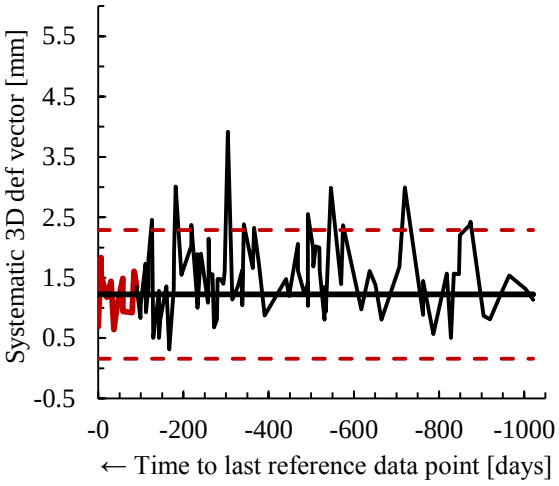
71. Gangsaas A, Astreinidou E, Quint S, Levendag PC, Heijmen B. Cone-beam computed tomography-guided positioning of laryngeal cancer patients with large interfraction time trends in setup and nonrigid anatomy variations. Int J Radiat Oncol Biol Phys. Elsevier Inc.; 2013 Oct 1;87(2):401–6.

72. Van Kranen S, van Beek S, Mencarelli A, Rasch C, van Herk M, Sonke J-J. Correction strategies to manage deformations in head-and-neck radiotherapy. Radiother Oncol. Elsevier Ireland Ltd; 2010 Feb;94(2):199–205.

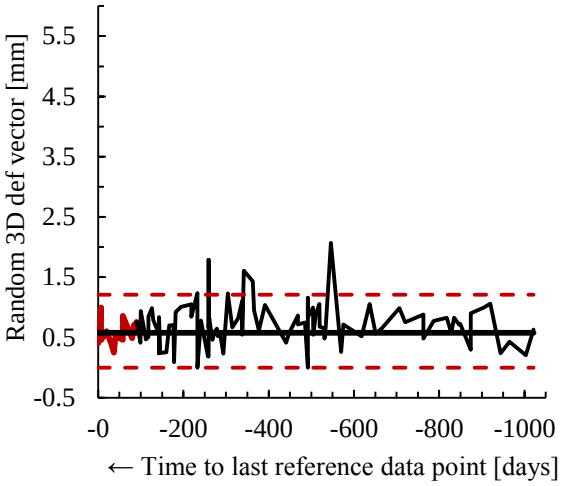
73. Van Beek S, van Kranen S, Mencarelli A, Remeijer P, Rasch C, van Herk M, et al. First clinical experience with a multiple region of interest registration and correction method in radiotherapy of head-and-neck cancer patients. Radiother Oncol. Elsevier Ireland Ltd; 2010 Feb;94(2):213–7.

APPENDIX A: Individual value SPC charts per ROI

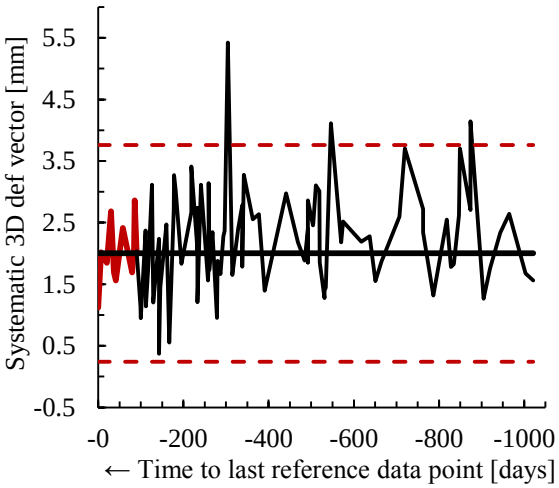
Patient systematic C3-C5 deformation



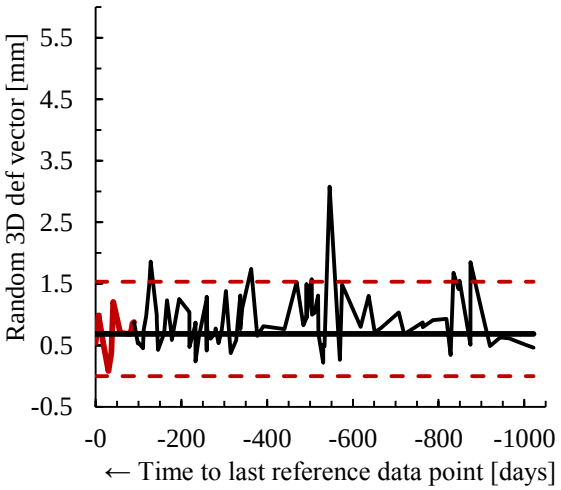
Patient random C3-C5 deformation



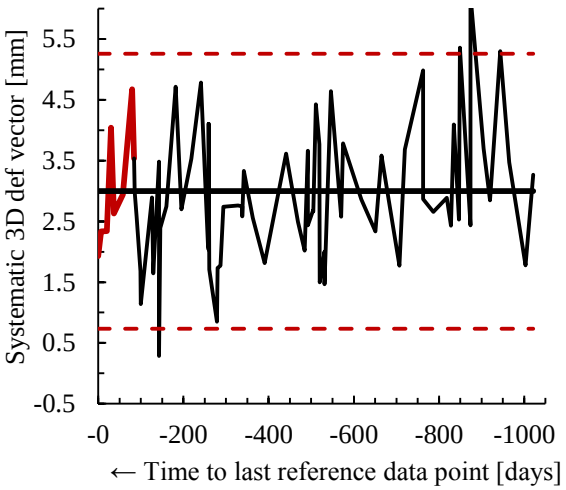
Patient systematic C5-C7 deformation



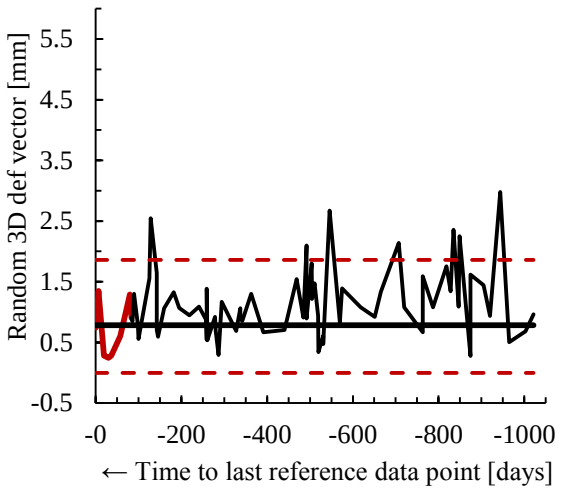
Patient random C5-C7 deformation



Patient systematic C7-caudal deformation



Patient random C7-caudal deformation



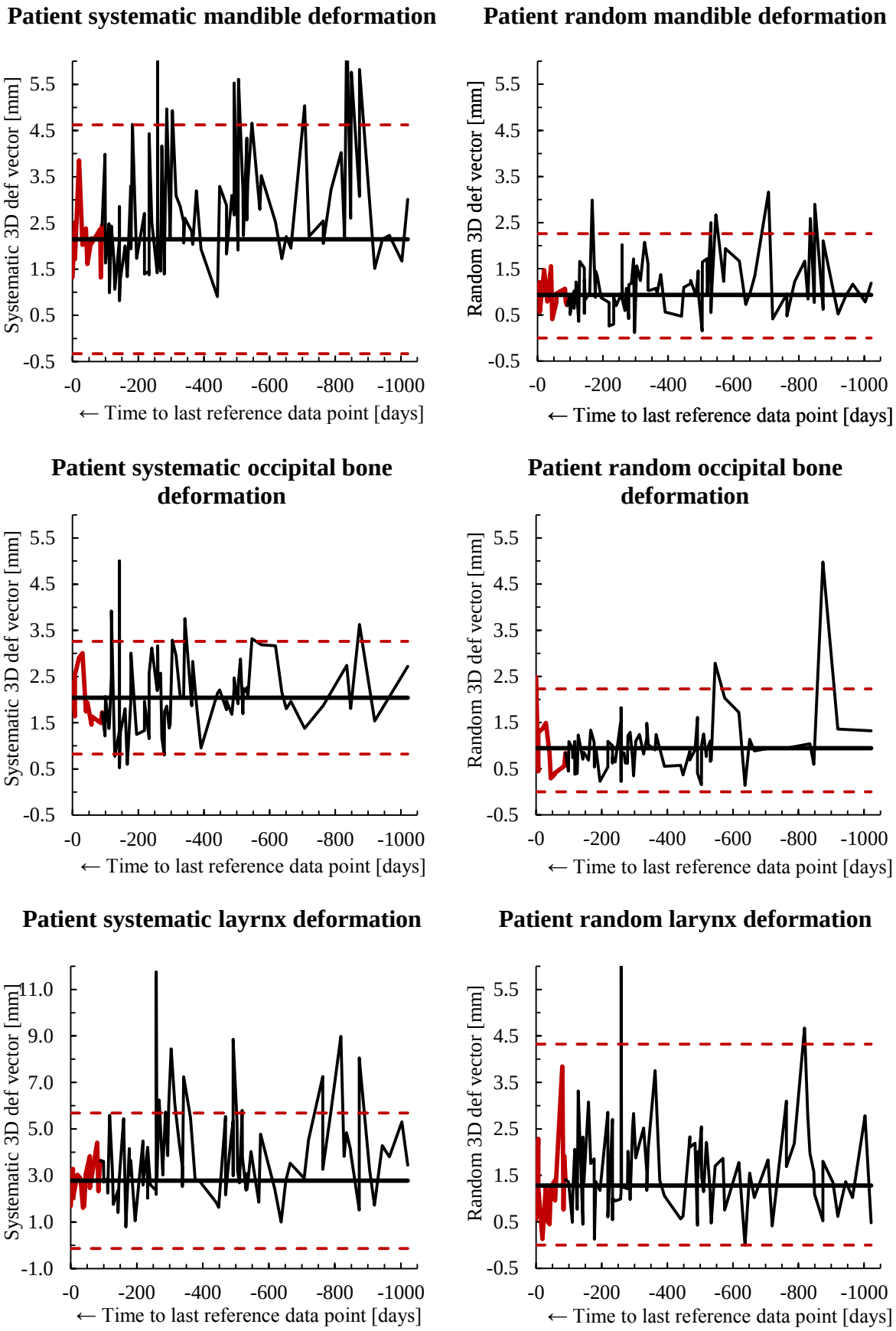
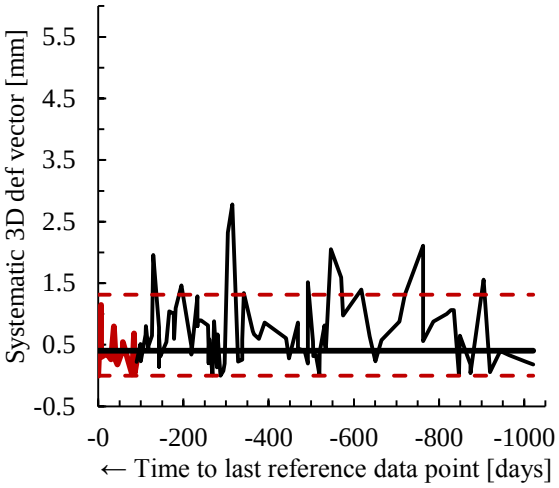


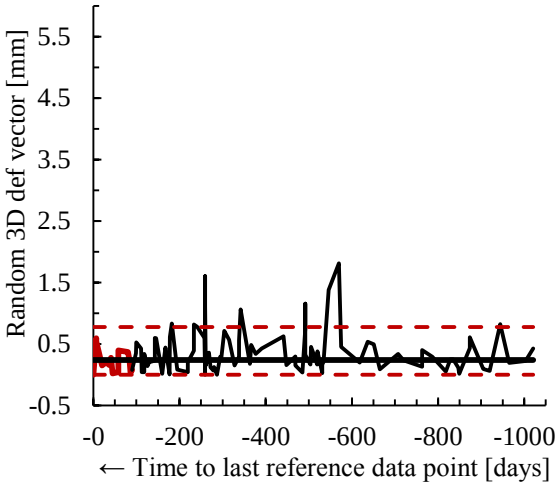
Figure A.1 – *x*-SPC charts for patient systematic and random deformation per ROI relative to C1-C3. Control group plotted in red, study group shown in black. Centre line charted in black; upper and lower process limits shown as dotted red lines.

APPENDIX B: Moving range SPC charts per ROI

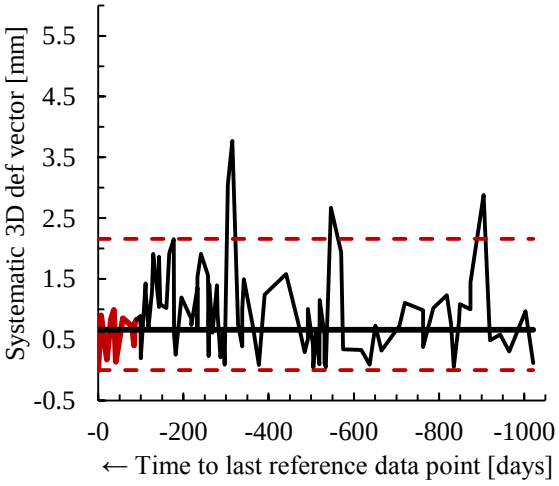
Patient systematic C3-C5 deformation



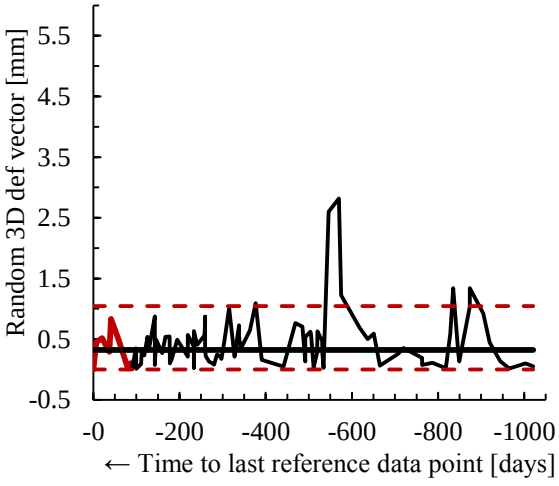
Patient random C3-C5 deformation



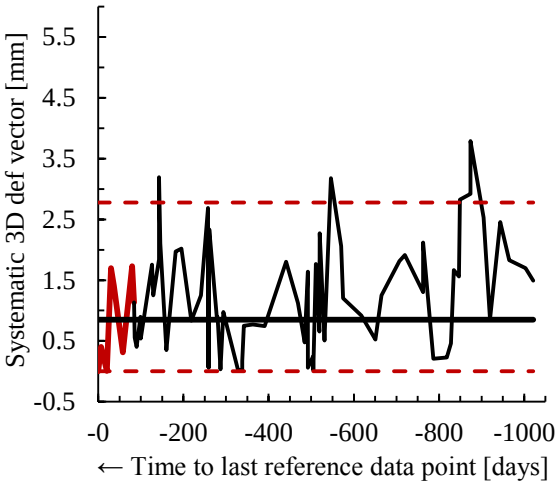
Patient systematic C5-C7 deformation



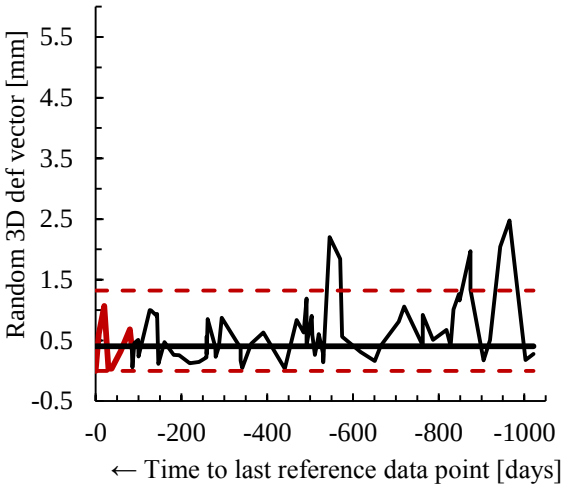
Patient random C5-C7 deformation



Patient systematic C7-caudal deformation



Patient random C7-caudal deformation



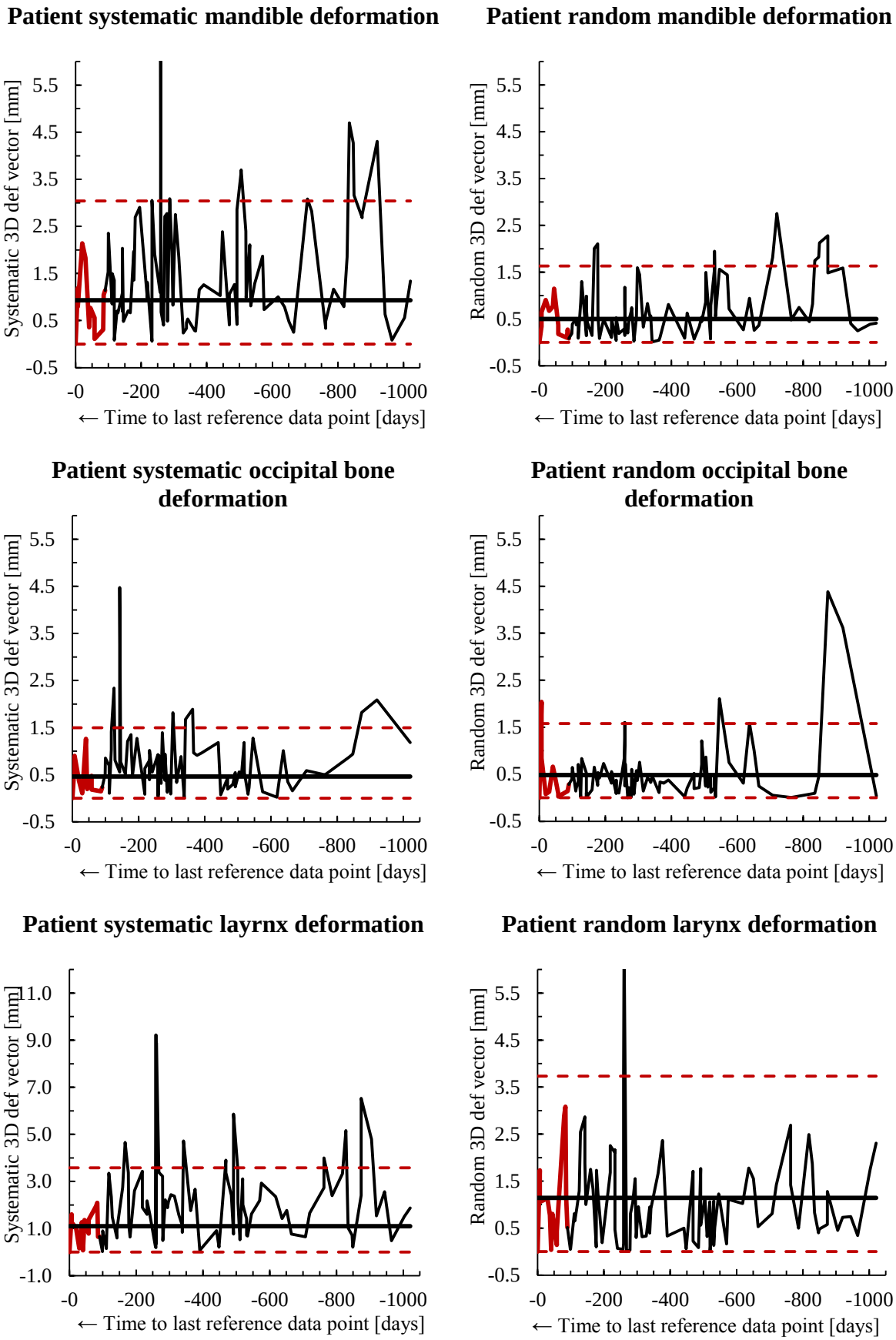


Figure B.1 – *mR*-SPC charts for patient systematic and random deformation per ROI, relative to C1-C3. Control group plotted in red, study group shown in black. Centre line charted in black; upper and lower process limits shown as dotted red lines.

APPENDIX C: Individual value and moving range SPC charts for weight loss

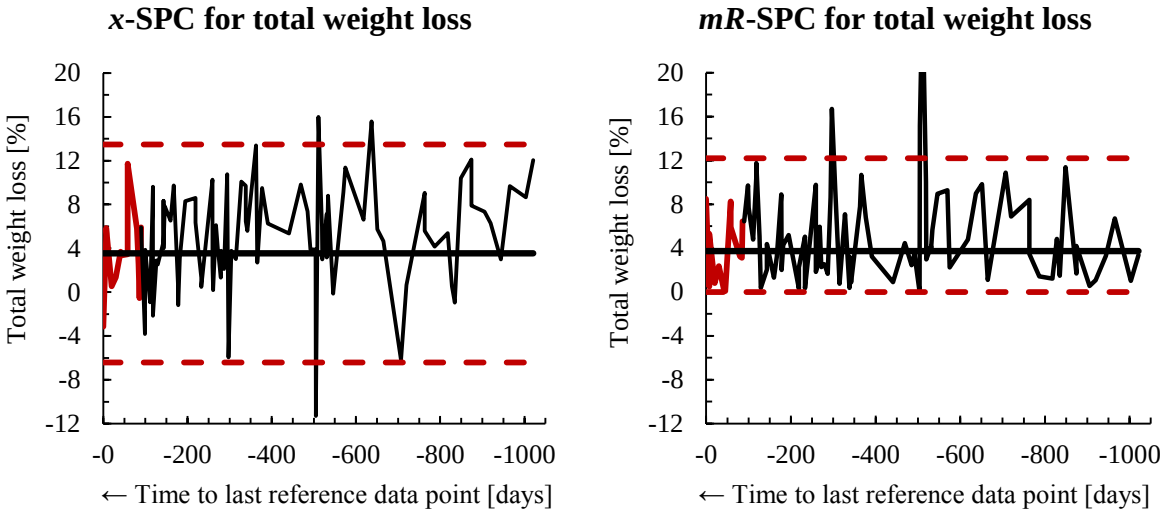
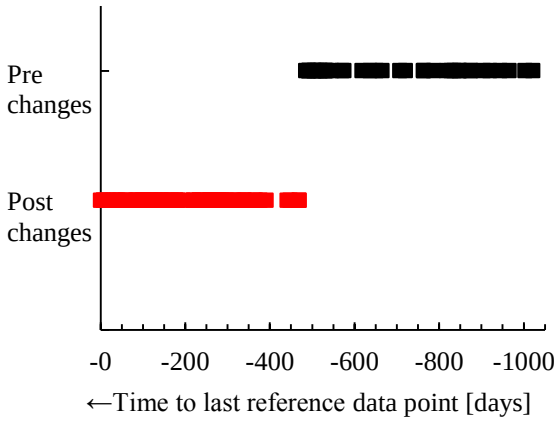


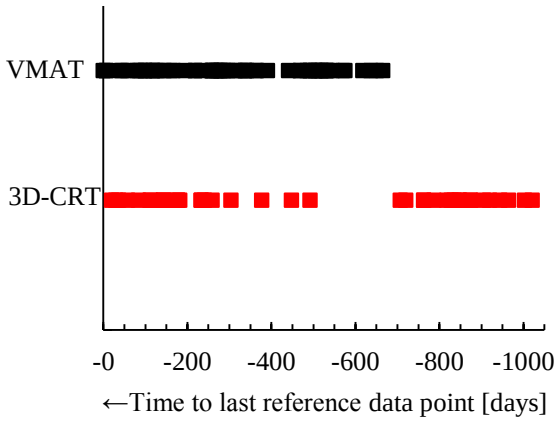
Figure C.1 – *x*-SPC and *mR*-SPC charts for total weight loss. *Control group plotted in red, study group shown in black. Centre line charted in black; upper and lower process limits shown as dotted red lines.*

APPENDIX D: Time plots for categorical variables

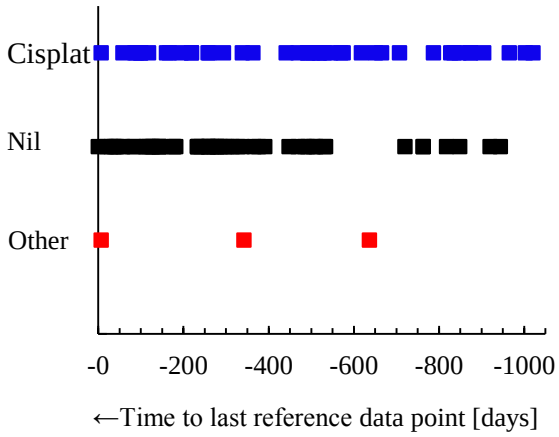
E **MDT changes versus time**



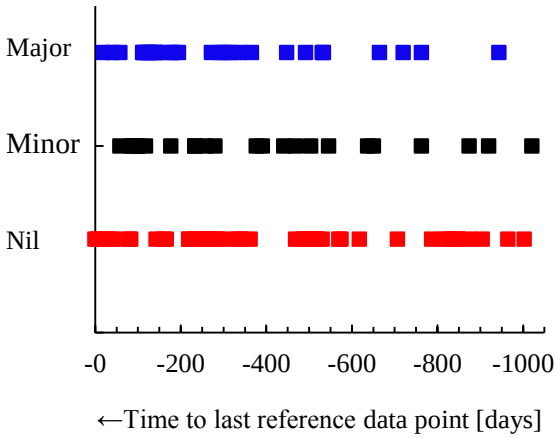
F **Treatment technique versus time**



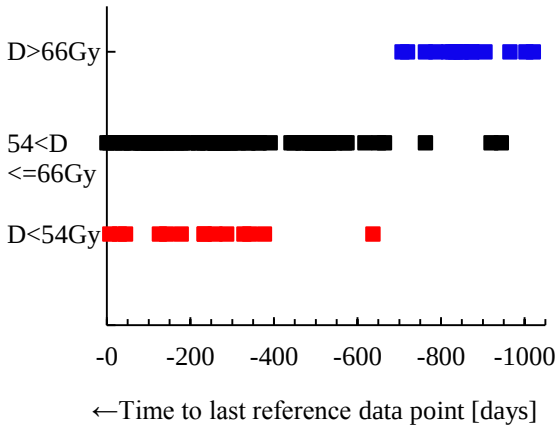
G **Chemotherapy versus time**



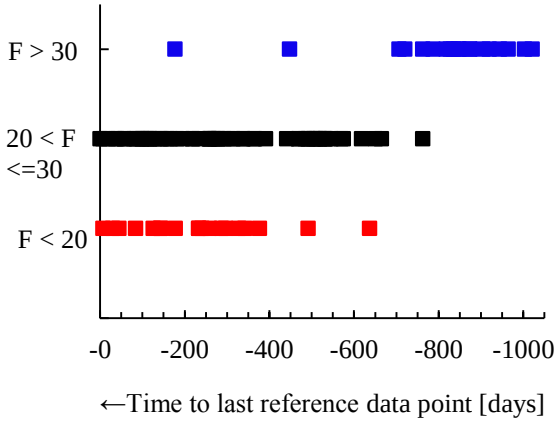
H **Surgery versus time**



I **Total dose versus time**



J **Number of treatment fractions versus time**



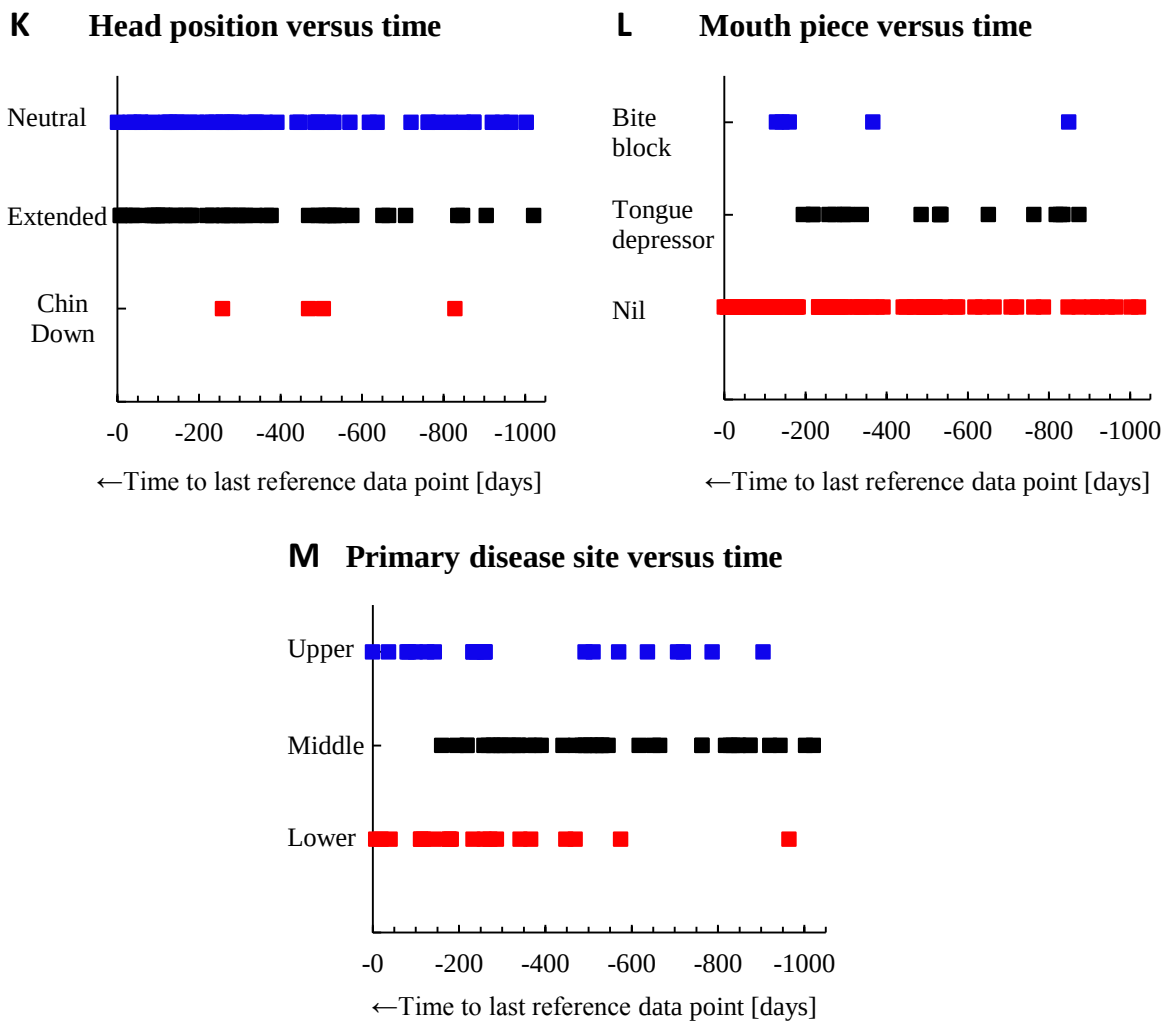


Figure D.1 – time plots E-F for variables not demonstrating changes coinciding with changes in setup accuracy. The x-axis represents the same reverse time base used in SPC analysis; the y-axis represents the different subgroups for each variable. Each individual point on the time plot represents a single incidence for the relevant subgroup variable, plotted against the reference date for that patient dataset. See text for additional information on each variable and subgrouping.