STRATEGIES FOR ADAPTIVE RADIOTHERAPY: TOWARDS CLINICALLY EFFICIENT WORKFLOWS

by

YIANNIS G. ROUSSAKIS

A thesis submitted to the
University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

PSIBS Doctoral Training Centre
College of Engineering and Physical Sciences
University of Birmingham
December 2015

Abstract

Adaptive radiotherapy (ART) aims to adapt the treatment plan to account for inter-fraction anatomical variations, based on online acquired images. However, ART workflows are not −yet− routinely used in clinical practice, primarily due to the dramatic increase of the workload required and the inadequate understanding of optimal methods to maximise clinical benefit. This thesis reports on investigations of procedures for the automation of the ART process and the identification of optimal adaptation methodologies.

Investigated auto-segmentation algorithms were found insufficient for an automated workflow, while a hybrid deformable image registration (DIR), incorporating both intensity-based and feature-based components, revealed the most accurate and robust performance. An evaluation method was proposed for interfraction treatment monitoring through dose accumulation following DIR. The robustness of several treatment methods to observable anatomical changes were investigated, highlighting cases whereby substantial dosimetric consequences may arise.

Offline ART workflows were explored, specifically investigating the effects of treatment monitoring frequency, adaptation method (simple re-plan or re-optimisation addressing cumulative dose), and adaptation timing. Contrary to simple re-planning, re-optimisation demonstrated its ability to compensate for under-/over-dose, however, non-uniform dose distributions and hot-spots may be generated. Therefore established planning techniques are applicable for re-planning while advanced approaches are required for treatment re-optimisation accounting for radiobiological consequences.

To my parents,

George and Domna Roussakis.

Acknowledgements

Foremost I would like to express my utmost gratitude to Dr Gareth Webster, my primary supervisor and mentor, for his invaluable contribution to my studies, the assistance and guidance for both this project and my development as a medical physicist.

I am grateful to Dr Hamid Dehghani, my second supervisor, for enabling the project, always being keen to advise, guide and propose new directions to this work. I feel fortunate I even met Prof. Stuart Green who served as my third supervisor. He would always find time in his busy schedule to discuss and advise on my project, while, his dedication for the advancement of radiotherapy has been a true inspiration.

A big thanks to the radiotherapy team at the Queen Elizabeth Hospital Birmingham for their support and advise. In particular Trevor Williams who provided helpful training in treatment planning and assisted in the work presented in Chapter 7.

I would also like to thank all colleagues, staff and academics of the PSIBS Doctoral Training Centre for the friendship, support and education provided.

Special thanks to my other half, Antonia Paraskeva, who offered endless support, love and encouragement. Lastly, I would like to thank my parents and family for their unconditional love and support for longer than I can remember.

This work was financially supported by the engineering and physical sciences research council (EPSRC) grant EP/F50053X/1 through a studentship from the PSIBS Doctoral Training Centre at the University of Birmingham.

Thank you!

Intelligence is the ability
to adapt to change.

– Stephen Hawking

Contents

List of Figures

List of Tables

List of Acronyms

3D	three-dimensional
ABS	atlas-based segmentation
ART	adaptive radiotherapy
BED	biologically effective dose
CBCT	cone-beam computed tomography
CI	conformity index
CNR	contrast-to-noise ratio
CRT	conformal radiotherapy
CT	computed tomography
CTV	clinical target volume
DDA	deformable dose accumulation
DICOM	digital imaging and communications in medicine
DIR	deformable image registration
DNA	deoxyribonucleic acid
DRR	digitally reconstructed radiograph
DSC	Dice similarity coefficient
DVF	dynamic vector field
DVH	dose-volume histogram
EPID	electronic portal imaging device
EUD	equivalent uniform dose

GTV gross tumour volume

H&N head and neck

HU Hounsfield unit

ICRU international commission on radiation units & measurements

IGRT image guided radiotherapy

IMAT intensity modulated arc therapy

IMPT intensity modulated proton therapy

IMRT intensity modulated radiotherapy

ITV internal targeet volume

kV kilo-voltage

LCV low contrast visibility

LDPE low-density polyethylene

LKB Lyman-Kutcker-Burman

LQ linear-quadratic

MBS model-based segmentation

MDC mean distance to conformity

MLC multi-leaf collimator

MVCT mega-voltage computed tomography

NPC nasopharingeal carcinoma

NTCP normal tissue complication probability

OAR organ at risk

PRV planning organ at risk volume

PTV planning target volume

QA quality assurance

RBE relative biological effectiveness

RIR rigid image registration

ROI region of interest

RTD radiotherapy dose

RTP radiotherapy plan

RTS radiotherapy structure

SNR signal-to-noise ratio

SPICE smart probabilistic image contouring engine

TCP tumour control probability

TNM tumour, nodes, metastases

TPS treatment planning system

UI uniformity index

VMAT volumetric modulated arc therapy

XVI x-ray volumetric imaging

List Of Publications

The following publications arose from work performed as part of this PhD thesis.

Papers in peer-reviewed journals

Roussakis YG, Dehghani H, Green S, Webster GJ. Validation of a dose warping algorithm using clinically realistic scenarios. Br J Radiol. 2014;88:20140691.

Roussakis YG, Zhang R, Heyes G, Webster G, Mason S, Green S, et al. Cherenkov emission portal imaging during CyberKnife® radiotherapy. Phys Med Biol. 2015;60:N419-N425.

Conference presentations

Roussakis YG, McWilliam A, Hartley A, Sangera P, Benghiat H, Hickman M, et al. Evaluation of multiple auto-segmentation solutions against inter-observer variability. Radiother Oncol. 2015;115(Suppl. 1):S507-508. (3rd ESTRO Forum: Poster presentation)

Roussakis YG, Zhang R, Heyes G, Webster G, Mason S, Green S, et al. Real-time Cherenkov portal imaging during radiation therapy. (UKRO 2015: Oral presentation)

Roussakis YG, Williams T, Sanghera P, Hartley A, Dehghani H, Green S, Webster GJ. Plan robustness to anatomical changes: protons (IMPT) vs photons (VMAT, TomoTherapy). (UKRO 2015: Oral presentation)

Roussakis YG, Dehghani H, Green S, Webster GJ. A framework for the validation of actual delivered dose estimation strategies. Radiother Oncol. 2014;111(Suppl. 1):S93. (ESTRO 33: Poster presentation)

Roussakis YG, Dehghani H, Green S, Webster GJ. Comparison of two strategies to calculate actual delivered dose at a radiotherapy treatment fraction. (UKRO 2013: Oral presentation)

Roussakis YG, Dehghani H, Green S, Webster GJ. Quantitative evaluation of cone-beam CT simulation from CT data. (UKRO 2013: Poster presentation)

Co-author

Williams T, Sanghera P, Hartley A, Dumbill A, Chalkley A, Roussakis YG, et al. An evaluation of IMPT versus rotational IMRT for nasopharyngeal carcinoma. Radiother Oncol. 2015;114(Suppl. 1):48. (5th ICHNO: Poster presentation)

Williams T, Roussakis YG, Sanghera P, Hartley A, Heyes G, Dumbill A, et al. Comparison of intensity modulated proton therapy versus rotational IMRT for nasopharyngeal carcinoma. Radiother Oncol. 2015;115(Suppl. 1):556. (3rd ESTRO Forum: Oral presentation)

Chapter 1

General Introduction

This chapter gives a brief introduction to the basics of cancer, the rationale for external beam radiotherapy and the various treatment modalities. Online image guidance and techniques potentially useful in adaptive radiotherapy workflows are also introduced, focusing on the aspects employed in the thesis. Finally, the scope and layout of the work presented herein are outlined.

1.1 Cancer

Cell division and multiplication occurs in most tissues of the body in a controlled manner, with a balance between new cell formation and cell death. Certain distractions of the DNA (mutations) can lead to uncontrolled cell division and the formation of abnormal offsprings and tissues (tumours). Tumours can be distinguished into benign or malignant. Benign tumours are not able to spread (metastasise) nor invade other tissues, while malignant tumours can, and are therefore characterised as cancerous. Cancer, therefore, refers to a collection of diseases in which some of the body's cells initiate uncontrolled cell division

and formation of abnormal cells or tissues that are able to metastasise and invade other tissues.

Cancer type, staging, and grading

Cancers are typically classified by the type and location of cell they originated from, including:

- **Carcinoma:** Originated from epithelial cells (e.g. skin, tissues that cover internal organs).
- **Sarcoma:** Originated from connective or supportive tissues (e.g. bone, muscles).
- **Leukaemia:** Originated from blood forming tissues (e.g. bone marrow).
- **Lymphoma and myeloma:** Originated from immune system.
- **Blastoma:** Originated from embryonic tissues.
- **Melanoma:** Originated from melanocytes.

Cancer staging describes the size and penetration of the tumour, if or how many lymph nodes it invaded and whether it has metastasised, typically employing the TNM (Tumour, Nodes, Metastases) system for solid tumour staging. Tumour grading, on the other hand, describes the level of tumour cell abnormality observed under microscopic examination as compared to normal cells.

The choice of potential treatment can depend on the type of cancer, its location, staging and grading, as well as the patient's health status. Treatment types may range from surgery, chemotherapy, radiotherapy, hormonal therapy and other options, or a combination of multiple techniques.

1.2 External beam radiotherapy

Radiotherapy is a form of treatment which employs radiation, typically used against cancer. External beam radiotherapy (hereafter radiotherapy) refers to the techniques where radiation, in the form of x-rays, gamma rays or accelerated particles, is delivered from a source outside the patient.

1.2.1 Theory

Radiotherapy aims to deliver adequate radiation dose to the tumour volume, causing enough damage to 'control' −preferably kill− the tumour, while at the same time minimising the dose received by the surrounding healthy tissues. Normal tissue cells are usually more efficient in repairing from radiation damage than cancerous cells, therefore, radiation treatment is typically delivered in a number of small fractions to further minimise their damage, allowing healthy tissue repair.

Radiobiology and radiobiological modelling

Radiobiology is the field that studies the effects of radiation to normal and tumour cells. Mathematical functions are often applied to patient treatment response and late effect assessment results, to enable thorough analysis and modelling of future treatments −termed radiobiological modelling. The probability a given amount of radiation, received by a healthy organ, to cause a certain form of complication is called normal tissue complication probability (NTCP). On the other hand, the probability of eliminating all clonogenic tumour cells at the end of the treatment with a given amount of radiation is called tumour control probability (TCP). As demonstrated in Figure 1.1, such knowledge is of great importance in order to design a radiotherapy treatment which would have a high probability of tumour control while keeping healthy tissue complication probability at an acceptably low level.

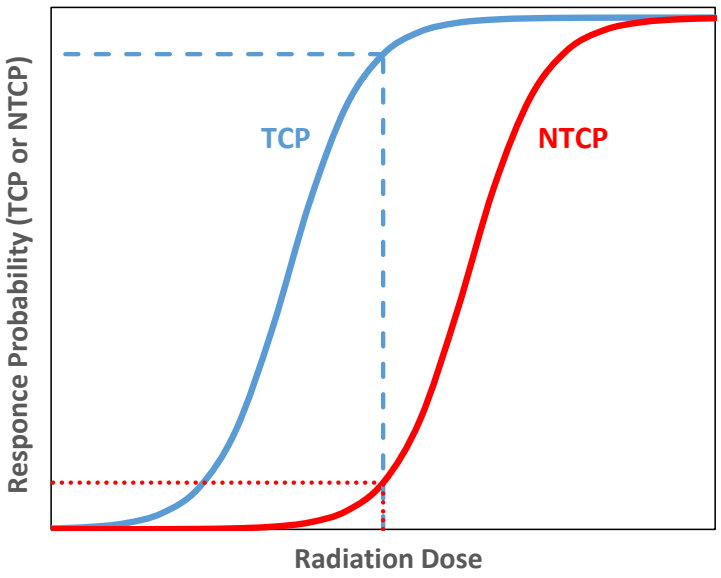

Figure 1.1: Generalised dose-response curve for tumour control probability (TCP) and normal tissue complication probability (NTCP).

One of the most widely used NTCP models is the Lyman-Kutcker-Burman (LKB) model:

$$NTCP = 1/\sqrt{2\pi} \int_{-\infty}^{t} exp(-t^2/2)\ dt \tag{1.1}$$

$$t = \frac{D_{eff} - TD_{50}}{m \cdot TD_{50}} \tag{1.2}$$

$$D_{eff} = \left(\sum_i v_i D_i^{1/n} \right)^n \tag{1.3}$$

where D_{eff} is the dose that, if given uniformly to the entire volume, will lead to the same NTCP as the actual non-uniform dose distribution, TD_{50} is the uniformly delivered dose that would result in 50% chance of complications, v is the fraction of the organ irradiated uniformly, n is the volume effect parameter, and m is a measure of the slope of the sigmoid curve.

The Poisson model with Linear-Quadratic (LQ) modelling of inter-fraction cell survival is a commonly used TCP model:

$$TCP = exp\left\{-N_0 \cdot exp\left[-\alpha D\left(1+\frac{\beta}{\alpha}d\right)\right]\right\} \tag{1.4}$$

where N_0 is the initial number of clonogenic cells, D is the total dose delivered in n equal fractions of size d, α and β are the parameters associated with the initial slope and degree of curvature of the LQ cell survival curve.

Biologically effective dose (BED) modelling is a method that allows the translation of a physical dose delivered to tumour or normal tissue, to the biological effect of this radiation to the corresponding tissue. A simplified BED formula is:

$$BED = nd\left(1+\frac{d}{\alpha/\beta}\right) \tag{1.5}$$

where n is the number of fractions and d the dose per fraction.

Radiotherapy volume definitions

The international commission on radiation units and measurements (ICRU) produced three reports −ICRU 50 [1], ICRU 62 [2], ICRU 83 [3]− which outline recommendations on prescribing, recording and reporting photon beam therapy. These recommendations, summarize the radiotherapy volumes as listed below and shown in Fig. 1.2.

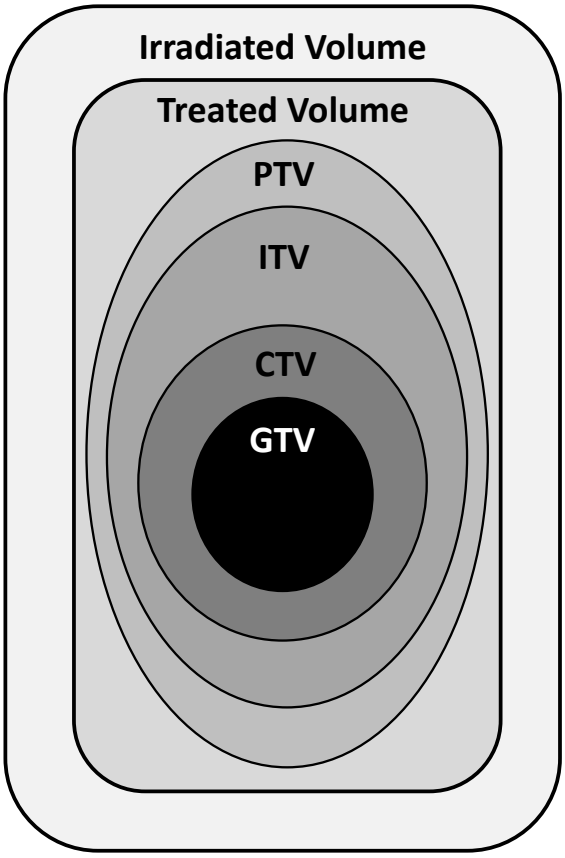

Figure 1.2: Radiotherapy prescription target volumes as defined by ICRU 50, ICRU 62 and ICRU 83 reports.

- Gross tumour volume (GTV): Extent of malignant tumour growth.
- Clinical target volume (CTV): Margin that encompasses sub-clinical microscopic malignant disease.
- Internal target volume (ITV): Margin around CTV to account for internal movement.
- Planning target volume (PTV): Margin that accounts for geometrical variations in the CTV and inaccuracies in treatment delivery.
- Treated volume (TV): The volume receiving the prescribed dose (which can be larger or smaller than the PTV) due to the inability to perfectly conform the dose to the often complex shape of the PTV, caused by limitations of the irradiation technique.

- Irradiated volume (IV): The volume receiving a dose considered significant in relation to normal tissue tolerance.
- Organ at risk (OAR): Healthy tissue whose radiation tolerance may influence a treatment.
- Planning organ at risk volume (PRV): Margin around OAR that accounts for geometric variations and inaccuracies in treatment delivery.

1.2.2 Treatment modalities

Three-dimensional conformal radiotherapy

As previously discussed, the aim of radiotherapy is to deliver adequate dose to the tumour while minimising dose to neighbouring healthy tissues. 3D conformal radiotherapy (3D-CRT) is the term used for those radiotherapy techniques that tailor the dose to the volumetric shape of the target. This dose conformation is typically achieved with the use of multi-leaf collimators (MLC's), which are collimators consisting of multiple individual leaves whose position can be independently changed, enabling arbitrary beam shapes.

Intensity modulated radiotherapy

Intensity modulated radiotherapy (IMRT) is a form of conformal radiotherapy during which the intensity of the delivered beam is fine-tuned (or modulated) across the beam's profile.

Volumetric modulated arc therapy

Volumetric modulated arc therapy (VMAT) or occasionally referred as intensity modulated arc therapy (IMAT) is an IMRT technique whereby radiation is delivered while dynamically rotating around the patient, while dose rate and MLCs also change dynamically.

TomoTherapy®

TomoTherapy® is the name of a commercial product from Accuray® (Sunnyvale, CA) which is capable of delivering rotational IMRT treatments while the treatment couch moves perpendicular to the treatment beam. Effectively, this equipment employs an IMAT technique to dynamically irradiate thin slices of the patient in a helical delivery process.

Intensity modulated proton therapy

The differences between photon and proton beam therapy lie in the fundamental physical interactions each experiences while travelling through matter. As opposed to photons, protons are slowed and eventually stopped, depositing the bulk of their energy at the end of their range −called the Bragg peak, Fig. 1.3. A spread-out Bragg peak can be created by the accumulation of a number of proton beams with different energies. Compared to photon treatments, this can enable efficient sparing of healthy tissues found in front or beyond the target region.

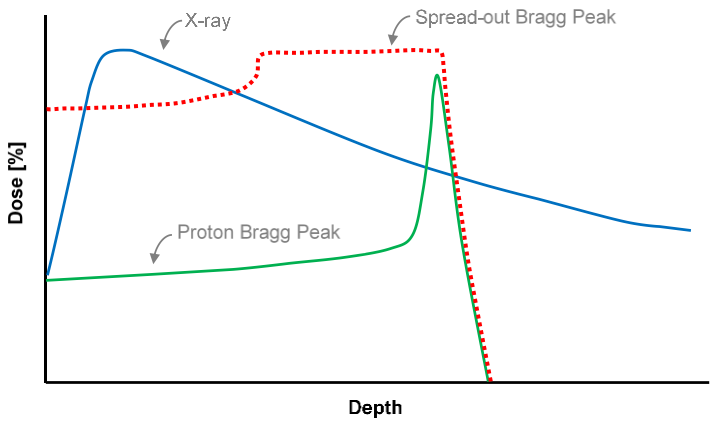

Figure 1.3: Schematic representation of percentage depth dose curves for x-ray photons and protons, showing the characteristic proton Bragg peak and a spread-out Bragg peak being the result of the accumulation of proton beams with different energies.

Intensity modulated proton therapy (IMPT) can be delivered by spot-scanning technique. This technique incorporates thin proton 'pencil' beams that are steered with the use of magnets to 'paint' the dose onto the target region.

1.3 Image guided radiotherapy

Typical radiotherapy practice currently involves the preparation of a treatment plan based on an initial high resolution CT scan of the anatomy to be treated. The patient then undergoes a treatment simulation process during which body or head immobilization devices are custom created if required and marks or tattoos drawn on the skin. The combination of immobilization devices, in-room laser lights and skin marks allows the precise alignment and accurate positioning of the patient for each treatment session. However, since the treatment and patient positioning are optimised for the anatomy on the planning CT image, any changes in patient's anatomy during the treatment course itslef (which may last for up to 8 weeks) could result in a suboptimal treatment. Also, given that laser-tattoo alignment is external, this might not reliably correlate to the position of an internal target at any given instance.

To account for interfraction movements and positioning uncertainties, images of the patient may be acquired immediately prior to the treatment while patient is on the treatment couch, allowing localization of the internal target and repositioning as required, thus ensuring the therapeutic beam is correctly targeted at that time. This process is termed image-guided radiotherapy (IGRT) and can be performed either immediately before treatment or in real-time throughout treatment for dynamic motion tracking. A wide variety of online imaging technologies have been developed over the past decade, including radiation-based systems like electronic portal imaging detectors (EPID), cone-beam CT

(CBCT), mega-voltage CT (MVCT); and non-radiation-based systems like ultrasound, camera-based and electromagnetic tracking based systems [4]. The feasibility of a novel approach for real-time portal imaging, using an optical camera and utilising the Cherenkov emission effect has been investigated as part of this study, with the experimental procedures and outcomes outlined in Appendix A (p. 169) [5].

All these systems, and especially the volumetric imaging techniques which provide 3D visualisation of bony and soft tissue anatomy, enable accurate daily positioning of patients based on soft tissue matching. This daily positioning is typically performed by shifting the treatment couch which only possesses four degrees of freedom, with the exception of certain more sophisticated equipment. It is therefore not possible to correct for certain anatomical changes such as organ deformation and sliding motion using solely simple couch shifts.

1.4 Adaptive radiotherapy

In 1997, Yan et al. [6] proposed the concept of adaptive radiotherapy (ART), suggesting the adaptation of the treatment plan to account for interfraction anatomical variations, based on online acquired images [7]. Such treatment adaptations are currently employed in routine clinical practice when significant anatomical changes are observed, such as substantial weight loss. State-of-the-art ART, on the other hand, aims to regularly monitor the treatment delivery and adapt when necessary (offline ART) [6]; predict the result and alter it before −or during− the treatment of that day (online ART) [8]; or even track the change or movement and adapt in real time (real-time ART). [9; 10]

Each of these implementations comes with its own advantages and limitations. Offline ART can efficiently adapt to gradual changes and can only compensate for random daily

changes, while, online ART can adapt to random changes but when progressive changes are observed many adaptations may be required. *Even though the majority of the tools and concepts discussed in the following sections are applicable to both ART implementations, this thesis will focus on offline techniques.* A flowchart of the major processes towards a potential offline ART workflow is shown in Fig. 1.4.

Interfraction anatomical changes can be random (e.g. level of bladder filling) or gradual (e.g. weight loss) or even a combination of effects. Consequently, optimum treatment monitoring would require frequent dosimetric evaluations and, potentially, multiple treatment adaptations. Certain limitations such as the dramatic increase in workload and lack of reliable automation tools currently prevent efficient routine monitoring of delivered dose throughout the treatment course.

An efficient ART workflow would benefit from robust automatic segmentation solutions which would decrease manual delineation time; fast and accurate deformable image registration solutions that would enable correct point-to-point matching of, say, the planning CT to the daily acquired image for treatment assessment and dose accumulation, while robust evaluation protocols would be required for quality assurance of these procedures. Additionally, further assessment of the clinical benefit and optimization of ART workflows is required.

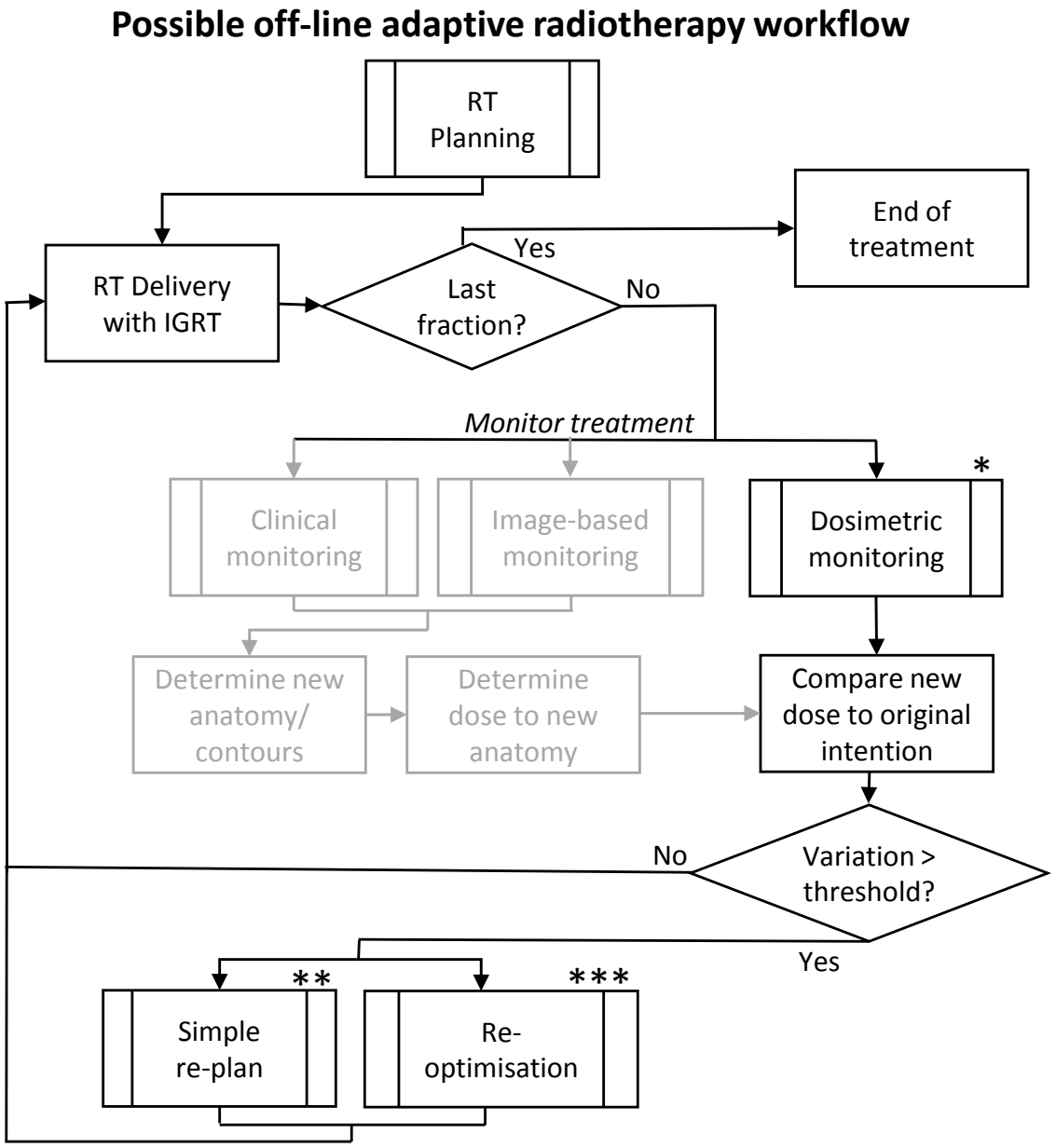

Figure 1.4: Flowchart of the major processes towards a possible offline adaptive radiotherapy workflow, investigated in this thesis.

*: Detailed description and investigation of dosimetric treatment monitoring techniques are outlined in Chapters 5 and 6. Procedures to assist, automate or evaluate steps towards the dosimetric monitoring process are discussed in Chapters 2, 3 and 4.

** & ***: The two treatment adaptation techniques, simple re-plan and re-optimisation, are discussed and investigated in Chapter 8.

1.4.1 Automatic segmentation

Automated segmentation solutions aim to reduce the manual delineation time, increase adherence to contouring guidelines, assist towards the reduction of inter- and intra-observer variability by producing consistent starting point structures and enable further implementation of adaptive radiotherapy procedures by minimising this time-consuming intervention by physicians. Auto segmentation algorithms can be identified in two general categories: those that do not require prior knowledge and those that utilise prior knowledge [11]. The first group contains threshold and active contour based algorithms which are generally used for outer body and bone extraction. On the other hand, atlas-based, model-based, machine learning or hybrid algorithms require prior knowledge and training.

Precise delineation of regions of interest (ROIs) has become more critical with the widespread implementation of highly conformal dose delivery techniques, such as IMRT and VMAT. Furthermore, repeat scanning for plan adaptation and re-optimisation when substantial anatomical changes are observed has been proven advantageous [12; 13]. Regrettably, manual contour delineation is a highly time-consuming process [14], which also suffers from crucial inter- and intra-observer variability [15; 16].

A number of commercial auto-segmentation solutions have become available in recent years, either as standalone software or integrated to treatment planning systems, while clinical investigations evaluated their performance [17–21]. Valentini et al. [22] recently highlighted the importance of such software in radiotherapy and outlined a number of recommendations for the employed evaluation procedures. The performance of several such commercial solutions will be investigated in Chapter 2, p.27.

1.4.2 Deformable image registration

Deformable image registration (DIR) aims to map the location of each individual voxel in a given image to another reference image, while, in contrast to linear rigid image registration (RIR) methods, it handles the registration task with a non-linear approach enabling elastic voxel-to-voxel mapping between images. Following DIR, a voxel-specific 3D mathematical transformation is typically generated, containing information for the elastic mapping of each image element to the reference image, often referred to as the deformation vector field (DVF).

DIR algorithms are considered very important for the automation and wider application of certain adaptive radiotherapy protocols, as the generated DVF can facilitate multiple processes, including:

(a) Auto-segmentation: Certain auto-segmentation approaches use DIR to register an anatomical atlas to a given image and then map the atlas contours to this image.

(b) Contour propagation: DIR can be performed between two images of the same patient, with the manually drawn contours on one image automatically propagated to the other.

(c) Dose mapping: The DVF following DIR between two images can be applied to the dose distribution in order to map this to a reference anatomy, or for the dose accumulation on a single frame of reference (deformable dose accumulation − DDA).

(d) Motion compensation: DIR can be applied to multiple images of the same patient (e.g, real-time 4D imaging) to provide further information on the nature of organ motion.

DIR algorithms can be identified in two main categories, according to their underlying physical basis [23]: those inspired by physical models (e.g. optical flow - Demon's [24],

viscous fluid flow [25]); and those influenced by approximation and interpolation theories (e.g. thin-plate spline [26], B-spline [27]), each of which have their own advantages and limitations. A brief discussion of DIR algorithm components and some advantages and limitations of each are outlined in Appendix B, p. 180. An exhaustive survey was recently produced by Sotiras et al. [28].

Several techniques have been proposed in the literature for performance evaluation of DIR algorithms, using phantoms or patient data. Physical phantoms are tools routinely used for quality assurance in medical imaging and radiotherapy applications in clinical settings and have, therefore, been extensively used for DIR assessment [29–36]. The use of phantoms benefits from their practicality, ease of use and, typically, knowledge of deformation 'ground truth', to which the DIR result can be compared for quantitative evaluation. However, it is practically impossible to model in great detail every anatomic and image acquisition factor resulting in real medical images, using a mechanical phantom. Namely, complex anatomical variations and organ deformations, volume changes due to weight loss or physiological processes such as rectum filling, various image acquisition and reconstruction artefacts prevent phantoms' ability to provide objective means of DIR evaluation.

Patient images have been chosen by many researchers to provide sufficient realism in evaluation conditions and seek the performance limits of the evaluated DIR [19; 37–45]. However, use of clinical data without access to the ground truth of the observed deformation obscures objective quantitative evaluation.

Patient images artificially deformed with known deformations has been proposed as an alternative evaluation method, preserving both access to the ground truth and the relevant image complexity [32; 33; 46] —sometimes referred as digital phantoms. This technique, however, has several limitations. The realism and complexity of the artificial deformations

may introduce bias in the evaluation procedure, if too simple or too complex and physically unrealistic artificial deformations are introduced. Furthermore, it is difficult to properly modify images in order to model different image acquisition parameters and random artefacts, introducing additional sources of uncertainty, especially when attempting to evaluate inter-modality DIR.

The evaluation frameworks can be identified in four main categories:

(a) Landmark tracking: Distance between anatomical landmarks.

(b) Image feature: Comparison of position and shape correspondence of image features.

(c) Contour comparisons: Volumetric or surface comparison of anatomical contours.

(d) Dynamic vector field (DVF): Mathematical evaluation of produced DVF to assess if DIR was physically realistic; comparison of two DVFs (artificial deformations).

Whichever of these approaches is chosen, of great importance is the similarity metric and statistical analysis to be employed. Certain measures are sensitive to local and others to global differences, so it is important to identify the one that is able to highlight the desired differences.

Deformable registration can be performed between images of the same or different modality, such as between a planning CT and an online acquired CBCT, complicating the registration and its evaluation methodology since each imaging modality may possess different characteristics. In this occasion, when physical phantoms are used these can simply be scanned using both imaging modalities in order to obtain the ground truth. Alternatively, if digital phantoms are chosen, a method to resemble, say, a CBCT from a CT image would be required −a technique investigated in Chapter 3, p.42. Intra- and inter-modality DIR are investigated in Chapter 4, p. 54.

1.4.3 Delivered dose estimation

Delivered dose estimation: rights and wrongs

On-line image guidance enables accurate positioning of patients immediately prior to treatment (discussed in Section 1.3, p.9). Three-dimensional image guidance techniques such as cone-beam CT (CBCT), mega-volatge CT (MVCT) and others offer good soft tissue contrast and therefore enable better patient positioning. Moreover, these images can be used for the estimation of potential dosimetric effects due to anatomical changes that can not be compensated by the couch shifts during patient positioning.

The widespread incorporation of such technologies in clinical practice triggered the investigation of dosimetric treatment monitoring and dose accumulation techniques, and subsequently treatment adaptation if required.

The often synonymous terms "dose mapping" [47; 48], "dose deformation" [49] and "dose warping" [50–52] have recently become apparent in the literature as a means of performing inter-fraction treatment monitoring or dose accumulation [53–58]. All three terms usually refer to the process of mapping a dose distribution associated with one image to another image (or frame of reference), typically following DIR between the two images. On the other hand, dose accumulation or deformable dose accumulation is the application of dose mapping of several instances of a patient's treatment to a single frame of reference, followed by a summation or accumulation of these dose distributions. However, a thorough reading of the literature can reveal subtle but important discrepancies in how these terms are practically employed for dose of the day calculation and inter/intra-fraction dose accumulation: one method is to recalculate the dose using a representation of the daily anatomy (usually a CT, CBCT or MVCT scan or even a planning CT deformed to much the daily anatomy as observed in the on-board acquired image) −or each phase of a 4-dimensional scan [53; 59; 60]; a second method (which is later shown to be wrong)

is to directly warp the original dose distribution according to the observed anatomical deformation (i.e. no dose recalculation step involved) [55–58]. This inconsistency observed in the literature is investigated in Chapter 5.1, p.69.

Actual delivered dose estimation using online CBCT scans

Kilovoltage CBCT (kV-CBCT) technology is integrated on many clinical linear accelerators and can be used for dosimetric treatment monitoring. Dose calculations require accurate determination of the relationship between image Hounsfield unit (HU) and material electron density, employing the so-called HU-to-density curves. This is easily achieved with modern CT scanners dedicated for radiotherapy treatment planning procedures. However, the use of kV-CBCT for such calculations is hindered by their lower image quality.

Many authors investigated techniques to overcome the inferior image quality and enable direct dose calculations on kV-CBCT scans [61–64] or a representation of the anatomy in these images [60; 65]. A review of the literature has been performed herein, and the proposed techniques repeated to identify the method of choice to be used in the following studies.

Phantom based HU-to-density curves A common method for the creation of HU-to-density curves for the planning CT images is the use of tissue equivalent phantoms with various regions of known electron density. This procedure involves the CT scan acquisition of the phantom of choice and the association of image HU to the known density from each region of the phantom. Similarly, CBCT scans of a phantom can be acquired with the HU in the acquired images used to create the CBCT specific HU-to-density curves.

Two commercial phantoms have been employed in the literature, the Catphan® (The Phantom Laboratory, Salem, NY) and Gammex RMI® (Gammex 467, Gammex, Middleton WI) [61; 64; 65], creating a separate HU-to-density curve for each choice of scanning

parameters used. The creation of CBCT-parameter-specific curves for each phantom in these studies resulted in IMRT dose calculations with accuracy of 97.5% to 99.0% as compared to dose calculations on planning CT images [61; 64; 65].

The advantage of this method is that the same workflow is applicable for the calculation of HU-to-density curves for both planning CT and CBCT images, while no user-subjectivity is introduced. However, the image quality of CBCT scans is influenced by the size of the imaged object and the scanning parameters used, introducing uncertainties in the choice of best HU-to-density curve.

Bulk density override Density override techniques typically involve the segmentation of a small number of tissues in the CBCT image and the override of their image values with the know HU of certain materials. In one such approach, Fotina et al. [64], performed automatic segmentation based on grey value thresholds, separating the CBCT images into three sections: tissue, air, bone. These sections were then overridden by the HU of water, air and bone, respectively. Their dose calculations revealed slightly better accuracy as compared to the results following phantom based HU-to-density curve generation.

The complexity, and potentially accuracy, of density override techniques can increase with the segmentation of more tissue types. On the other hand, these techniques require more manual interventions or alternatively the reliance on auto-segmentation algorithms.

Patient/tissue based HU-to-density curves HU-to-density curves may be calculated from patient CT and CBCT data [61; 64]. In this approach, volumes of interest are drawn in nearly homogeneous density areas within certain anatomical regions in both CT and CBCT images. Based on the HU-to-density relationships determined from the CT data, the respective CBCT relationships can be derived.

DIR based density mapping Taking advantage of the optimised dose calculations on planning CT quality images, an alternative option is the deformable registration of

the CT image to the online acquired image, here the online CBCT, and the calculation of dose on the registration result ('Dose_Recalc' method in Section 5.1.1, p. 71) [60; 65]. This technique can be fully automated but the dose calculation accuracy would depend on the registration accuracy.

The techniques described above are evaluated and compared in Chapter 5.2, p. 84.

1.4.4 Dose warping for accumulation

Estimation of actual delivered dose during a radiotherapy fraction, can provide important information on the dosimetric consequence of individual inter-fraction anatomical changes. Thorough treatment monitoring would require accumulation of these dose distributions on a single frame of reference, for the identification of cases where treatment adaptation is required. This procedure requires a robust mapping of each anatomical point on the online-acquired image to the reference anatomy and the subsequent mapping of daily dose distributions.

Deformable image registration (DIR) algorithms are essential for the accurate implementation of these processes, considering all degrees of freedom of underlying anatomical changes. The accuracy of DIR algorithms is therefore of critical importance and has been the subject of investigation by several researchers, as discussed in Chapter 4, p. 54.

Dose warping and deformable dose accumulation have been employed in a number of clinical investigations, including a dose feedback technique in adaptive radiotherapy frameworks [66], the assessment of planning target volume (PTV) margins [67] and the examination of parotid gland dose-effect relationships [68], based on dose distributions recalculated on daily or weekly scans and the accumulation on a single frame of reference. Consequently, quality assurance and evaluation techniques have been investigated in order to validate the applicability of this dose warping concept. Previous work

has investigated mathematical models to directly convert DIR errors into dose warping uncertainties, through the use of patient images and mechanical or digital phantoms [38; 47; 48; 69; 70], while a number of deformable dosimetric and non-dosimetric gel phantoms have been produced enabling the experimental evaluation of both DIR and dose warping [49; 50; 52; 71; 72]. Even though some of these studies revealed promising results, they have not convinced the radiotherapy community that these uncertainties are adequately understood [51].

Such studies, and especially the use of deformable dosimeters for the evaluation of dose warping, are very important as they can reveal the 3D dosimetric impact due to uncertainties of a given DIR algorithm. However, they possess three important limitations: Firstly, typical physical dosimetric phantoms present limited image complexity and would not assess the performance limits of the DIR algorithm under evaluation in clinical scenarios. Second, plan delivery, intrinsic dosimetric and dose reading uncertainties are present when using any type of dosimeter in physical phantom measurements. The third limitation is the fact that even where such approaches can offer high precision dosimetric uncertainty evaluation, they cannot directly inform users about the potential extent of those uncertainties in practical clinical cases. These issues are addressed in Chapter 6, p. 92.

1.4.5 Plan robustness to anatomical changes

Implementation of inter-fraction dosimetric treatment monitoring workflows in routine clinical practice requires substantial time and capital investments. Specialised software are required to assist in a semi-automated workflow with subsequent costs and requirements for training of all users. As discussed previously, procedures including automatic segmentation, deformable image registration, delivered dose estimation and dose warping

for accumulation can assist towards an efficient treatment monitoring and subsequently offline ART workflow, nevertheless, physicists and clinicians need to perform thorough evaluations and establish robust quality assurance procedures.

On the other hand, certain clinically observable anatomical changes may visually seem considerable but may not result in a clinically critical dosimetric change. Moreover, the same anatomical change can have a dissimilar dosimetric impact to different treatment techniques, depending on the properties of radiation involved (e.g. photons vs particle radiotherapy) or the delivery method (e.g. step-and-shoot vs rotational vs helical).

It is therefore necessary to investigate the importance and necessity of inter-fraction treatment monitoring and subsequently treatment adaptation for individual cohorts of patients and treatment techniques. In some of these cases, simple adjustments, considerations or additional constraints to the original plans could make them more robust to the most common anatomical changes. For the cases were no further optimisations are available, dosimetric treatment monitoring and treatment adaptations can be beneficial.

These issues are addressed in Chapter 7, p. 107, were the robustness of several treatment plans is assessed under clinically realistic anatomical changes.

1.4.6 Offline adaptive radiotherapy implementations

Offline ART aims to monitor treatment and if a substantial patient-specific change occurs, adapt the plan for the following fractions. For the implementation of such workflows three steps need to be decided, based on the available tools, online imaging technique and other resources. These are: (a) the type of observation that would trigger the need for treatment adaptation, if any (i.e. anatomical or dosimetric), (b) the details of the re-planning procedure and (c) adaptation timing and frequency.

Observation

Observation of anatomical change that would trigger plan adaptation or just further investigation can be acquired in three ways:

- ***Clinical observation:*** When a change is observed or suspected clinically, as for example substantial weight loss or change of a superficial or palpable tumour.
- ***Observation on image:*** When a change is observed on a repeat CT scan or an online acquired image [12].
- ***Dosimetric triggering:*** When a dosimetric change is observed based on recalculation of the original plan on a repeat CT or online acquired image [73]. This dosimetric trigger may come from a single dosimetric evaluation or following accumulation of 'daily doses' on a single frame of reference.

Re-planning methodology

For the generation of a new plan one of two options is to be chosen:

- ***Simple re-plan:*** New plan creation based on new anatomy, aiming to make subsequent fractions deliver the intended dose without considering what dose was actually delivered up to the point of adaptation.
- ***Re-optimisation addressing cumulative dose:*** Re-plan with the same endpoint dosimetric targets as the original plan while addressing the actual delivered dose up to the point of adaptation, to compensate for potential underdose of target regions and/or overdose of organs at risk.

When and how often to adapt

Action levels and frequency of adaptation may be affected by both the clinical significance of observed differences and the available resources. The plan might only be adapted when

substantial change has been observed, or, it may be adapted as many times as required. However, in cases of a clinical trial environment, a research study [74] or a hospital with limited resources, the time and frequency of adaptation may need to be pre-defined.

The influence of each of these steps to actual clinical scenarios of prostate cancer treatments is investigated in Chapter 8, p. 130.

1.5 Scope and layout of work

Inter- and intra-fraction anatomical changes during the course of external beam radiotherapy can lead to sub-optimal treatment outcomes. Adaptive radiotherapy workflows can restore or re-optimise the treatment based on the observed anatomical changes. However, ART is not −yet− routinely used in clinical practise. Some of the main reasons that hinder wider implementation of ART workflows are: the dramatic increase of workload required; lack of robust and reliable means of automation; insufficient understanding of uncertainties involved in certain procedures and methods of assessing these uncertainties; and inadequate understanding of the optimal methods to maximise clinical benefit. The aim of this thesis is to assess the steps involved towards a clinically efficient offline ART workflow, while investigating further evaluation techniques.

Chapter 2 evaluates several commercial auto-segmentation solutions against inter-observer variability in an attempt to identify their fitness for incorporation into a semi-automated ART workflow without or with minimal manual intervention.

Chapter 3 evaluates a tool which allows the simulation of CBCT images from CT datasets, which would be a valuable tool for the evaluation of certain procedures towards an ART workflow; such as auto-segmentation on CBCT scans, DIR between CT and CBCT images and others.

Chapter 4 outlines the evaluation of DIR algorithms in both CT-to-CT and CT-to-CBCT registration, while further discussion of DIR algorithms is outlined in Appendix B. Such algorithms are of great importance in offline ART workflows as they can facilitate auto-segmentation and inter-fraction dose accumulation by determining voxel-to-voxel correspondence between two images.

Chapter 5 investigates an important inconsistency observed in the literature regarding the use of DIR algorithms for the estimation of actual delivered dose at a radiotherapy fraction and examines methods for the calculation of daily delivered dose using on-line acquired CBCT scans. A method for the validation of dose warping is proposed in Chapter 6 and used for the evaluation of a commercial algorithm. The dosimetric consequences of observed anatomical changes can be reliably inspected with accurate dosimetric monitoring and allow for informed decision regarding potential treatment adaptation.

Chapter 7 examines the robustness of several treatment techniques to clinically observable anatomical changes, in an attempt to identify the combinations of treatment techniques and anatomical changes more likely to result in clinically important dosimetric alterations.

The findings of previous chapters are incorporated in Chapter 8 which explores workflows towards offline ART, concentrating on the investigation of the effects of treatment monitoring frequency; the particular method of plan adaptation (i.e. simple re-planning or re-optimisation accounting for cumulative dose); and adaptation timing.

Finally, Chapter 9 outlines the general outcomes and conclusions of this doctoral thesis, along with discussion on potential future directions.

Supporting work performed during the PhD period can be found in appendices. Appendix A describes the investigation into the feasibility of real-time Cherenkov emission portal imaging during radiation therapy using an optical camera [5]. This technique is

of particular importance in radiotherapy techniques that lack alternative portal imaging options, such as Cyberknife®.

Appendix C outlines the experimental investigation for the production of a patient-specific 3D printed phantom.

Chapter 2

Automatic Segmentation

This chapter outlines the evaluation of multiple commercial solutions for automatic segmentation in head-and-neck and male pelvis anatomy against inter-observer variability. The preliminary results of this study have been presented at the 3rd ESTRO Forum [75] and UKRO 2015 conferences.

Manual segmentation of regions-of-interest is a highly time consuming process performed by clinicians before treatment planning can proceed. Adaptive radiotherapy often involves the preparation of a new plan on a newly acquired CT requiring a new manual segmentation. Accurate and robust auto-segmentation solutions could dramatically reduce manual delineation time, therefore, enabling wider clinical implementation of ART.

2.1 Evaluation of multiple auto-segmentation solutions

This study aimed to quantitatively evaluate the accuracy of several commercial auto-segmentation algorithms for head-and-neck (H&N) and male pelvis anatomies −on CT

datasets− and compare them against inter-observer variability in terms of spatial accuracy, to identify their suitability for clinical use without or with minimal manual intervention.

2.1.1 Methods

Data selection

Twenty four radiotherapy planning CT scans, 12 of H&N and 12 of prostate cancer patients, were randomly selected for the evaluation of several auto-segmentation solutions. Gold-standard contours for use in this study were drawn by a senior site-specialist consultant. A further 30 H&N and 15 prostate patient planning CT datasets were selected, the main regions of interest (ROI) contoured and reviewed by a senior site-specialist consultant, for use as the atlas datasets for the atlas-based auto-segmentation algorithms used in this study. All CT scans were acquired on a Philips Brilliance Big Bore scanner and had a slice thickness of 3 mm while H&N patient scans were contrast enhanced.

Auto-segmentation algorithms

Five different auto-segmentation algorithms from four commercial software solutions have been evaluated in this study:

(a) OnQ rts 2.0 (OSL, Shrewsbury, UK)

(b) Atlas-based segmentation in RayStation 4.5 (RaySearch, Stockholm, Sweden) (RayStation ABS)

(c) Model-based segmentation in RayStation 4.5 (RayStation MBS)

(d) Smart probabilistic image contouring engine (SPICE) in Pinnacle^3 9.8 (Philips, Madison, WI)

(e) AutoSegment in MultiPlan 5.0 (Accuray, CA)

The auto-segmentation module in OnQ rts employs an atlas-based framework with user defined atlases. The algorithm picks the best matching atlas based on mutual information criterion [76], and specifically by comparing the digitally re-contracted radiographs (DRRs) of the image and each atlas. Rigid followed by deformable image registration (DIR), using an intensity-based 'Demons' algorithm [24; 32; 77; 78], is then performed between the atlas and the CT scan to be auto-contoured, with the organ outlines transferred to the new anatomy after a post-processing step. The Demon's DIR algorithm is based on the Maxwell's Demons model of the diffusion process. The image boundaries are seen as membranes through which the image diffuses under the influence of demons forces, utilising image-intensity similarity metrics.

RayStation offers both atlas-based and model-based segmentation algorithms. The atlas-based algorithm requires user defined atlases, which for this work were identical to those in OnQ rts. This algorithm identifies the best matching atlas through rigid registration. Deformable registration (DIR) is then performed between the chosen atlas and the CT image. The deformation matrix computed during the DIR process is then applied to the contours associated with the atlas dataset and mapped to the CT image.

In the model-based RayStation algorithm a number of organ models are offered by the software. Each shape model has been created by an average of 10-50 expert drawn contours while the main shape variations have been identified by statistical analysis. The algorithm adjusts the models to match the underlying anatomy using rigid and deformable registration, without violating the shape constraints included in the model. Following model-based auto-segmentation, the user may manually add 'hint contours' in order to drive the segmentation in areas of poor result (RayStation MBH_h).

SPICE algorithm combines atlas-based and model-based segmentation, with non user-defined atlases and multiple steps of DIR. Two pipelines are employed, one for thorax and H&N [79–81] and another for abdominal and prostate anatomies. It should be noted that SPICE produces multiple contours for certain organs, namely three sets for parotids, and two for prostate and rectum. For the purpose of the evaluation, the best matching set for each patient case was chosen.

AutoSegment in MultiPlan is a semi-automatic model-based algorithm, with manually defined initialisation points required for the auto-segmentation to proceed. The user is asked to place a minimum number of initialisation points in specific areas to identify the area covered by each organ. The algorithm copies the pre-defined organ models to the image, based on the location of the initialisation points and adjusts the model based on these points and surrounding image intensity values. It does not currently support H&N auto-segmentation and was therefore only evaluated here for the male pelvis region. The required initialisation points used here were defined by a specialist radiographer.

It should be noted that parameters such as registration resolution in all ABS algorithms and flexibility of models in MBS algorithms, which can be adjusted by the user, were kept to the default values.

As the main purpose of auto-segmentation procedures in clinical practice are labour and time saving, the manual procedures required for MultiPlan and optimisation of results in RayStation MBS_h were kept to a minimum. Specifically, for auto-segmentation in MultiPlan, 14 initialisation points were defined in prostate and 12 in seminal vesicles, requiring 1-2 minutes. On the other hand, addition of hint contours following model-based segmentation in RayStation was limited to bladder and rectum in the male pelvis region, and spinal cord in the H&N region requiring 2-3 (mean: 2.5) and 2-5 (mean: 4) minutes, respectively.

The datasets were also manually re-contoured by two additional observers to quantify inter-observer variability against the 'gold-standard'. The observers were all site-specialists, with experience ranging from senior consultants to experienced registrar and an experienced radiographer, all of whom are trained and responsible for manual contouring in the clinical setting at our institution. To strike a balance between relevance and operator time, manual re-contouring was limited to the most critical regions of interest in each investigated anatomy, namely brainstem, spinal cord and parotids in H&N, and bladder, prostate and rectum in male pelvis.

Evaluation

The evaluation of auto-segmentation results was performed by calculating the absolute volumetric differences, the volume-wise and slice-wise Dice similarity coefficient (DSC) and mean distance to conformity (MDC) indices, considering the original manually drawn contours as the reference. DSC is defined as:

$$DSC = \frac{2|V_A \cap V_R|}{|V_A| + |V_R|} \tag{2.1}$$

where V_A is the auto-segmented volume and V_R is the reference volume (Fig. 2.1). DSC can take values between 1 and 0, with 1 revealing perfect overlap and 0 indicating no overlap at all between the two contours. Good agreement is generally considered for values between 0.6 and 0.8, while values above 0.8 indicate near excellent agreement [82]. MDC is defined as the mean distance that all outlying points in V_A must be moved in order to achieve perfect conformity (overlap) with V_R [83] and is measured in units of distance, i.e. mm, (Fig. 2.1). The volume-wise analysis considers the contours as three-dimensional volumes while slice-wise analysis only compares the underlying volumes in each slice where both are present, ignoring mismatch in the z-direction. Volumetric analysis was performed using ImSimQA software (Oncology Systems Limited, Shrewsbury, UK).

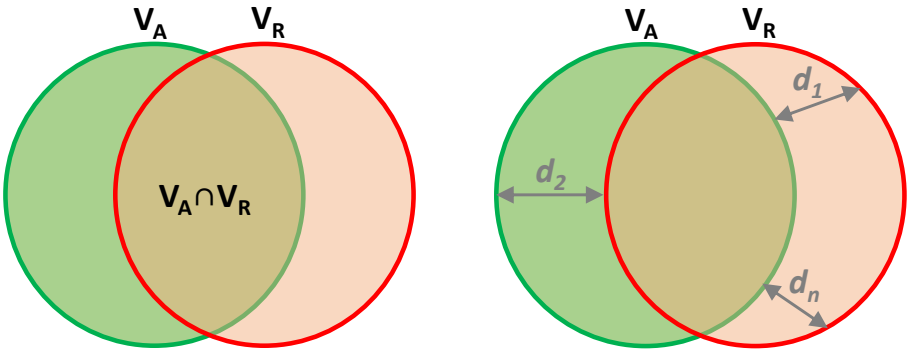

Figure 2.1: Schematic diagram demonstrating how the Dice simularity coefficient (DSC) and mean distance to conformity (MDC) are calculated.

The two-sided non-parametric Wilcoxon rank-sum test has been employed for examination of statistical significance in the slice-wise MDC differences. Mean inter-observer MDC was compared against each individual algorithm as well as against the mean MDC of atlas-based (OnQ rts, RayStation ABS), model-based (RayStation MBS, SPICE) and semi-automatic model-based (RayStation MBS_h, MultiPlan) algorithms. Similar algorithms were also compared against each other.

2.1.2 Results

H&N patients

Figure 2.2 shows the mean ($\pm$standard deviation) volume-wise DSC and MDC absolute volumetric differences for the commercial algorithms under evaluation, and inter-observer variation across all 12 H&N datasets.

The slice-wise MDC for the H&N ROIs across the full cohort, together with statistical significance of observed differences against mean inter-observer variability are shown in Figure 2.3.

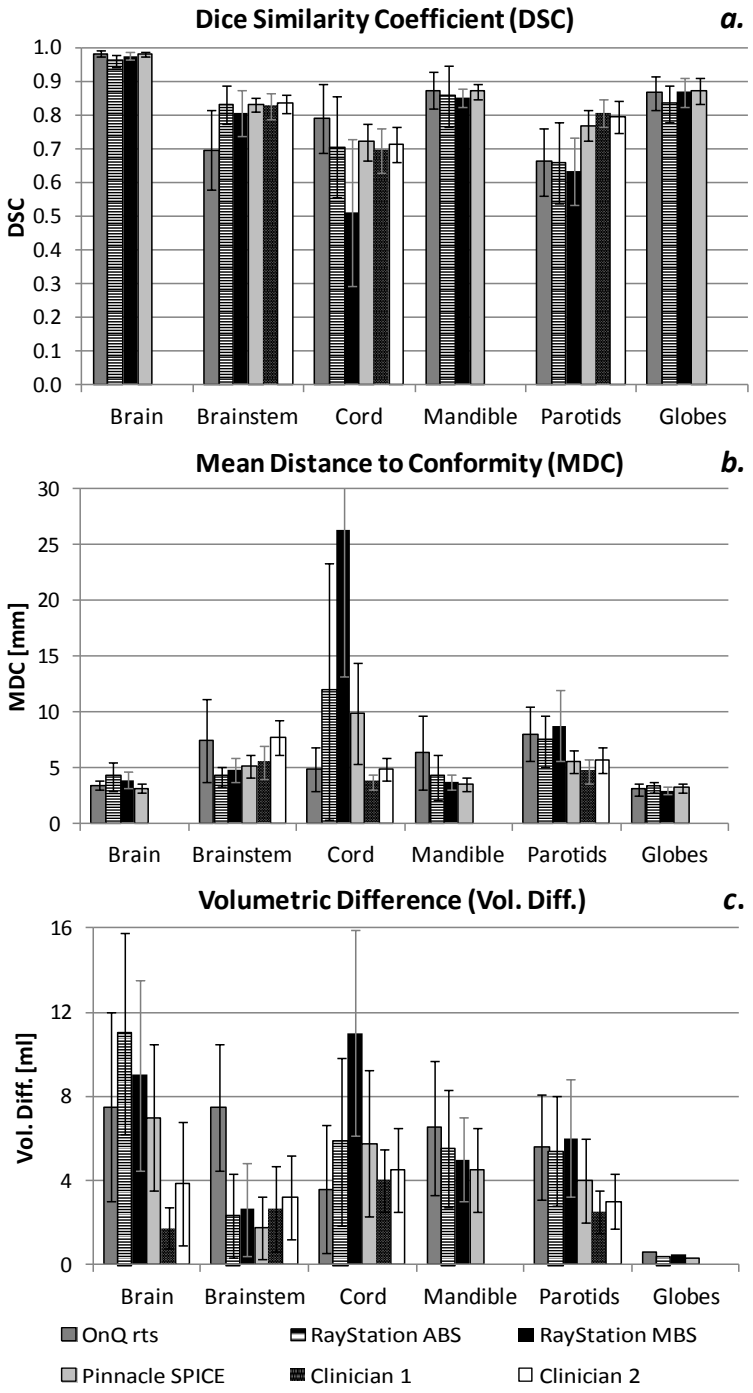

Figure 2.2: ***a.*** Mean and standard deviation of Dice similarity coefficient (DSC), ***b.*** mean distance to conformity (MDC), ***c.*** absolute volumetric differences for head-and-neck region.

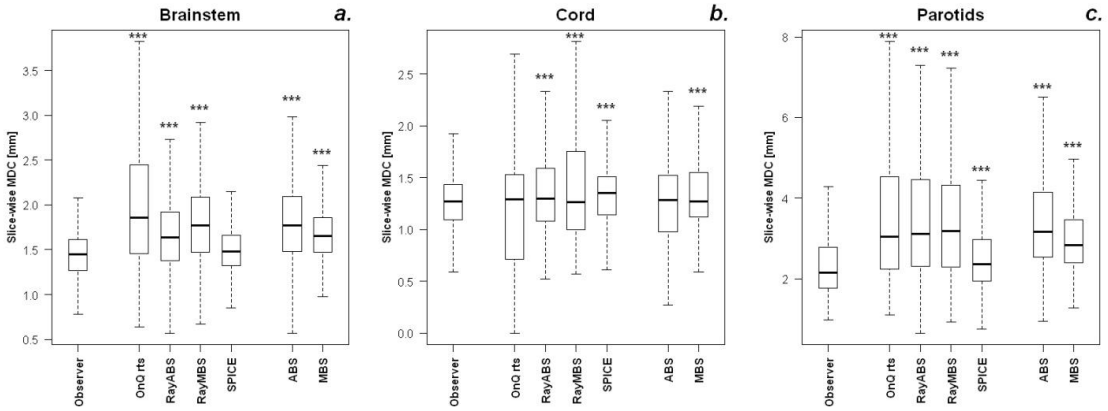

Figure 2.3: Mean distance to conformity (MDC): mean inter-observer variability (Observer) against individual algorithms and mean of atlas-based (ABS) and model-based (MBS) algorithms for ***a.*** brainstem, ***b.*** spinal cord, and ***c.*** parotids. * denotes P-values following comparison against inter-observer variability (*:$P \leqslant 0.05$; **:$P \leqslant 0.01$; ***:$P \leqslant 0.001$)

Figure 2.4 shows the slice-wise MDC histograms as calculated for brain, brainstem, spinal cord, parotids and mandible contours, respectively. These histograms reveal the percentage of slices in which a given MDC was observed throughout all test cases. They are presented in logarithmic scale while only revealing the observed errors in 90% of the examined slices in order to highlight the most relevant differences between the evaluated algorithms in terms of likely subsequent manual correction. Figure 2.4d includes an additional curve which represents the improvement seen after the introduction of hint contours following RayStation MBS.

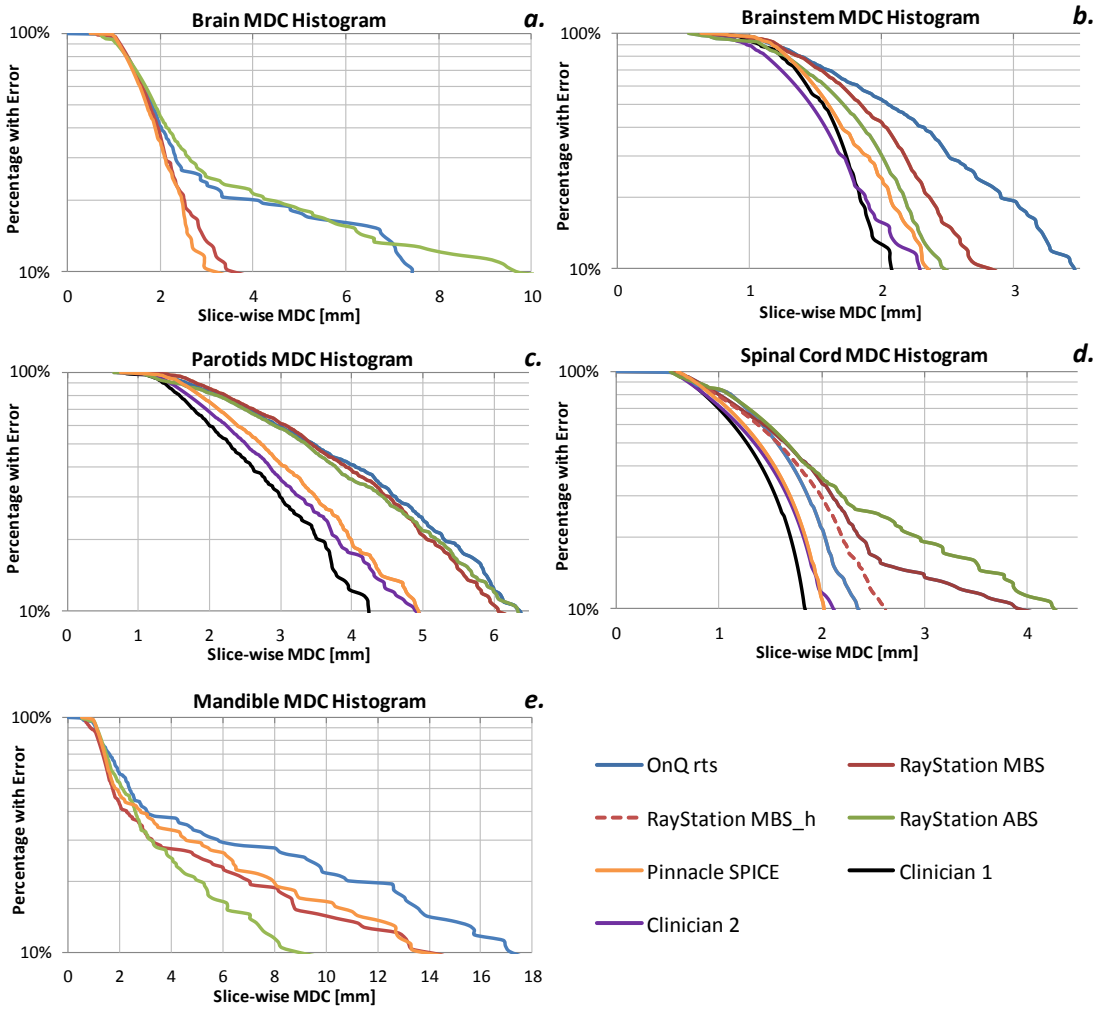

Figure 2.4: Slice-wise mean distance to conformity (MDC) histograms for ***a.*** brain, ***b.*** brainstem, ***c.*** parotids, ***d.*** spinal cord, and ***e.*** mandible, across all test cases.
(e.g. for brain the histogram shows that 80% of slices gave an MDC of $\leqslant$ 4 mm for OnQ rts and RayStation ABS, with MDC of $\leqslant$ 3 mm for RayStation MBS and SPICE.)

Prostate patients

Figure 2.5 shows the mean ($\pm$ standard deviation) volume-wise DSC, MDC and absolute volumetric differences for the commercial algorithms under evaluation and inter-observer variation, for all 12 male pelvis datasets. The slice-wise MDC and statistical significance against inter-observer variation for the male pelvis ROIs are shown in Figure 2.6. Figure 2.7 shows the slice-wise MDC histograms as calculated for bladder, prostate, and rectum contours, respectively.

Of the two atlas-based algorithms in H&N cases, OnQ rts revealed significantly lower MDC (i.e. better agreement with the gold-standard observer) in brain ($P < 0.001$), spinal cord ($P < 0.001$) and mandible ($P < 0.001$) with significantly higher MDC in brainstem ($P < 0.001$). For prostate cases, RayStation ABS exposed significantly higher MDC in bladder ($P < 0.001$) while significantly lower in prostate ($P < 0.01$).

Of the two fully automated model-based algorithms under evaluation, RayStation MBS and SPICE, the latter revealed significantly lower MDC in brain ($P < 0.05$), brainstem ($P < 0.001$) and parotids ($P < 0.001$) for H&N cases and rectum ($P < 0.001$) for prostate cases.

Comparing the semi-automatic algorithms, RayStation MBS_h and MultiPlan, with minimal time spent for manual intervention, RayStation MBS_h revealed significantly lower MDC in bladder ($P < 0.001$) with significantly higher MDC in prostate ($P < 0.001$) and rectum ($P < 0.001$).

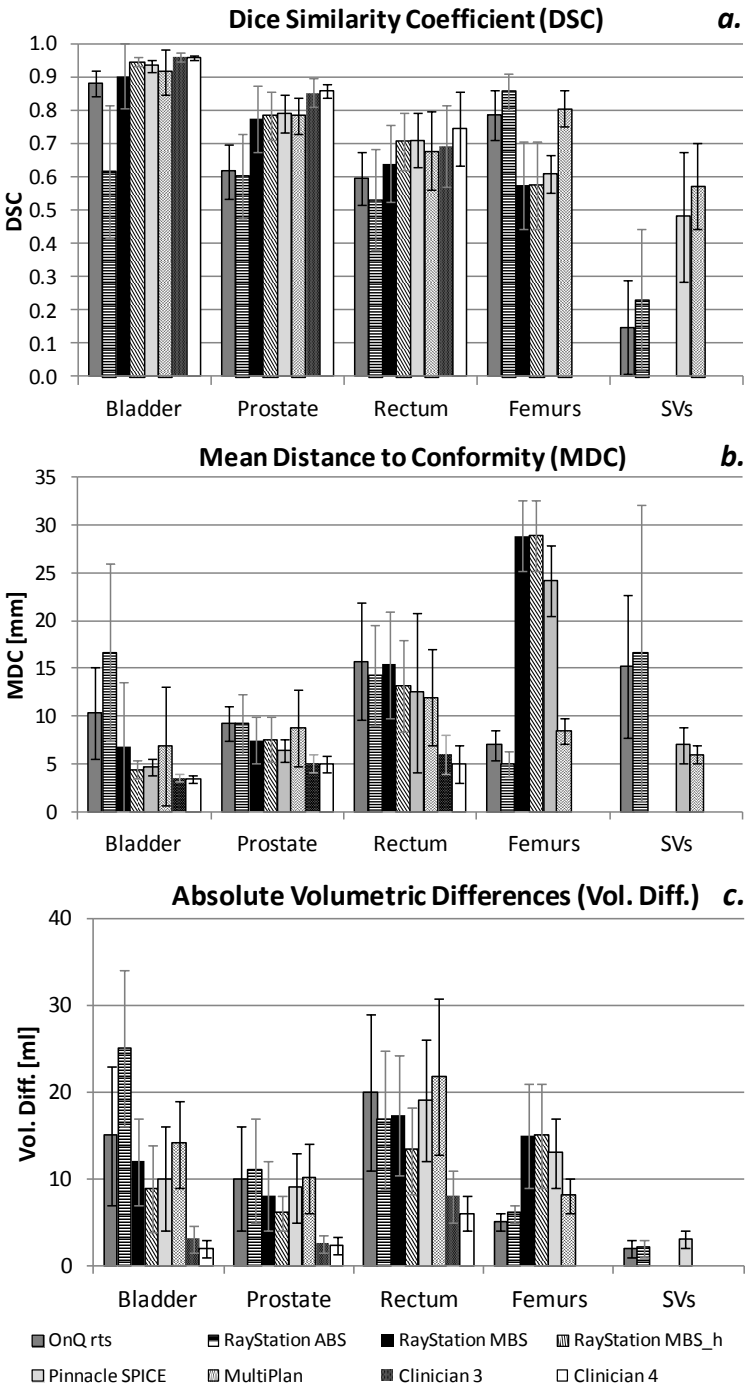

Figure 2.5: ***a.*** Mean and standard deviation of Dice similarity coefficient (DSC), ***b.*** mean distance to conformity (MDC), ***c.*** absolute volumetric differences for male pelvic region.

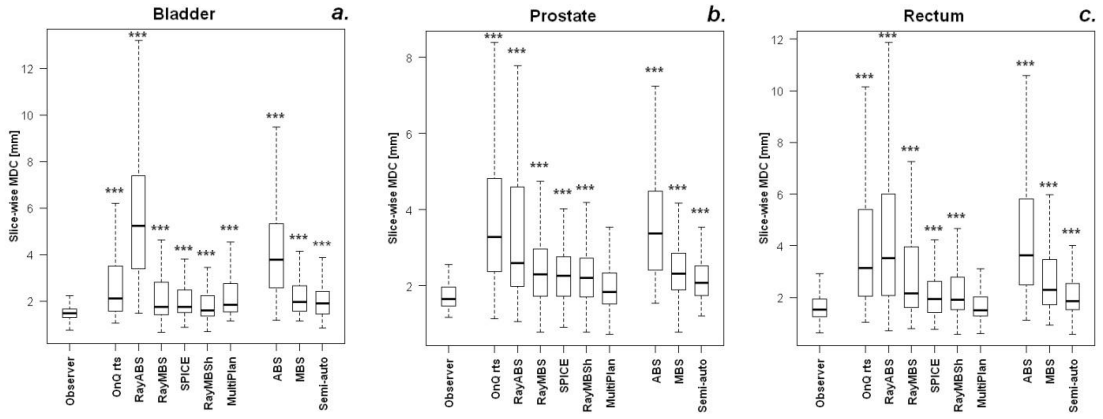

Figure 2.6: Mean distance to conformity (MDC): mean inter-observer variability (Observer) against individual algorithms and mean of atlas-based (ABS), model-based (MBS) and semi-automated MBS (Semi-auto) algorithms for ***a.*** bladder, ***b.*** prostate, and ***c.*** rectum. * denotes P-values following comparison against inter-observer variability (*:$P \leqslant 0.05$; **:$P \leqslant 0.01$; ***:$P \leqslant 0.001$)

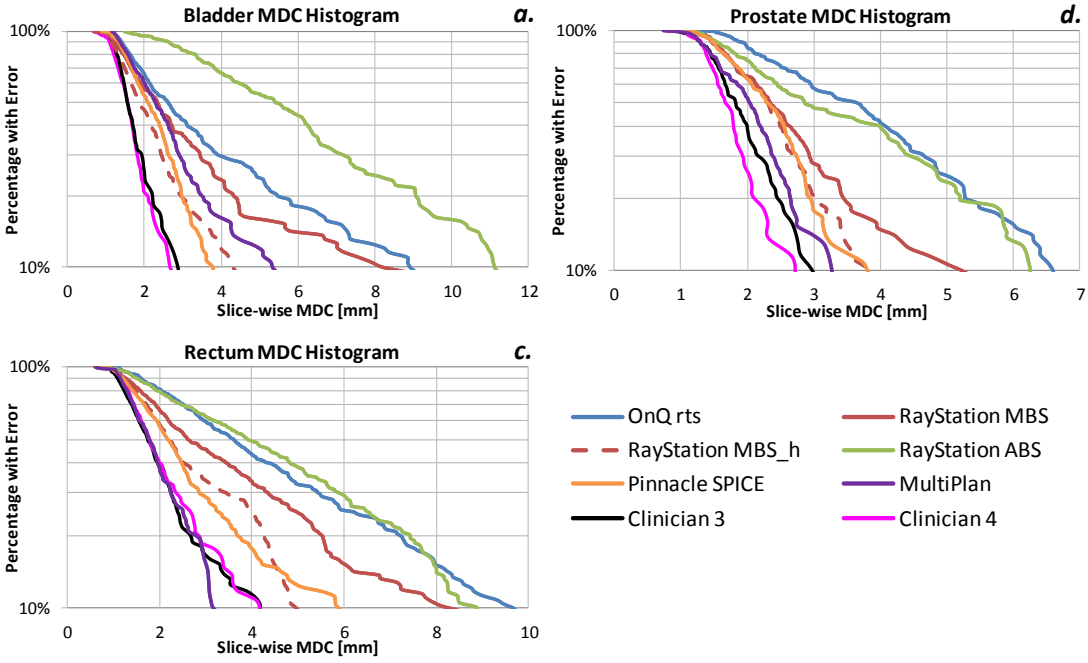

Figure 2.7: Slice-wise mean distance to conformity (MDC) histograms for ***a.*** bladder, ***b.*** prostate, and ***c.*** rectum, across all test cases.
(e.g. for prostate the histogram shows that 80% of the slices gave an MDC of $\leqslant 5.5$ mm for OnQ rts and RayStation ABS, with of $\leqslant 3.5$ mm for the other algorithms.)

2.1.3 Discussion

Performance of various commercial auto-segmentation algorithms has been assessed by spatial comparison of the resulting contours against manually drawn gold-standard contours produced by senior site-specialist consultants, employing the volume-wise DSC and MDC, slice-wise MDC metrics and absolute volumetric differences. The observed uncertainties were also compared against inter-observer variability, using observers with a range of expertise representative of that found in many clinics (i.e. suitably trained and performing manual contouring in the clinical setting). This approach is intended to represent the extent of inter-observer variability in busy radiotherapy departments.

Volume-wise DSC provides a measure of volume overlap but is prone to averaging effects when compared volumes are either too big or too small relative to the differences between them. Volume-wise MDC exposes the mean surface-to-surface distance between two volumes in three dimensions, but may overestimate the contour mismatch in cases where one of the volumes is either shorter or longer. This is particularly important in organs like the spinal cord and rectum, where, depending on the institution's internal guidelines, contour lengths may vary. Therefore, even though the slice-wise MDC also employed here ignores differences in superior and inferior directions, it provides detailed information for the mismatch in the transverse plain, which is arguably the most relevant in terms of reducing clinician delineation time as this process takes place primarily in the transverse plane.

A limitation of this study is the use of contours that have been manually drawn by a single observer as the 'gold-standard' contours. It could be argued that a commonly approved delineation by two or more consultants would be a more appropriate and bias-free method, however, the generally low inter-observer variability observed herein may justify this choice. Besides, evaluating several auto-segmentation software solutions in

terms of spatial accuracy, the aim of this study was the comparison of their uncertainty to inter-observer variability. Under the assumption that a finite inter-observer variability typically exists, it is herein considered the benchmark.

Furthermore, since the primary aim of auto-segmentation solutions is to save manual delineation time required by clinicians, it could be argued that this time saving should be quantified during the evaluation of such algorithms. To do this, clinicians would need to calculate the time required for manual correction of automatically generated contours. This investigation was not performed primarily due to limited availability of clinicians, but also due to the potential bias introduced due to the limited familiarity of clinicians in 'correction tools'. Most clinicians have years of experience in manual delineation tools offered by clinical software, while only occasionally use the correction tools to amend certain errors. Moreover, manual delineation is occasionally performed in alternate slices with automatic tools employed for interpolation of the discontinuous contours. On the other hand, the correction of automatically drawn contours would need to be performed in every image slice.

Due to the same reasoning, specific values of percentage of slices with a given MDC that would indicate clinical efficacy in performing manual corrections could not be objectively identified. It could even be argued that small errors of the order of 1-2 voxels (i.e. 1-3 mm) could be overlooked when evaluated by specialist clinicians, especially in areas of low contrast.

As mentioned previously, the time spent for the introduction of initialisation points in MultiPlan and hint contours in RayStation MBS_h was restricted, so further gains may be possible if additional time were taken, but would also hinder the potential time saving.

Adaptive radiotherapy workflows have been shown to be advantageous for H&N and prostate cases [12; 13], but this can involve re-planning on follow-up CT scans, thereby dra-

matically increasing the required time for manual organ delineations. Auto-segmentation algorithms capable of producing results with uncertainties comparable to inter-observer variability could therefore be used with minimal manual intervention, enabling wider implementation of treatment adaptations. This study has shown that some algorithms are close to this target for specific structures but that substantial further improvements are required before this becomes a clinical reality.

2.1.4 Conclusions

A range of auto-segmentation solutions for H&N and male pelvis anatomy have been evaluated with certain algorithms suggesting comparable uncertainties to inter-observer variability for specific anatomical structures. However, the generally inferior performance of these algorithms suggests that further improvements are required before their application can progress from reducing clinician delineation time to facilitating semi-automated treatment planning and adaptive radiotherapy workflows.

Chapter 3

Cone-beam CT Simulation

This chapter describes the evaluation of a tool that simulates CBCT images out of CT datasets, offered by a commercial product (ImSimQA, OSL, UK). The results of this study have been presented at UKRO 2013 conference.

An accurate and robust CBCT simulation tool would be valuable for the evaluation of certain procedures towards a semi-automated adaptive radiotherapy workflow, introducing into the evaluation the additional complexity posed by these images while preserving the 'ground truth'. Evaluation procedures such a tool may facilitate include: automatic segmentation of CBCT images, deformable registration between CT and CBCT datasets, dose calculations on CBCT scans or deformable dose accumulation from an on-line acquired CBCT to a reference CT.

3.1 Evaluation of cone-beam CT simulation from CT datasets

This study evaluated a tool, offered by ImSimQA software, that allows simulation of CBCT images out of CT datasets by introducing various levels of noise, filtering and cropping original images according to specified field-of-view. Specifically, the similarity of simulated CBCT images out of CT datasets as compared to actual CBCT scans was qualitatively and quantitatively assessed.

3.1.1 Methods

CBCT simulation in ImSimQA

The CBCT simulation module in ImSimQA includes two steps, the addition of noise and the definition of field-of-view (FOV). The CBCT noise generator has 3 components. First, a Gaussian noise pattern is added for which the user can control its parameters (mean value and standard deviation of the noise). Second, a ring noise pattern is used as a texture generating filter −to make the image intensity darker or brighter, for which the user can control its intensity. The third component combines and processes the combined pattern. A cylindrical FOV mask is then used to define the new CBCT FOV, the diameter of which can be controlled by the user.

Data collection

Computed tomography (CT) and CBCT scans of a Catphan® phantom (The Phantom Laboratory, Salem, NY) were acquired on a Philips Brilliance (Philips Healthcare, Netherlands) and an Elekta Synergy® equipped with an X-ray volumetric imaging devise (XVI) (Elekta AB, Sweden). The CT and CBCT scans were acquired with the standard in-house

protocol for head-and-neck (H&N) imaging and daily image guidance, respectively. For CBCT scanning the *S20* collimator, *F0* filter (i.e. no filter) and *Small* field-of-view panel position were employed, acquiring two identical repeat scans with the isocentre positioned in the centre of the *CTP404 module* section of the phantom and a third acquisition with the phantom offset by 50 mm superiorly in the scan plane. Furthermore, a CT scan and three daily on-board CBCT scans of a H&N and a prostate cancer patient were randomly selected.

The CT images of the Catphan® phantom and the patient scans were then imported into ImSimQA, where varying levels of 'CBCT' and Gaussian noise were added. During the noise addition process, visual comparison of the artificial CBCT images (aCBCT) and the original CBCT scans was being performed. Five aCBCT images with different combinations and levels of noise passed the visual assessment and were used for further evaluation.

Under the assumption that the random photon noise is the only source of noise during CT imaging, Brooks et al. [84] proposed a formula for the estimation of dose delivered to the centre of a cylindrical object (D) as a function of the signal-to-noise ratio (SNR), the slice thickness (b), and the spatial resolution (ε) of the resulting image:

$$D \propto \frac{(SNR)^2}{\varepsilon^3 b} \tag{3.1}$$

Quantitative analysis

To enable a quantitative evaluation of the CBCT simulation results a number of image quality measures employed in the literature were used [85; 86]:

Contrast-to-noise ratio (CNR):

$$CNR = \frac{\overline{CT}_{insert} - \overline{CT}_{background}}{\sigma_{background}} \tag{3.2}$$

Low contrast visibility (LCV):

$$LCV = 6.5 / \frac{\overline{CT}_{polystyrene} - \overline{CT}_{LDPE}}{(\sigma_{polystyrene} + \sigma_{LDPE})/2} \tag{3.3}$$

Uniformity index (UI):

$$UI = |\overline{CT}_{periferal\_ROI} - \overline{CT}_{central\_ROI}| \tag{3.4}$$

Signal-to-noise ratio (SNR):

$$SNR = \frac{\overline{CT}_{insert}}{\sigma_{insert}} \tag{3.5}$$

Noise percentage ($\sigma_{\%}$):

$$\sigma_{\%} = \frac{\sigma_{background}}{\overline{CT}_{background} - \overline{CT}_{air}} \cdot 100\% \tag{3.6}$$

where $\overline{CT}$ is the mean pixel value and σ the standard deviation of pixel values within a certain region of interest (ROI).

Relatively high levels of noise and non-uniformity are typically observed in CBCT images and therefore some of the above quality measures can be affected by the choice of ROI within the background. In order to expose the range of potential values due to this uncertainty, multiple calculations were made for some quality measures using different regions in the background.

Additional qualitative evaluation was performed by plotting the profiles of original (CT and CBCT) and artificial images through certain slices in the uniformity (*CTP486 Module*) and sensitometry (*CTP404 Module*) regions of the Catphan® phantom, shown in Figure 3.1 and regions in the patient datasets.

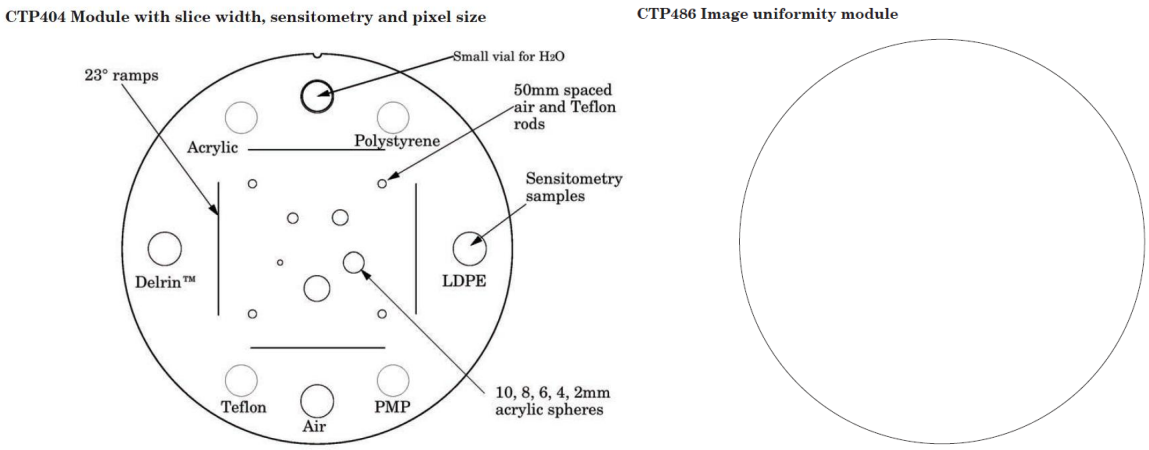

Figure 3.1: The two modules of the Catphan® phantom used during the analysis

The analysis was carried out with the use of the open source software ImageJ. All images were loaded to ImageJ and circular ROIs manually drawn in the middle of certain phantom inserts and the background. The drawn ROIs had smaller radius (80 mm) than the inserts (120 mm) to ensure they would be completely contained within the inserts. ImageJ macros were then used for the calculation of mean and standard deviation of pixel values within each ROI.

In the patient cases, rectangular ROIs were drawn spanning laterally in a central transverse plane slice (yellow regions seen in Fig. 3.6, p. 50, and Fig. 3.7, p. 51). The analysis in these cases was restricted to the qualitative evaluation using profile plots through the ROI regions. This was to avoid uncertainties due to potential misalignment of small ROIs, the positioning of which in different images would not be reproducible.

3.1.2 Results

Catphan® phantom

Figure 3.2 shows the CNR values calculated for polystyrene and low-density polyethylene (LDPE) inserts against the background. Figure 3.3 shows the LCV, UI, SNR and percentage noise values for the measured and artificial CBCTs. These results reveal that variations in objective quantitative image quality metrics in the artificial CBCT images were within the range established by the measured CBCT scans.

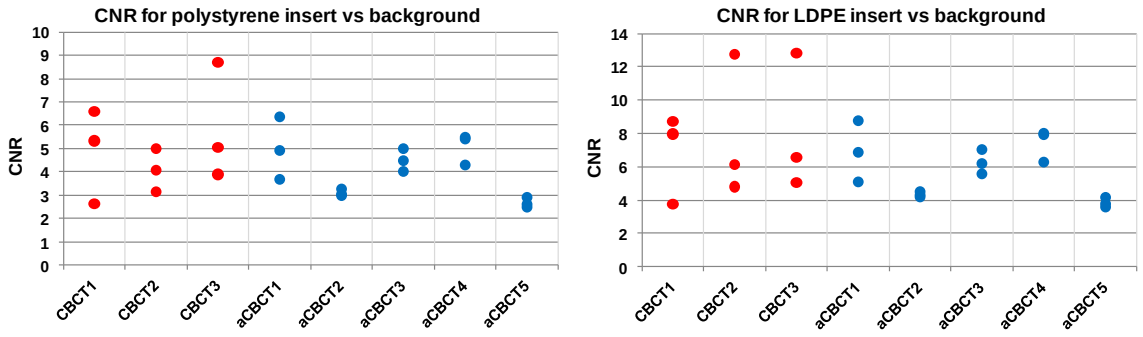

Figure 3.2: Contrast-to-noise ratio (CNR) values for the polystyrene and low-density polyethylene (LDPE) inserts as calculated for the 3 measured CBCT scans and 5 artificial CBCTs (aCBCT). Three CNR values were calculated for each image using 3 regions in the background.

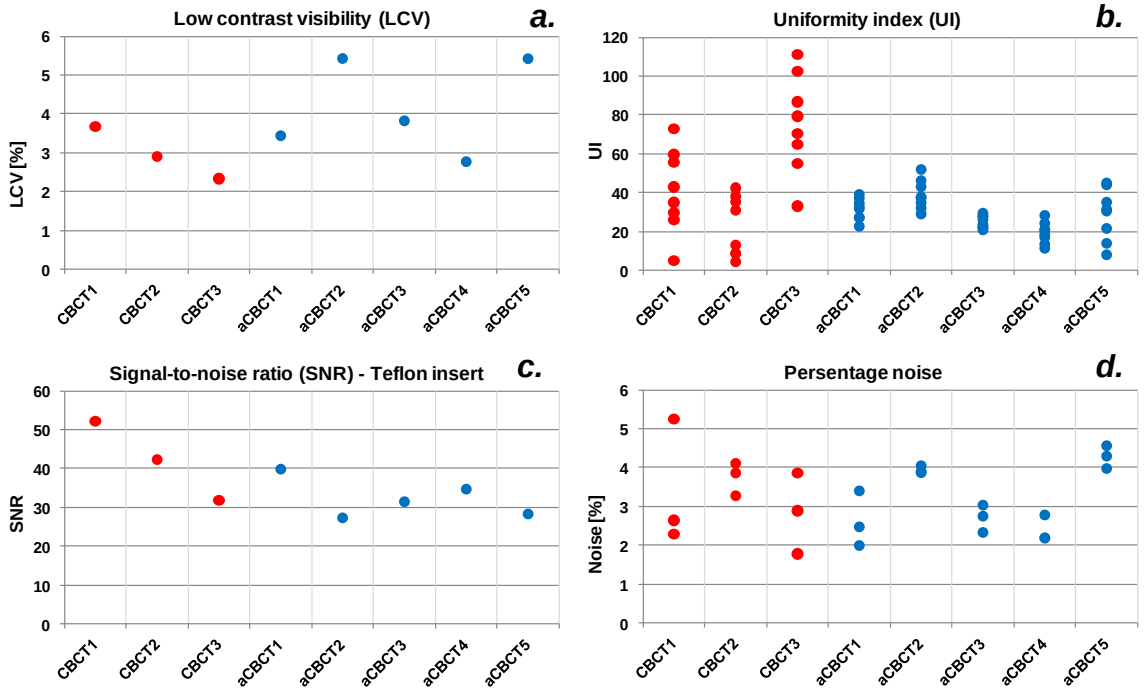

Figure 3.3: ***a.*** Low contrast visibility (LCV), ***b.*** uniformity index (UI), ***c.*** signal-to-noise ratio for the teflon insert, and ***d.*** percentage noise values as calculated for the 3 measured CBCT scans and 5 artificial CBCTs (aCBCT). Multiple values were calculated for UI and percentage noise in each image using different regions in the background.

Figure 3.4 shows profile plots through the sensitometry region of the Catphan® phantom for the original CT, the 3 measured CBCTs and one of the artificial CBCTs. These plots demonstrate representative CBCT simulations, particularly in the central region. The inclusion of the profile plots for the original CT scan highlights the changes introduced by CBCT noise simulation. Figure 3.5 shows profile plots through the uniformity region, exposing a slight systematic deviation from measurement at the periphery of the phantom.

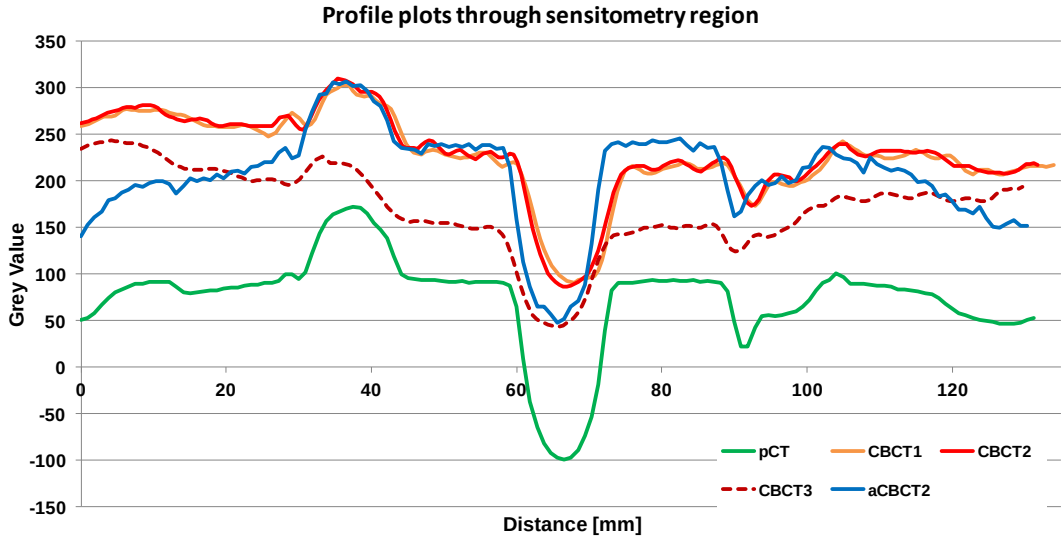

Figure 3.4: Profile plots through the sensitometry region (*CTP404 Module*) of the Catphan® phantom as calculated for the original CT (pCT), the 3 measured CBCTs and one of the artificial CBCTs (aCBCT) images.

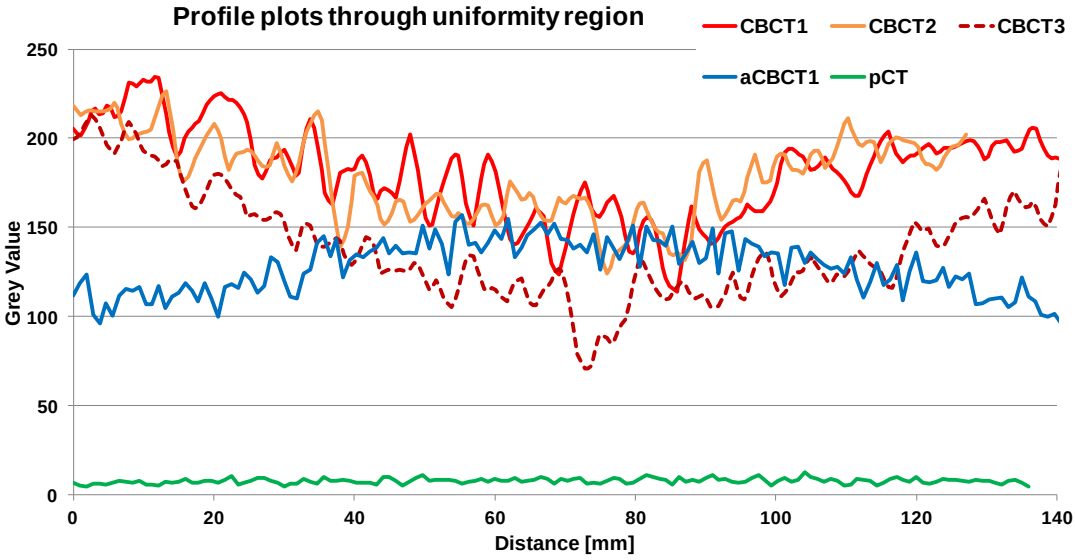

Figure 3.5: Profile plots through the uniformity region (*CTP486 Module*) of the Catphan® phantom as calculated for the original CT (pCT), the 3 measured CBCTs and one of the artificial CBCTs (aCBCT) images.

Patient cases

Figure 3.6 shows a transverse slice from a H&N patient's original CT and the respective profile plots through the marked region for the original CT, 3 measured CBCTs and one of the artificial CBCTs. Good agreement is observed between the measured and artificial images, while the changes introduced to the original CT are highlighted (i.e. difference between green line to the rest of the profiles).

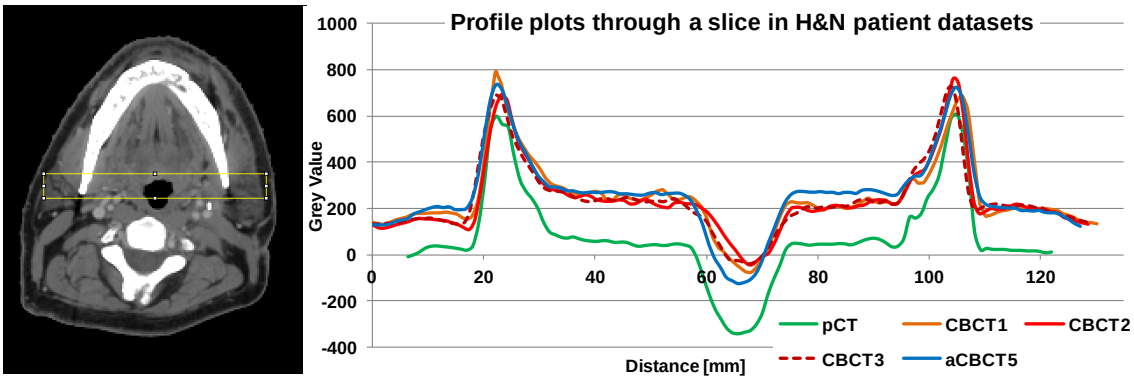

Figure 3.6: Profile plots through the one slice of the H&N cancer patient datasets -yellow region shown on the left hand side- as calculated for the original CT (pCT), the 3 measured CBCTs and one of the artificial CBCTs (aCBCT) images.

Similarly, Fig. 3.7 shows the corresponding profile plots for the prostate cancer patient case, revealing a slightly less good agreement between the measured and artificial images. This can be attributed to the small deviation in uniformity observed previously, noise and artefacts. The larger anatomy in the prostate cancer patient case, compared to the small H&N anatomy make the accurate uniformity modelling more important. Accordingly, additional noise and artefacts are also observed when the anatomy is larger.

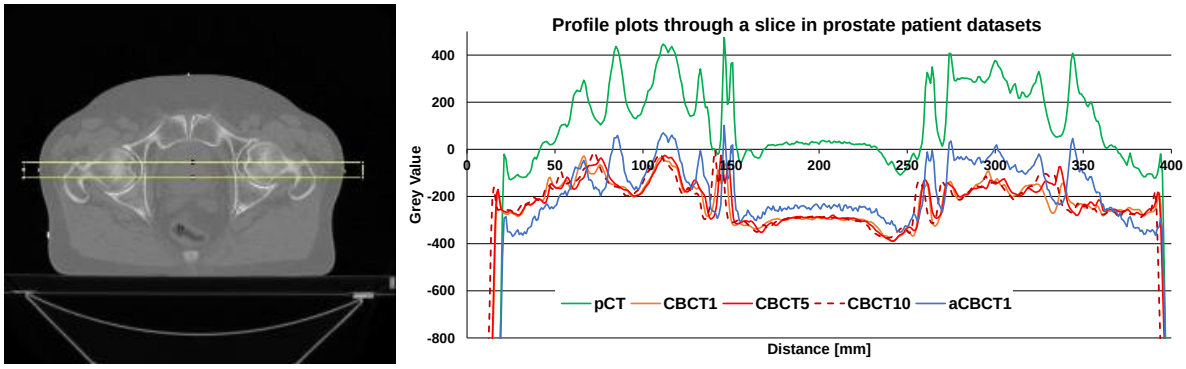

Figure 3.7: Profile plots through the one slice of the prostate cancer patient datasets -yellow region shown on the left hand side- as calculated for the original CT (pCT), the 3 measured CBCTs and one of the artificial CBCTs (aCBCT) images.

3.1.3 Discussion

This study evaluated a tool offered by a commercial software, which facilitates the simulation of CBCT images out of CT datasets, in phantom and patient cases using qualitative and quantitative measures. In the phantom examination, quantitative image quality metrics between measured and simulated CBCT images revealed good agreement, while further qualitative evaluation revealed a slight systematic deviation in the simulation of non-uniformity observed in CBCT scans. In the patient cases, qualitative evaluation between the measured and artificial CBCTs showed good agreement in the H&N case with less good agreement in the male pelvis case.

ImSimQA's CBCT noise simulation tool could therefore by used to facilitate the evaluation of automatic segmentation of CBCT images, deformable registration between CT and CBCT datasets (Section 4.2, p.64), or even deformable dose accumulation from a CBCT to a reference CT, under realistic conditions. Robust evaluation of such procedures would greatly benefit from the preservation of 'ground truth' the simulation of CBCT images from CT datasets provides. Evaluation of dose calculation on artificial CBCT datasets would require a more accurate noise simulation algorithm.

In the patient cases investigated it was observed that the relationship between the HU in the planning CT and CBCT scans for prostate patient case was opposite to that in the H&N patient case. This may be attributed to the different filtering employed in each scanning protocol. In the H&N cases our department's protocol utilises the *Small* field-of-view and no beam filtering. On the other hand, the protocol for pelvis scanning utilises the *Medium* field-of-view and a bowtie filter. The absence of bowtie filter in the H&N scanning protocol may result in overexposure and potentially saturation of areas of the detector panel [87; 88].

CBCT scans have a larger degree of scatter as compared to diagnostic fan-beam CT, primarily due to the larger FOV [65; 89]. Furthermore, large scanned objects will result in additional scatter and hence greater noise and differences in Hounsfield units [90], while, beam hardening artefacts are also greater for bigger objects [91]. Bow-tie filters, anti-scatter grids and correction algorithms can minimize these effects to a certain extent. A relatively simple noise simulation algorithm (i.e. with only few degrees of freedom) such as the one evaluated in this study, can not adequately simulate all these processes.

The evaluation of worfklows for direct dose calculation or CBCT scans may require more sophisticated CBCT simulation algorithms. Due to the complexity and range of artefacts and noise observed in typical CBCT images, a very accurate simulation would require ray-tracing or even Monte Carlo approaches. In fact, Jia et al. recently developed an algorithms which combines both of these techniques for the accurate simulation of CBCT projections out of CT datasets while also incorporating geometric and mechanical factors from individual commercial CBCT systems [92; 93].

3.1.4 Conclusions

A commercial tool that enables the simulation of CBCT images using CT datasets was evaluated in phantom, head-and-neck and male pelvis cases. Representative images were generated in all cases. Inaccuracies were observed in the simulation of non-uniformity in the phantom experiment and the simulation of noise and artefact levels in the presence of large anatomy (i.e. male pelvis).

Chapter 4

Deformable Image Registration

This chapter outlines the evaluation of deformable image registration (DIR) algorithms for use in adaptive radiotherapy workflows. The first section describes the evaluation of several commercial DIR solutions in CT-to-CT registration. The second section describes the evaluation of these DIR solutions in CT-to-CBCT registration.

DIR algorithms have a key role in certain adaptive radiotherapy workflows, as, by determining voxel-to-voxel correspondence between two images they can facilitate automatic segmentation and dose accumulation.

4.1 Evaluation of commercial solutions for CT-to-CT DIR

This study aimed to quantitatively evaluate the accuracy of several commercial DIR algorithms, using artificial digital phantoms, in CT-to-CT registration for head-and-neck (H&N) and male pelvis anatomy.

4.1.1 Methods

Evaluated algorithms

The algorithms under evaluation were:

(a) OnQ rts (OSL, Shrewsbury, UK)

(b) RayStation (RaySearch, Stockholm, Sweden) Hybrid DIR without using controlling regions-of-interest (ROI)

(c) RayStation Hybrid DIR using controlling ROIs

The DIR algorithm in OnQ rts employs an image intensity-based non-linear optimisation based on Demon's algorithm [24]. A multi-resolution iterative procedure is performed whereby a similarity measure is evaluated in each iteration. The similarity of the images to be registered is initially evaluated at a coarse image resolution. The boundaries of objects within the image are seen as membranes through the 'moving' image 'diffuses' to maximise the similarity. The procedure is then repeated for multiple iteration (the number of which can be controlled by the user) for finer image resolutions. The estimated deformation forces are regularised by simple Gaussian smoothing.

The hybrid DIR algorithm in RayStation is based on a mathematical formulation linearly composed by four non-linear terms: (1) an image similarity term; (2) a term to ensure smoothness and invertability of the deformed image grid; (3) a term to maintain anatomically reasonable deformation computation when ROIs are used; and (4) a penalty term −employed when controlling ROIs are used− ensuring structures in the two images will be matched. The first two terms operate under similar principles as the algorithm in OnQ rts. The third and fourth terms introduce a feature-based component to the algorithm. In addition to the intensity-based similarity metric calculated in each iteration to

drive the registration, the addition of ROIs allows the algorithm to be also driven by the spatial correspondence of respective ROIs.

Digital phantom creation

Each patient's planning CT scan (pCT) was transferred to ImSimQA (OSL, UK, v.3.0.77) software where clinically realistic artificial deformations have been introduced, following a procedure previously described by Varadhan et al. [46]. Briefly, ImSimQA offers 'global' and 'local' deformation simulation options. The user defines control points on the CT images that can then be displaced individually or in clusters in order to simulate anatomical movements.

All artificial deformations were based on uncommon but actual clinical observations during on-line volumetric image guidance at Queen Elizabeth Hospital, Birmingham, UK. Simulations were confirmed as clinically plausible and realistic following visual inspection by a site-specialist consultant and a site-specialist radiographer.

H&N cases: Twelve H&N cancer patient CT scans with the associated structure sets (RTS) were randomly selected. Using the 'global deformation' option and by moving the chin and back of the head by 10-15 mm in opposite directions, backward (DefHN1) and forward (DefHN2) neck flexion scenarios were simulated for each patient CT scan.

Male pelvis cases: Twelve prostate cancer patient CT scans with the associated structure sets (RTS) were randomly selected, and three artificial deformations have been applied to each of these scans. The first deformation (DefPr1) involved the identification of a rectal 'gas' pocket in the planning CT scans and its 'local' radial expansion by 5-10 mm, with a subsequent expansion of the rectum. Bladder expansion (DefPr2) was simulated by 'locally' expanding the bladder upwards by 10-15 mm and by 2-5 mm towards the prostate and rectum. For the third artificial deformation (DefPr3) an artificial rectal gas pocket was introduced and expanded with a subsequent radial deformation of the rectum.

For the introduction of the rectal gas pocket, average Hounsfield Units were measured from gas pockets observed in the CT scans of other patients and this value was used to 'paint' a small region in the rectum. DefPr3 can be considered as a closer approximation of the potential clinical scenario, but is also more challenging for intensity matching DIR algorithms, such as the Demon's, due to the introduction of new features.

Deformable image registration evaluation

The original and artificially deformed CT images, along with the associated DICOM structure sets (RTS) were then transferred to OnQ rts and RayStation systems. DIR was performed with the original CT scans set as the target (or stationary) images and the artificial scans set as the source (or moving) image that is deformed to match the target image. For RayStation's hybrid DIR implementation using controlling ROIs, the body and spinal cord were used in H&N cases, while in male pelvis cases the bladder, prostate and rectum were employed.

The RTS associated with the source images were automatically deformed with the resulting deformation matrix and transferred to the target images. DIR performance was therefore evaluated by comparing the original and DIR-mapped RTS sets, employing the Dice similarity coefficient (DSC) (Equation 2.1, p. 31) and mean distance to conformity (MDC) (definition in p. 31) metrics.

Furthermore, the reproducibility of the registrations was also evaluated. Each DIR, described above, was performed twice and two sets of DIR-mapped contours compared using the DSC and MDC metrics.

4.1.2 Results

H&N cases

Figure 4.1 shows the slice-wise MDC histograms calculated for brain, brainstem, parotids, spinal cord and mandible, following DIR and contour mapping of both sets of artificial deformations to the original anatomy.

Considering the logarithmic scale of the graphs, it is observed that OnQ rts and RayStation's 'Hybrid' implementation without the use of controlling ROIs (RayStation H_noROIs) revealed similar performance. On the other hand, the use of controlling ROIs improved the registration result for all investigated organs.

It is also observed that regions like the brain and mandible, which are usually defined by bony boundaries and hence one would expect better registration performance, revealed a larger range of errors. This can be attributed to the fact that these regions were 'moved' more as part of the artificial deformations applied.

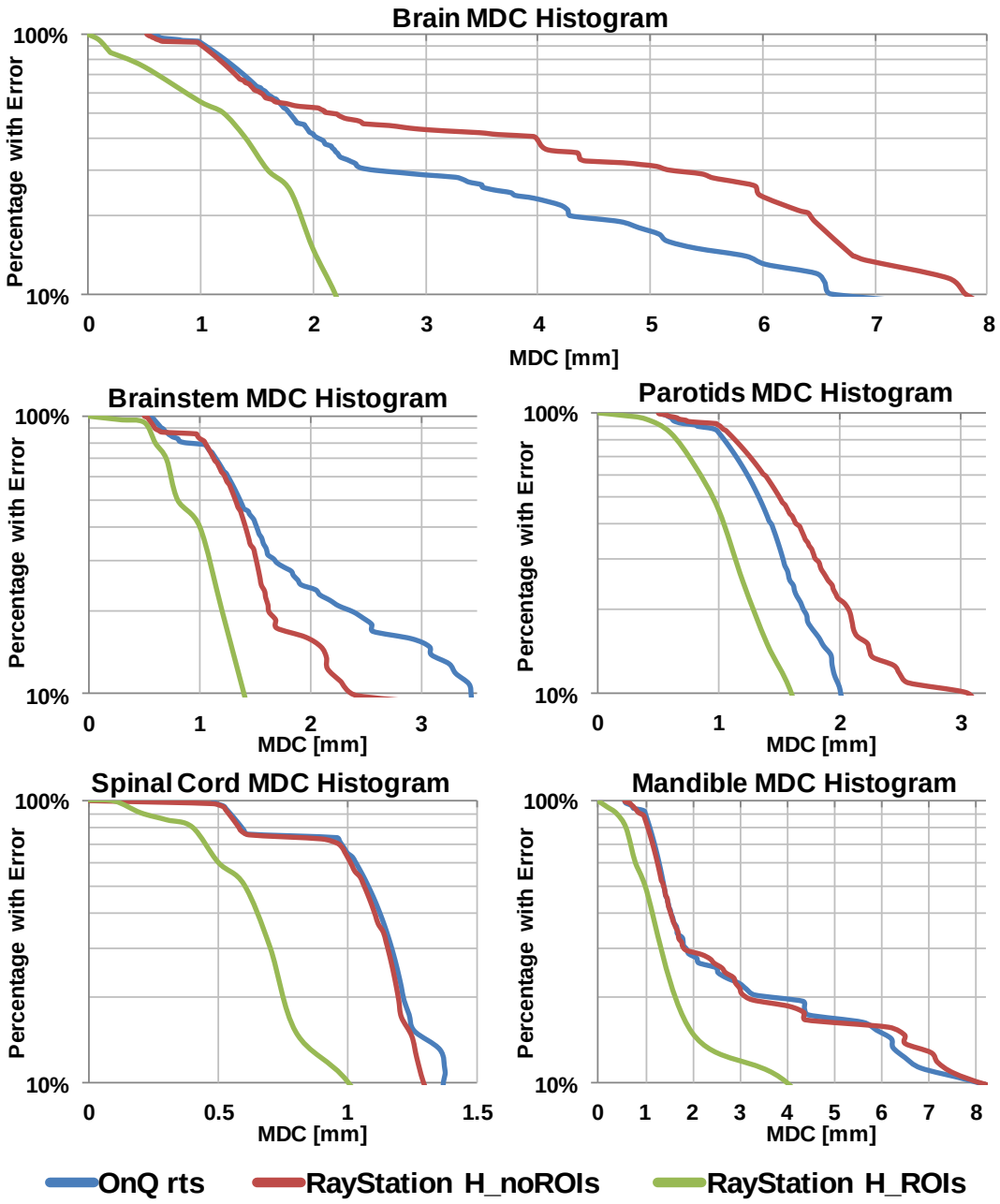

Figure 4.1: Mean distance to conformity (MDC) histograms for brain, brainstem, parotids, spinal cord and mandible, following deformable image registration (DIR) of both sets of artificial images.

Male pelvis cases

Figure 4.2 shows the slice-wise MDC histograms calculated for bladder, prostate and rectum, following DIR and contour mapping of all three sets of artificial deformations.

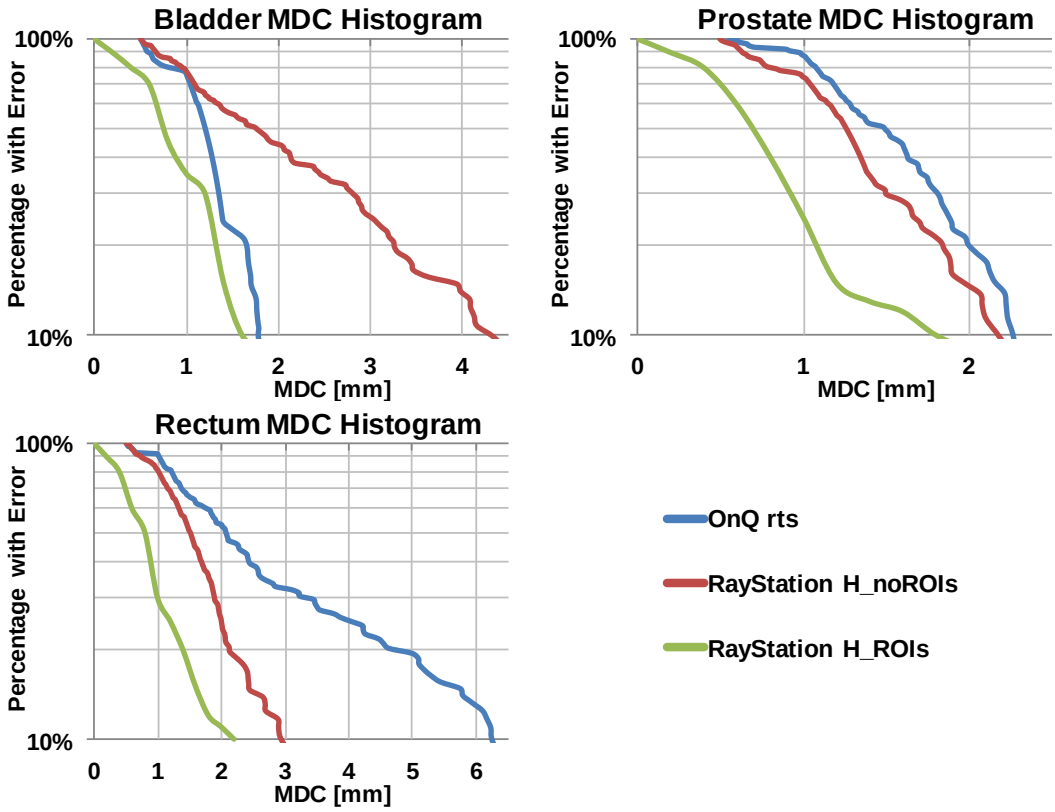

Figure 4.2: Mean distance to conformity (MDC) histograms for bladder, prostate and rectum, following deformable image registration (DIR) of all three sets of artificial images.

Figure 4.3 shows the average volume-wise MDC values calculated following DIR of each artificial deformation to the original anatomy, individually. It is observed that the use of controlling ROIs with the 'Hybrid' implementation improved the registration result for all three organs, even after the introduction of new features in the artificial image (i.e. DefPr3, Fig. 4.3c).

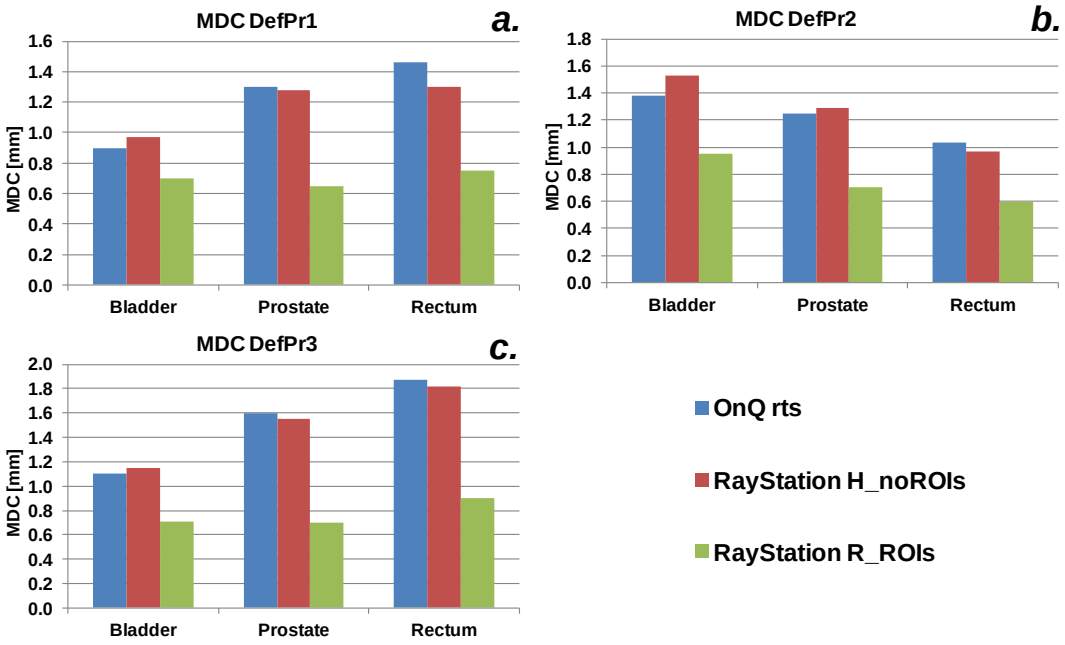

Figure 4.3: Average mean distance to conformity (MDC) charts for bladder, prostate and rectum following deformable image registration (DIR) between original CT and each set of artificial deformations.

DefPr1: Existing rectal pocket expansion.

DefPr2: Bladder expansion.

DefPr3: Introduction of artificial rectal pocket and rectal expansion.

Reproducibility

All three investigated algorithms were proven reproducible, with perfect overlap of the DIR-mapped contours (i.e. DSC = 1, MDC = 0 in all cases).

4.1.3 Discussion

This study evaluated three DIR algorithms, for CT-to-CT registration of H&N and male pelvis anatomy, using clinically realistic artificial deformations. The use of the artificial deformations provided the ground truth to which the results have been compared in terms of spatial accuracy.

The results showed that all algorithms performed reasonably well in the investigated scenarios, with the majority of slice-wise errors being less than 2 mm. However, larger errors ($2 - 8$ mm) were observed. A closer look at the areas with large errors confirmed that these were primarily found in areas with low contrast (e.g. brain-brainstem and prostate-rectum boundaries) or in areas were large deformations were applied during the artificial deformations (e.g. mandible and top of bladder). In general, the inclusion of controlling ROIs in the Hybrid implementation of RayStation DIR greatly improved the registration result and allowed the algorithm to cope with the addition of new features in the artificial image.

Of note is the relatively poor performance of image intensity-based algorithms when new features appear (or disappear) on the deformed image (i.e. OnQ rts and RayStation H_noROIs). This type of algorithms generally attempt to determine a voxel-to-voxel correspondence between the two images to be registered, with the typical underlying assumption that mass in conserved. For this reason, contour-based or hybrid DIR algorithms are expected to perform better in these scenarios, provided that accurate contours exist in both image sets to be registered.

The evaluation method herein was based on the evaluation of the original contours (RTS) against the mapped contours, following DIR between the original and the artificial CT images. This means that any potential registration errors within each organ would be ignored. This is arguably of less importance when DIR is used to map contour sets from

one instance to another but can be of great importance when DIR is used for, say, dose mapping. Such spatial evaluation when using artificially created deformations, could be performed by comparing the inverse of the applied three-dimensional deformation matrix (also called dynamic vector field, DVF) against the DIR-calculated DVF. This procedure was not performed in this study as not all matrices could be exported from the evaluated systems. Furthermore, no DICOM file format exists, yet, for the standardisation of the DVF files. The development of such format and potentially software algorithms to enable direct comparison of such data becomes more important with the increasing use of DIR procedures.

4.1.4 Conclusions

This study evaluated the performance of three DIR algorithms, in CT-to-CT registration of H&N and male pelvis anatomy, using clinically realistic artificial deformations. All three algorithms revealed good performance under the evaluation conditions, with the majority of errors being less than 2 mm. In areas of low contrast or when new features were added to the artificially deformed image, RayStation's hybrid algorithm with the use of controlling ROIs retained its good performance while the other two methods exposed larger errors (2-8 mm).

4.2 Evaluation of commercial solutions for CT-to-CBCT DIR

This study aimed to quantitatively evaluate the three DIR algorithms employed in Section 4.1 in CT-to-CBCT registration of male pelvis cases. Only this anatomical site was chosen as manual contouring on CBCT images was not practical (time-wise, experienced radiographer or specialist physician availability, and limited tissue contrast), while auto-contouring on CBCT images was not considered sufficient in neither H&N nor male pelvis anatomies.

4.2.1 Methods

Ten prostate cancer patients previously treated at our institution with daily image guided IMRT were randomly selected, picking the planning CT with the associated structure set (pCT, pRTS) and one online acquired CBCT image. Contours were manually drawn on the CBCT for the critical organs (cbctRTS), namely the prostate gland, bladder and rectum, by an experienced radiographer. These contours were subsequently considered the 'ground truth' structure sets.

The planning CT images and structure sets of the 12 prostate patients used in Section 4.1, together with the 3 artificially deformed images created using their planning CT scans (as described in Section 4.1.1, p. 56), were also employed. Noise was then added to the artificially deformed images, using ImSimQA software in order to convert them into an approximation of CBCT images, following the workflow described in Chapter 3, p. 42. Specifically, various levels of Gaussian noise were added changing the mean and standard deviation noise intensity values, followed by varying ring noise patterns to introduce texture and simulate the non-uniformity of CBCT scans. Finally, the field-of-

view (FOV) of the image was altered to match that of the actual CBCT FOV.

The registration and evaluation procedures, described in Section 4.1, were then performed.

4.2.2 Results

CT-to-CBCT DIR

Figure 4.4 shows the slice-wise MDC histograms calculated for bladder, prostate and rectum, following DIR and contour mapping between the CT and measured CBCT images. OnQ rts and RayStation H_noROIs reveal similar performance, and inferior to RayStation H_ROIs performance.

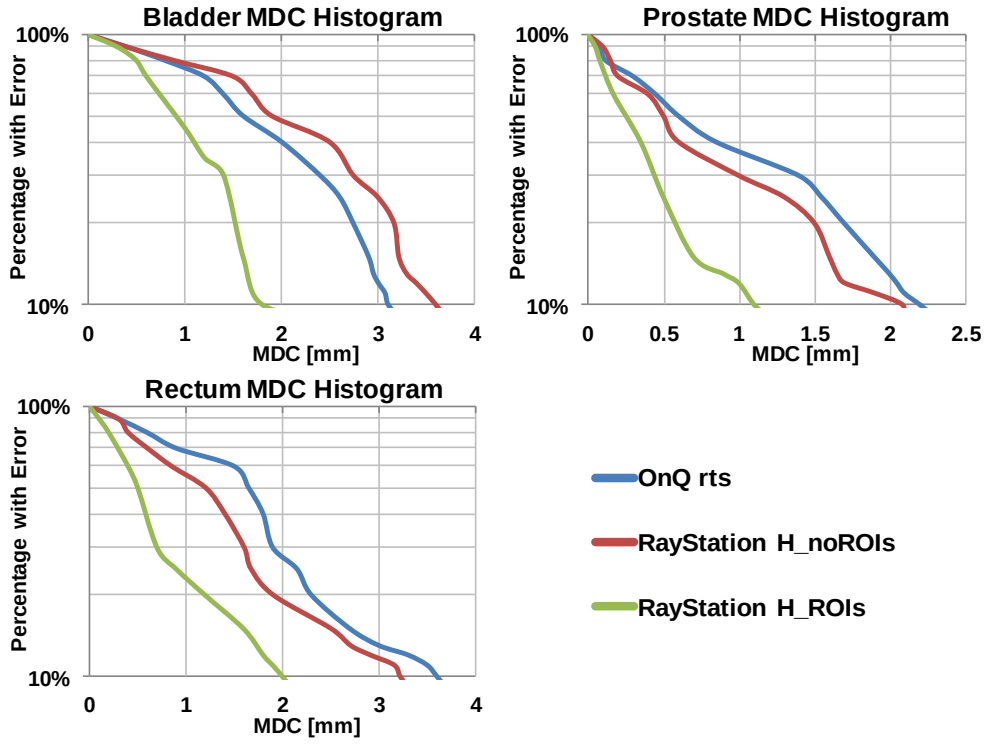

Figure 4.4: Mean distance to conformity (MDC) histograms for bladder, prostate and rectum, following deformable image registration (DIR) of all CBCT images.

CT to artificial CBCT DIR

Figure 4.5 shows the slice-wise MDC histograms calculated for bladder, prostate and rectum, following DIR and contour mapping between the CT and artificially deformed images with added CBCT noise simulation. It should be emphasised that the artificially deformed images used for this experiment are the same as those employed in Section 4.1.1 (p. 56). In other words, the results in Fig. 4.5 correspond to the ones in Fig. 4.2 (p. 60) but with the addition of CBCT simulated noise. Comparing the two Figures it is observed that both image intensity-based-only algorithms (i.e. OnQ rts and RayStation H_noROIs) revealed inferior performance in the presence of CBCT simulated noise. On the other hand, the use of controlling ROIs in RayStation's hybrid DIR implementation revealed similar results.

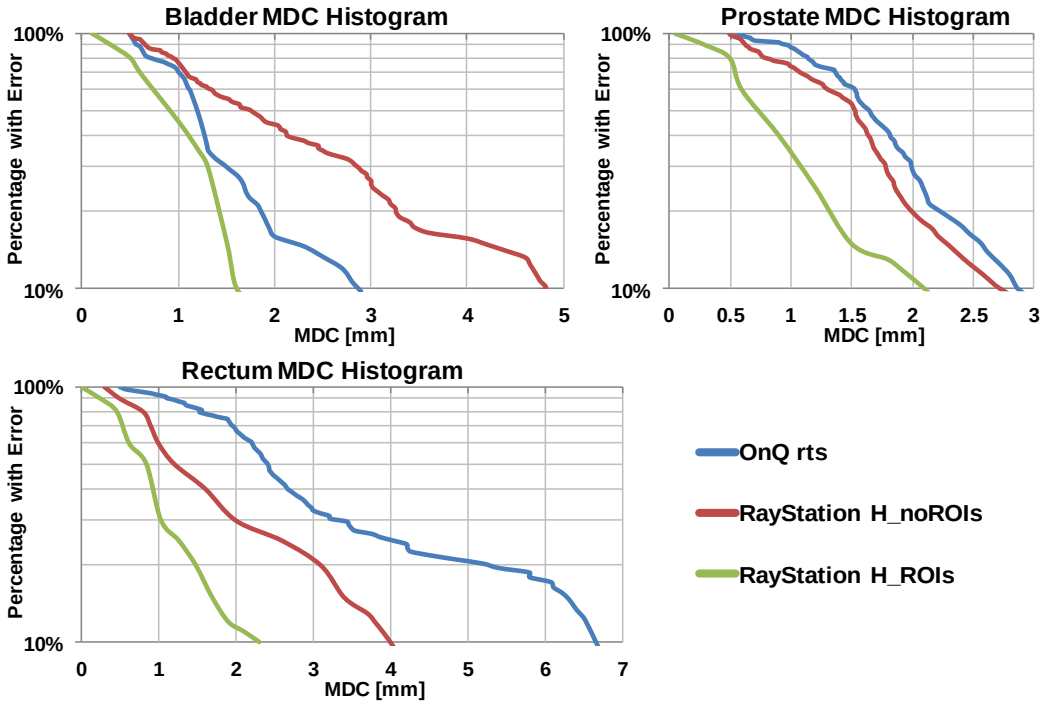

Figure 4.5: Mean distance to conformity (MDC) histograms for bladder, prostate and rectum, following deformable image registration (DIR) of all three sets of artificially deformed images with CBCT noise simulation.

4.2.3 Discussion

This study evaluated three DIR algorithms, for CT-to-CBCT registration of male pelvis anatomy, using both clinical datasets and artificial datasets.

The use of clinical CT and CBCT data enabled the evaluation of the DIR algorithms under real conditions. However, an absolute unconditional evaluation is hindered by the uncertainties involved during the manual delineating of ground truth contours on low image quality CBCT scans. Delineation uncertainties are diminished by the use of artificially deformed CT datasets with CBCT noise simulation, which nonetheless are merely a simplistic simulation and not actual CBCT scans.

Comparing the performance of the three DIR algorithms between the same images with and without the addition of CBCT noise simulation it is observed that the performance of RayStation H_ROIs was barely affected, while the performance of OnQ rts and RayStation H_noROIs was inferior after the addition of noise. Regrettably, however, this hybrid DIR methodology may not be practical in busy radiotherapy departments as it involves an additional ROI segmentation step, which could either be manual or automated but with the requirement of visual post-validation.

4.2.4 Conclusions

Three DIR algorithms have been evaluated in CT-to-CBCT registration of prostate patient datasets. All three algorithms (OnQ rts, RayStation's Hybrid algorithm with and without controlling ROIs) revealed low errors with the majority being up to 2-3 mm. The use of controlling ROIs with RayStation's Hybrid algorithm revealed superior performance while being consistent in areas of low contrast or even compared to the registration of the same artificial deformations prior to the addition of CBCT simulated noise.

Chapter 5

Dosimetric Treatment Monitoring

This chapter investigates workflows for the estimation of actual delivered dose during each radiotherapy fraction (i.e. the 'dose of the day'). The fist section describes a study aiming to exemplify a certain inconsistency observed in the literature on the use of dose deformation for delivered dose estimation, the preliminary results of which have been presented at UKRO 2013 and ESTRO 33 [94] conferences. The second section evaluates four techniques to allow direct dose calculation on the online acquired cone-beam CT scans or the representation of the anatomy in these images.

Dosimetric treatment monitoring is a critical part of certain adaptive radiotherapy workflows. The dosimetric impact of observed anatomical changes can be inspected by calculating the treatment plan on the new anatomy, while, the cumulative impact of gradual and/or random anatomical changes can be examined by accumulating 'daily delivered' dose distributions on a reference anatomy.

5.1 Delivered dose estimation: rights and wrongs

This study investigates and evaluates two strategies to allow calculation of 'dose of the day' during fractionated radiotherapy. As discussed in Section 1.4.3 (p. 17) two methods have been employed in the literature: one method is to recalculate the dose on a scan representing the daily anatomy; while a second method (which is later shown to be misguided) is to directly warp the original dose distribution to the new anatomy following DIR, without dose recalculation [55–58]. It should be emphasised that this work will not investigate the alternative use of dose warping, for the accumulation of recalculated dose distributions on a single frame of reference.

5.1.1 Methods

Creation of digital phantoms

Twelve head-an-neck (H&N) cancer patients, previously treated using intensity modulated radiation therapy (IMRT) with a prescribed dose of 65 Gy in 30 fractions, were retrospectively selected for this study. Clinically realistic artificial deformations simulating forward and backward neck flexion and shoulder movement, have been applied to each patient's planning CT scan using a commercial simulation software (ImSimQA, OSL, UK) as previously described in Section 4.1.1 (p. 56).

Backward and forward neck flexion: For the simulation of backward and forward neck flexion, the mandible-nose region and back of head were moved in opposite directions by 10-15 mm using 'global deformation', as shown in Fig. 5.1*a*. This global movement resulted in smaller displacement of the inner structures like the spinal cord, larynx, brain and brainstem, in the order of 3-6 mm.

Upward shoulder movement: Upward shoulder movement was simulated by 'lo-

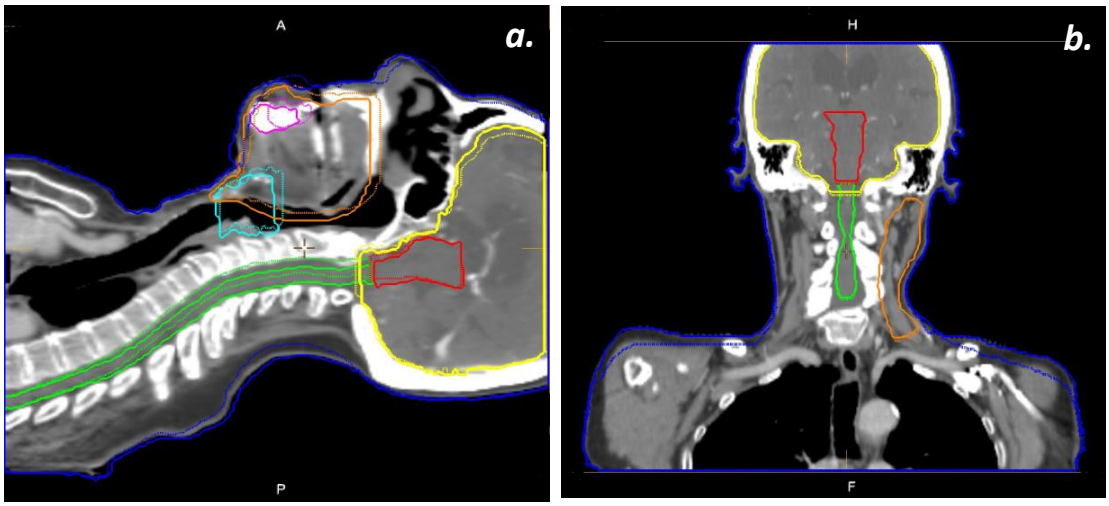

Figure 5.1: Examples of artificial deformations applied to patient datasets. The CT images show the artificial deformation results. Solid lines show the deformed and dashed lines the original structure sets for ***a.*** forward neck flexion, and ***b.*** upward shoulder movement simulation.

cally' moving the shoulders by 10-15 mm, as shown in Fig. 5.1*b*. The local deformation did not have any effect on inner structures, but displaced the region around the planning target volume (PTV) by 5-10 mm in cases where this was very close to the shoulders.

The aim for all three artificial deformations was to represent clinically observable anatomical changes while ensuring volume conservation, since part of the evaluation is based on dose-volume histogram (DVH) analysis. All simulations were based on actual clinical observations and verified through visual inspection by a specialist clinician and a specialist radiographer. Volume conservation was quantitatively assessed by comparison of original and deformed structure volumes and found to be conserved to $\geqslant 99\%$. Even though internal anatomy volumetric changes are often observed in clinical setting, such simulations were not employed in this study as this would make evaluation difficult, while not being essential for the evaluation of the strategies to be investigated.

Deformable image registration and delivered dose estimation

The workflow used here for the evaluation of strategies for actual delivered dose estimation under anatomic deformation is summarised in Fig. 5.2. The three clinically realistic artificial deformations described above have been applied to the planning CT scan (CT1) of each of the twelve H&N patient datasets employed for this study, creating 48 patient-specific phantoms (CT2), each of which can be assumed to represent a patient's anatomy during a hypothetical treatment fraction. Each CT2 was then imported into Monaco treatment planning system (Elekta AB, Sweden, v3.20) where the original treatment plan was applied (in order to ensure identical conditions to the original dose calculations, the original plan was applied as a 'QA plan'). The dose calculated on CT2 was considered the 'true' distribution (***Ground_Truth***) since this would have been the dose received by the new anatomy after the delivery of the plan on a hypothetical treatment fraction, assuming perfect conditions.

Each CT2 and the CT1 with the associated structure set (RTS1) and dose distribution (RTD1) were sent to OnQ rts (OSL, UK, v2.0). Rigid followed by deformable image registration, using the Demon's [24] algorithm, was then applied to the 'moving' CT1 scan in order to match the anatomy in the 'reference' CT2 scan, exporting the result as dCT1. The calculated deformation matrix was then automatically applied to both the RTS1 and RTD1 to generate the DIR-warped structures (dRTS1) and dose (***Dose_Deform***), respectively. Dose_Deform is therefore the direct deformation of the dose distribution calculated on the planning CT, using the DIR-generated deformation matrix deforming CT1 to match CT2.

The registration result image, dCT1, and structure set, dRTS1, were exported to Monaco where the original treatment plan was applied and new dose distribution calculated (***Dose_Recalc***). Hence, Dose_Recalc is essentially the dose calculated using the

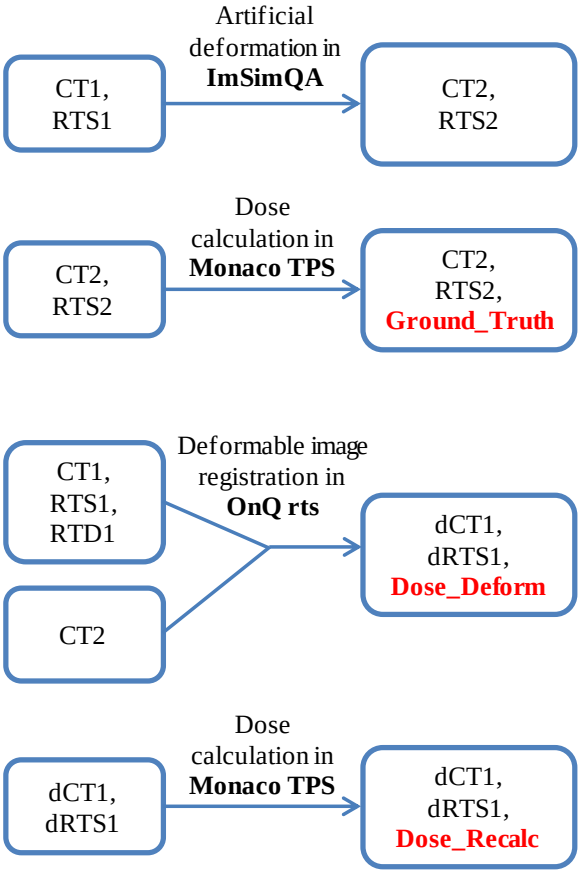

Figure 5.2: Workflow diagram illustrating the experimental procedure.

original plan on the planning CT scan deformed to match the artificially created image (n.b. although Dose_Recalc is then subject to errors in the DIR algorithm, these are also present in Dose_Deform method, thereby permitting fair comparison against Ground_Truth).

The robustness of both Dose_Deform and Dose_Recalc methodologies, using OnQ rts for H&N patients was then assessed by direct comparison with Ground_Truth.

Dose comparison and statistical analysis

Dose distributions generated by the three methods described above (i.e. Ground_Truth, Dose_Deform and Dose_Recalc) were transferred to ProSoma (MedCom, Germany, v3.3.252)

for evaluation. Each RTD file was individually loaded onto the appropriate artificially deformed CT (CT2) and the associated RTS for the generation of DVHs for brain, brainstem, spinal cord, contralateral parotid, optic chiasm, mandible and high-dose PTV. The agreement between the generated DVHs was visually inspected for qualitative analysis.

DVH information was then transferred into R programming language (www.r-project.org, v3.0.1) for quantitative and statistical analysis, comparing clinically relevant dose metrics such as the mean, median, maximum or minimum dose received by each structure. Specifically, the mean absolute differences between Dose_Recalc and Dose_Deform against Ground_Truth for these measures have been computed. Any differences between Dose_Recalc and Ground_Truth will quantify the impact of registration errors, while any additional differences between Dose_Deform and Ground_Truth will relate to the errors inherent in the Dose_Deform methodology. The non-parametric two-sided Wilcoxon signed-rank test was employed for statistical analysis, examining each artificial deformation separately to preserve statistical independence.

Further quantitative analysis was performed by measuring the 3D gamma passing rate, with 3% dose difference and 3 mm distance-to-agreement passing criteria ($\gamma_{3\%/3mm}$), of Ground_Truth against Dose_Deform and Dose_Recalc, using CERR [95] Matlab (MathWorks, Inc., Natick, Massachusetts) toolkit.

5.1.2 Results

Figure 5.3 shows a typical example of DVH comparison between the two delivered dose estimation strategies against the 'Ground_Truth' for brain, brainstem, spinal cord, contralateral parotid, mandible and PTV at a single fraction scenario for one patient. It is clear from this that Dose_Recalc (DR) overwhelmingly agrees with the Ground_Truth (GT) curve, which provides confidence in the accuracy of the DIR algorithm in this

scenario as any substantial errors would be apparent in this comparison. On the other hand, notable differences are observed between GT and Dose_Deform (DD). Of note, good agreement was observed between Dose_Deform and the DVHs for the original planned dose distribution, which is not included in the figure so as to maintain clarity. A considerable mismatch is observed at the low dose region in the brain DVH. Following thorough examination it was revealed that this was caused by a technical limitation of the employed software, whereby it is unable to maintain and process very small contoured regions (i.e. smaller than several voxels). During the deformation process, when a small contoured region gets shrunk to a region smaller than several voxels it gets eliminated.

Figures 5.4*a-c* illustrate the mean absolute differences between Ground_Truth against Dose_Deform and Dose_Recalc, comparing single fraction maximum and mean dose to spinal cord, brain, brainstem, optic chiasm and contralateral parotid for the three artificial deformations applied. Figure 5.4*d* shows a similar plot summarising the differences in the minimum and mean dose to the PTV. Again, the inaccuracies od Dose_Deform compared to Ground_Truth can be clearly observed.

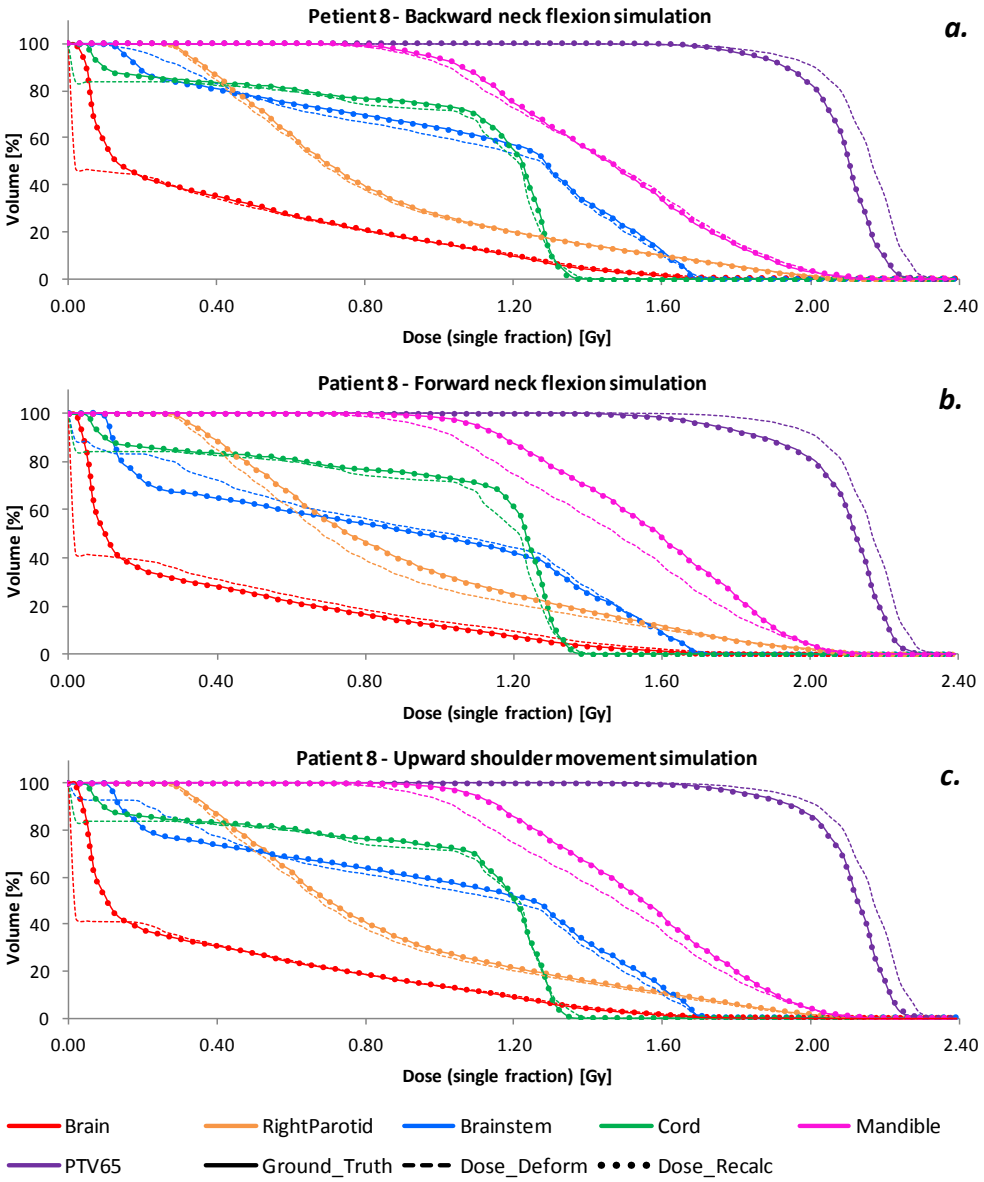

Figure 5.3: Dose volume histograms (DVH) comparing Ground_Truth (GT), Dose_Deform (DD) and Dose_Recalc (DR) methods for the four different artificial deformations (***a.-c.***) of a single patient.

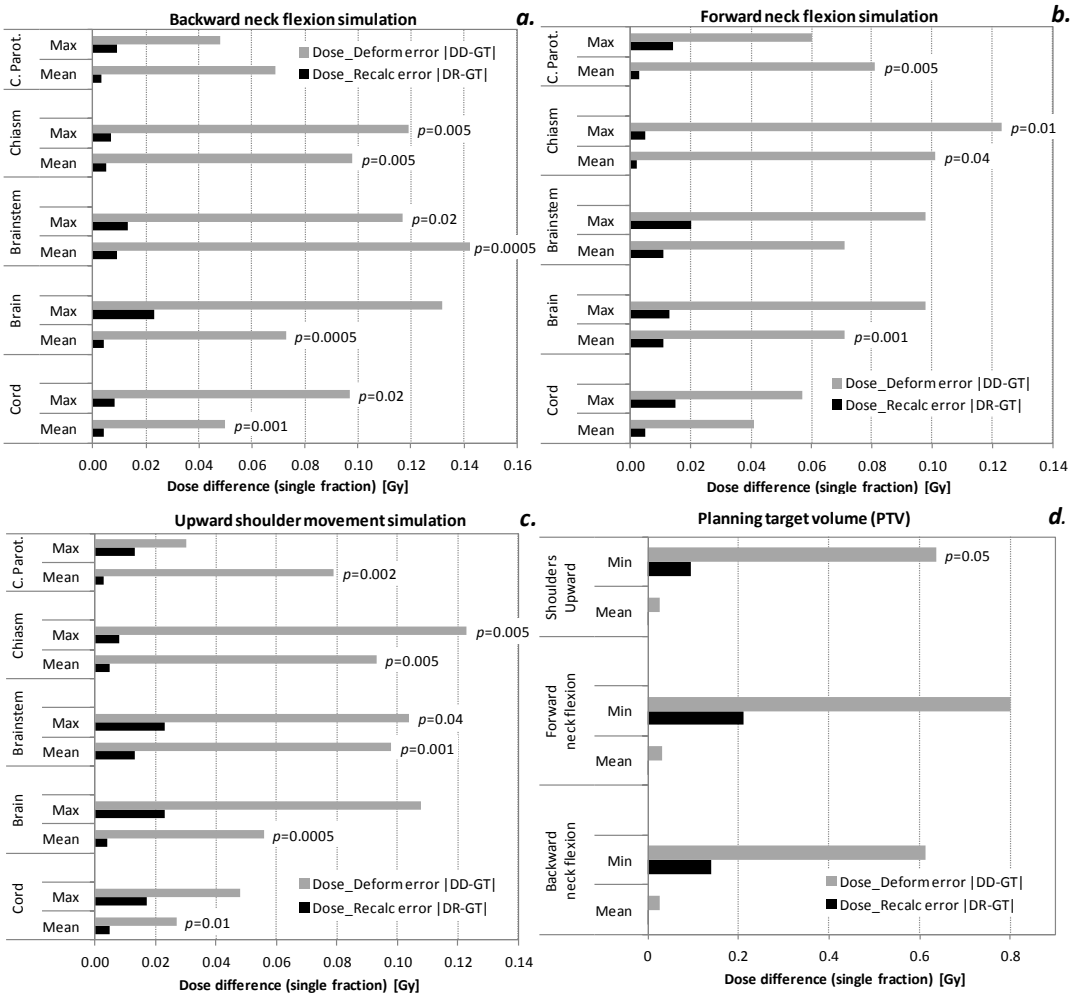

Figure 5.4: Mean absolute difference of maximum and mean doses for Ground_Truth (GT) against Dose_Deform (DD) and Dose_Recalc (DR), for ***a.*** backward neck flexion; ***b.*** forward neck flexion; ***c.*** upward shoulder movement; ***d.*** similar plot showing mean absolute difference of minimum and mean dose to the planning target volume (PTV). The P value is shown when the Wilcoxon test revealed $P \leqslant 0.05$ comparing DDvsGT and DRvsGT.

As observed in both the DVHs in Fig. 5.3 and the histograms in Fig. 5.4, Dose_Deform revealed poor agreement with the Ground_Truth with the differences revealing statistical significance (Wilcoxon test, $P \leqslant 0.05$) in many cases. Conversely, Dose_Recalc always had better agreement with Ground_Truth with no statistical significance in the observed differences. Indeed, extrapolating the mean discrepancies in Fig.5.4*a*, in estimating the

maximum dose to the brainstem due to neck flexion, over a 30 fraction course of treatment when using Dose_Deform methodology with the Demon's algorithm employed here, gives an expected value of approximately 3.6 Gy. This process revealed errors of between 0.3 and 7.8 Gy for the whole range of deformations and datasets used in this work. On the other hand, if Dose_Recalc method was used, the estimation of maximum dose to the brainstem would have a mean error of 0.5 Gy (range: 0.0 to 1.1 Gy).

The 3D γ analysis supported the above results: poor agreement was observed between Ground_Truth and Dose_Deform with mean $\gamma_{3\%/3mm}$ passing rate of 83.7%. Conversely, good agreement was found between Ground_Truth and Dose_Recalc with mean $\gamma_{3\%/3mm}$ passing rate of 99.8%. Local 3D γ analysis was also performed within various organs and regions of interest in order to expose local deviations, with the results summarised in Table 5.1. An example of 3D γ map is shown in Fig.5.5, in which backward neck flexion was applied (Def1). A high mismatch is observed between Ground_Truth and Dose_Deform in the anterior region of the patient and the penumbra behind the PTV (Fig. 5.5, top row), which was found to be due to the anatomy traversing the dose gradient present in the region as neck flexion was applied. Conversely, good agreement is observed between Ground_Truth and Dose_Recalc in the whole irradiated region (Fig.5.5, bottom row).

Local gamma analysis employs the local dose as a normalisation point for the determination of the percentage difference. On the other hand, global gamma analysis uses the point of maximum dose throughout the treatment plan as the normalisation point. This means that differences in regions of much lower dose than the maximum can be underestimated and potentially neglected when global gamma analysis is used. Therefore, local gamma was employed in this study.

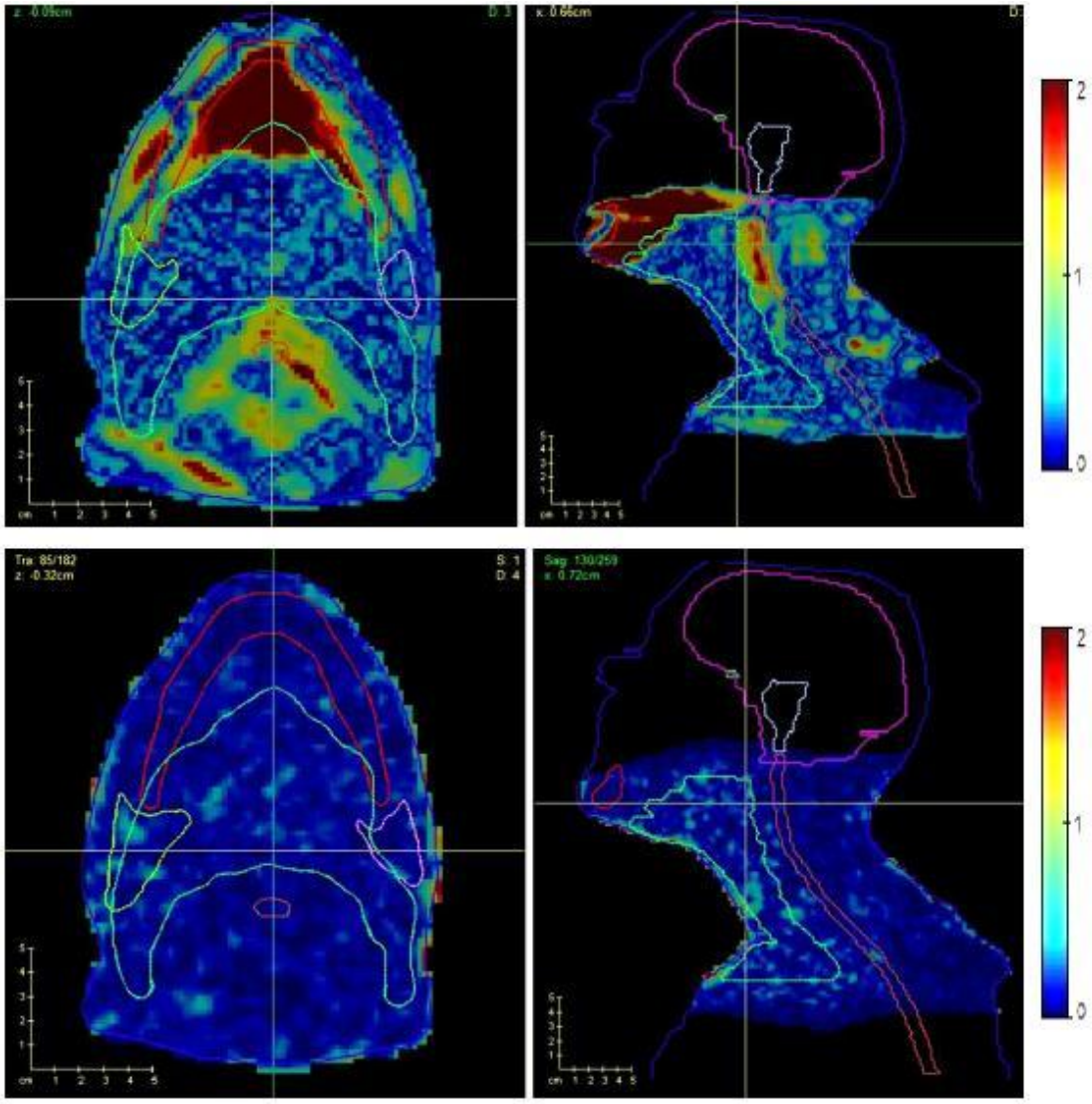

Figure 5.5: Examples of 3D gamma maps with $3\%/3mm$ passing criteria for: *top:* the warped dose case (i.e. comparison of Dose_Deform against Ground_Truth); and *bottom:* the dose recalculation on the DIR result case (i.e. Dose_Recalc against Ground_Truth).

Table 5.1: 3D gamma passing rate (with 3%/3mm criteria) for Ground_Truth against Dose_Deform and Dose_Recalc methods, in the entire irradiated volume (Whole volume) and individual organs and regions of interest.

	Ground_Truth vs Dose_Deform		Ground_Truth vs Dose_Recalc	
Anatomy	Mean [%]	Range [%]	Mean [%]	Range [%]
Whole volume	83.7	70.8 - 96.8	99.8	98.9 - 99.9
Brain	76.5	22.2 - 99.7	100	99.9 - 100
Brainstem	79.7	42.2 - 99.9	100	99.9 - 100
Spinal cord	85.1	54.9 - 99.9	100	99.9 - 100
PTV	95.6	77.7 - 99.9	99.9	98.8 - 100
C. parotid	85.1	55.1 - 99.9	99.9	99.7 - 100
Mandible	77.3	27.4 - 98.1	99.9	99.5 - 100

5.1.3 Discussion

This study utilised patient datasets and clinically realistic artificial deformations, as well as clinically optimised Monte Carlo dose calculations in an attempt to validate the dose warping process for the estimation of actual delivered dose at a radiotherapy treatment fraction, using methods that have been previously employed in the literature. Direct deformation of the planning dose to the daily anatomy following DIR, without a dose recalculation stage (Dose_Deform) is associated with statistically significant errors, while dose recalculated on the DIR result image (Dose_Recalc) was more accurate.

In the clinical setting, CBCT scans are often used for daily image guidance. As it is more challenging for DIR algorithms to register CT-to-CBCT than CT-to-CT images, DIR result in the former case is generally expected to be inferior. As the main focus of this study was the examination of the validity of direct dose deformation for actual

delivered dose estimation (Dose_Deform), the more straightforward case is employed to determine the inherent errors associated with Dose_Deform.

As part of this work, the accuracy of the DIR employed in this study was also quantified by allowing comparison of the artificially deformed structure sets (RTS2) with the DIR result RTS (dRTS1). The chosen evaluation metrics were the mean distance to conformity (MDC, described in Section 2.1.1, p. 31) and conformity index (CI):

$$CI = \frac{V_A \cap V_R}{V_A \cup V_R} \tag{5.1}$$

where V_A is the artificially deformed volume and V_R is the registration result volume. A perfect registration woulkd reveal $MDC = 0mm$ and $CI = 1$. Our registrations revealed a MDC always less than 3 mm with a mean value of 2 mm, while CI was greater than 0.70 in all cases with 0.75 mean value. This confirmed that the DIR algorithm used here performed well in the current scenarios. As noted earlier, the three dose distributions were copied on the appropriate artificially deformed image and the DVHs were created using the artificially deformed structure sets. This prevented any potential dosimetric uncertainties due to differences in structure volumes, resulting from small DIR inaccuracies.

The principal finding of this work, (i.e. that Dose_Deform is flawed and inaccurate) can be extended to the general case of dose warping for delivered dose estimation by considering the performance of a 'perfect' DIR algorithm using each methodology, as illustrated in Fig. 5.6. If a typical treatment plan aims to deliver $> 95\%$ of the prescribed dose to the entire target, then the boundaries of a given target broadly align with the isodose line for $> 95\%$ of dose to be delivered. In the presence of a real deformation, the actual target will not be identical to the target in the treatment plan and it may, therefore, be under-dosed when treatment is delivered (first row in Fig. 5.6). A hypothetical 'perfect' DIR algorithm will accurately deform the planning image to match the anatomy in the daily image and,

when the same deformation matrix directly applied to the planned dose distribution, as with Dose_Deform, it would perfectly warp the dose to the new anatomy demonstrating target coverage identical to the original treatment plan (second row in Fig. 5.6). Conversely, recalculating the dose distribution on the dataset resulting from DIR between the planning and daily image, with the same hypothetical 'perfect' DIR algorithm, would correctly reveal the actual delivered dose distribution (bottom row in Fig. 5.6). The same principle is applicable to any region of interest or organ at risk in realistic clinical scenarios. It should be noted that the results of the current work support this theoretical analysis, finding excellent agreement between both Dose_Recalc and Ground_Truth, as well as between Dose_Deform and the original planned dose distribution.

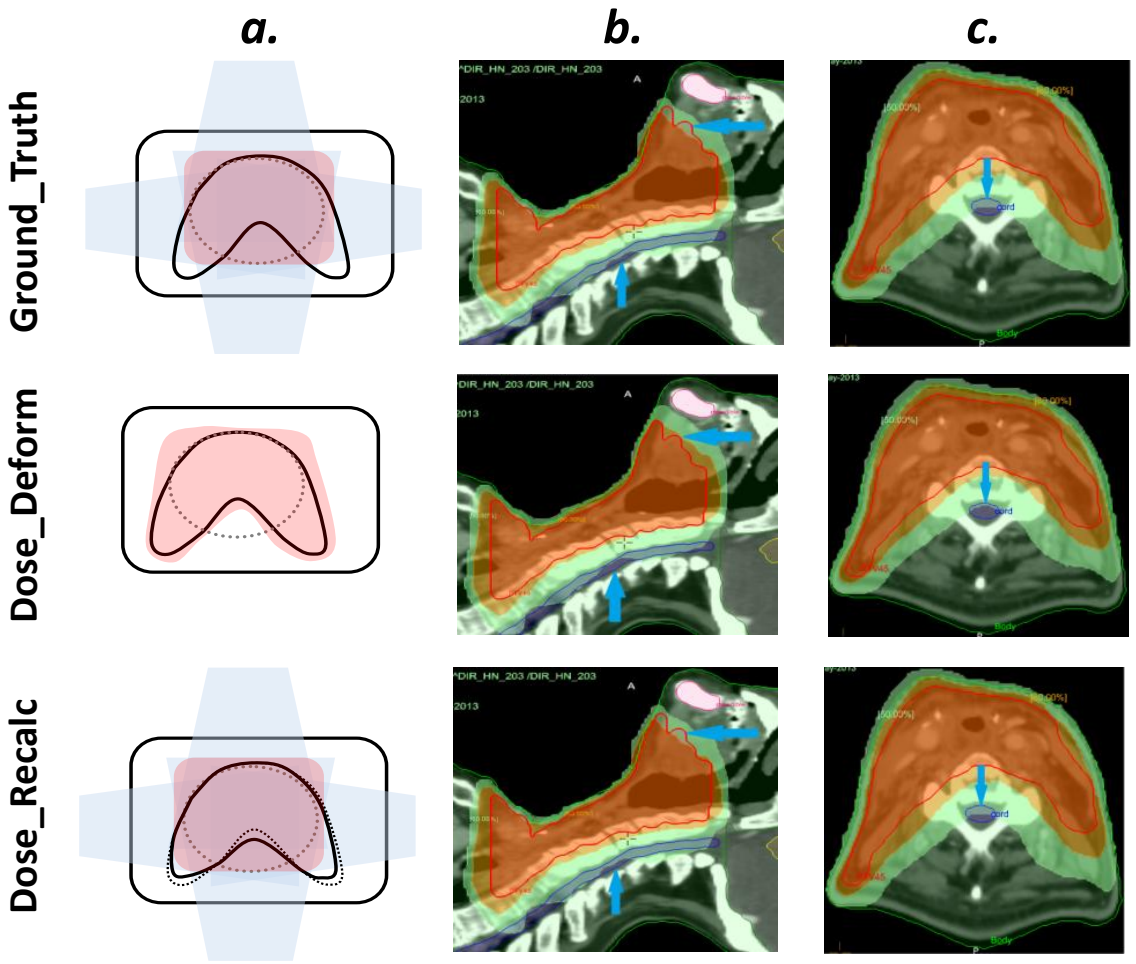

Figure 5.6: ***a.*** Schematic diagrams showing an original target (dashed circle), target after deformation (solid black line), target after deformable image registration (dashed line) and 95% isodose region (pink) for a simple exaggerated scenario. ***b.-c.*** Actual data acquired during this simulation study emphasizing the observed effect on the PTV and spinal cord. The 95% (orange) and 60% (green) isodose regions are shown, while arrows are pointing to regions with important differences between the methods under investigation.

With the increasing development of various DIR algorithms we would suggest that the better the performance of a given DIR, the better the actual delivered dose estimation using 'Dose_Recalc' method can get, while the opposite is true for 'Dose_Deform' method. It should be emphasised that direct deformation of the calculated dose to a reference anatomy is an essential step in any dose accumulation process (discussed in Chapter 6,

p. 92). Accordingly, the aim of this work has not been to refute the use of dose warping, but merely to highlight that precise adaptive workflows can produce invalid results if not properly understood.

5.1.4 Conclusions

Calculating the actual delivered dose at a radiotherapy fraction (dose of the day) by direct deformation of the original planning dose to the deformed anatomy following DIR, without a dose recalculation stage (Dose_Deform), was shown to be associated with statistically, and perhaps clinically significant errors. Dose recalculated on the original image as deformed to match the new anatomy (Dose_Recalc) was more accurate. These findings demonstrate the potential for misapplication of dose warping procedures observed in the literature and potentially in clinical practice.

5.2 Actual delivered dose estimation using on-line CBCT scans

This study investigates four techniques to allow direct dose calculation on on-line acquired CBCT images or the representation of the anatomy in these images, for the estimation of actual delivered dose at a given radiotherapy fraction.

5.2.1 Methods

Data selection

CT (Philips Brilliance, Philips Healthcare, Netherlands) and CBCT (XVI^{TM}, Elekta AB, Sweden) images of the Catphan® and Rando® (The Phantom Laboratory, Salem, NY) phantoms were acquired. Two VMAT plans were created for the Rando phantom, one simulating a high risk prostate cancer treatment (with a prescribed dose of 74 Gy to the primary PTV, 71 Gy to the secondary and 60 Gy to the tertiary PTV, in 37 fractions) and one simulating a head-and-neck cancer treatment (with 65 Gy to the primary PTV, 60 and 54 Gy to the secondary and tertiary PTVs, in 30 fractions).

Moreover, the planning CT and the online CBCT acquired prior to the first fraction of a head-and-neck and a prostate cancer patient, previously treated with similar plans, were retrospectively selected. The relevant patient anatomy in these CBCT scans was visually inspected and confirmed as almost identical to that observed in the planning CT images. All datasets were transferred to RayStation (RaySearch Laboratories AB, Stockholm, Sweden) where dose calculations were performed.

Preparation of CBCTs for dose calculation

Four approaches to allow direct dose calculation on CBCT datasets, previously outlined in Section 1.4.3, p. 17, were evaluated in this study:

- ***Phantom based HU-to-density curves*** were created using the CBCT scans of the Catphan® phantom. The image HU in regions with known electron density are associated for the creation of CBCT specific HU-to-density curves.

- ***Bulk density override*** was performed by assigning water equivalent HU to the whole region of the phantoms seen in the CBCT scans.

- ***DIR based density mapping*** methodology was performed as previously described ('Dose_Recalc' method in Section 5.1.1, p. 71) but using both an artificially deformed CBCT and CT images. Briefly, the CT and CBCT scans were artificially deformed with the same clinically realistic deformations. The original CT was then matched to the deformed CBCT following DIR. The treatment plan was then calculated on the artificially deformed CT and the DIR result CT, for direct comparison.

- ***Tissue based HU-to-density curves*** were created with the manual method offered by RayStation system. This approach offers sliding bar windows for the adjustment of tissue specific density thresholds based on the image's grey level histogram (Fig. 5.7) and directly visualised on the image as shown in Fig. 5.8.

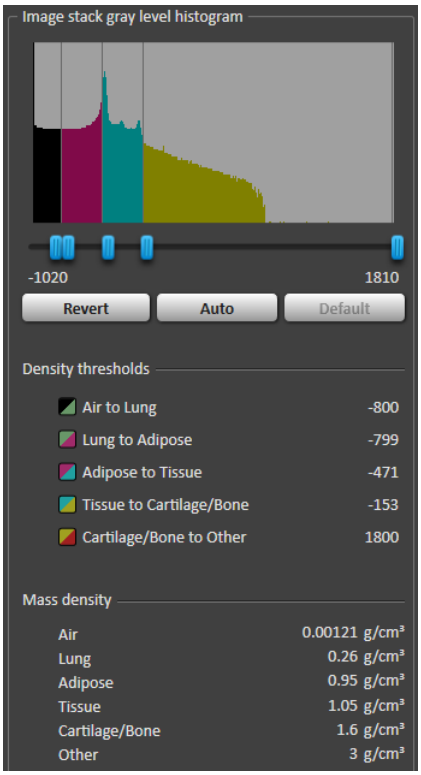

Figure 5.7: RayStation's (RaySearch Laboratories AB, Stockholm, Sweden) implementation for the generation of tissue specific HU-to-density curves for CBCT images.

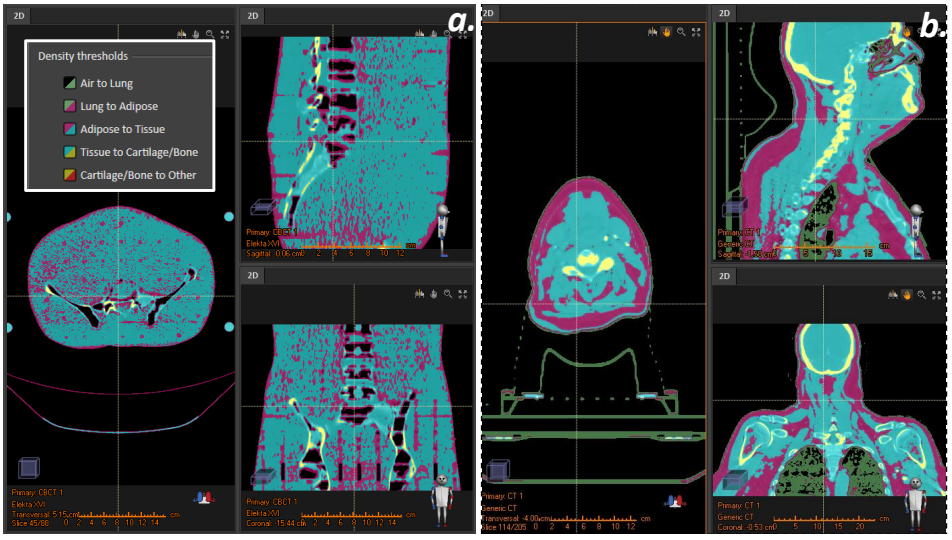

Figure 5.8: Examples of tissue segmentation following density thresholding on ***a.*** the pelvic area of Rando® phantom, and ***b.*** a head-and-neck cancer patient.

Evaluation of dose calculation on CBCT images

Original treatment plans were transferred to the CBCT images, following rigid image registration and recalculated with identical parameters, for all CBCT preparation methods listed above. Original dose distributions calculated on the CT datasets and those recalculated on the CBCT images were exported to CERR Matlab toolkit [95] and compared using global and local 3D gamma analysis with 1%/1mm ($\gamma_{1\%/1mm}$) and 0.5%/0.5mm ($\gamma_{0.5\%/0.5mm}$) criteria.

5.2.2 Results

The results of global and local 3D gamma analysis for the prostate cancer treatment scenarios are listed in Table 5.2, with the respective results for head-and-neck cancer treatment scenarios listed in Table 5.3. It is observed that, for the test cases used herein, all techniques for dose calculations using online CBCT images provide acceptable results (i.e. vast majority, $> 95\%$, of voxels in recalculated dose distributions agree with original dose within 1%/1mm criteria).

In phantom based HU-to-density curve generation, for which the Catphan® phantom was employed, the majority of dosimetric differences were observed around bony anatomy in both pelvic and head-and-neck scenarios. This observation correlates with Guan et al. [62] results, who exposed higher dosimetric differences when the Catphan® phantom was used and attributed these to its Teflon insert which cannot accurately mimic bone, as well as the small size of inserts that cannot provide correct HU under high scatter conditions.

Table 5.2: Percentage of voxels failing the 3D gamma analysis criteria comparing original dose distributions calculated on planning CT against dose recalculated on CBCT, following each CBCT preparation approach, in prostate cancer treatment scenarios.

		Phantom based		Bulk density		Tissue specific		DIR based	
Case	Site	$\gamma_{1\%/1mm}$ [%]	$\gamma_{0.5\%/0.5mm}$ [%]	$\gamma_{1\%/1mm}$ [%]	$\gamma_{0.5\%/0.5mm}$ [%]	$\gamma_{1\%/1mm}$ [%]	$\gamma_{0.5\%/0.5mm}$ [%]	$\gamma_{1\%/1mm}$ [%]	$\gamma_{0.5\%/0.5mm}$ [%]
	Body	0.6	1.5	1.3	4.6	0.0	0.3	0.2	0.6
Rando	Prostate	4.3	11.6	7.7	52.6	0.0	0.1	0.0	0.3
pelvis	Bladder	2.4	6.1	5.6	23.0	0.0	0.4	0.2	0.9
	Rectum	3.2	8.7	21.3	41.2	0.0	1.1	0.9	1.7
	Body	0.5	1.6	1.5	5.1	0.0	0.4	0.4	1.1
Prostate	Prostate	4.6	12.7	8.4	58.9	0.0	0.3	0.3	0.8
patient	Bladder	2.3	5.8	5.8	22.4	0.0	0.7	0.9	1.3
	Rectum	2.9	7.3	23.2	42.7	0.0	1.6	1.4	1.9

DIR: Deformable image registration

Table 5.3: Percentage of voxels failing the 3D gamma analysis criteria comparing original dose distributions calculated on planning CT against dose recalculated on CBCT, following each CBCT preparation approach, in head-and-neck cancer treatment scenarios.

		Phantom based		Bulk density		Tissue specific		DIR based	
Case	Site	$\gamma_{1\%/1mm}$ [%]	$\gamma_{0.5\%/0.5mm}$ [%]	$\gamma_{1\%/1mm}$ [%]	$\gamma_{0.5\%/0.5mm}$ [%]	$\gamma_{1\%/1mm}$ [%]	$\gamma_{0.5\%/0.5mm}$ [%]	$\gamma_{1\%/1mm}$ [%]	$\gamma_{0.5\%/0.5mm}$ [%]
	Body	0.8	1.7	1.5	6.2	0.0	0.3	0.3	0.7
Rando	PTV	3.2	10.6	6.6	44.2	0.0	0.5	0.3	0.8
H&N	Cord	1.8	9.1	5.3	35.1	0.0	0.2	0.5	1.5
	Brainstem	0.2	0.7	0.9	1.7	0.0	0.0	0.1	0.4
	Body	1.3	3.4	1.9	7.4	0.0	0.5	0.5	1.0
H&N	PTV	3.8	10.2	8.0	55.3	0.0	0.5	0.7	1.9
patient	Cord	2.1	8.3	8.3	59.6	0.0	0.3	0.6	1.1
	Brainstem	0.2	0.5	1.4	5.2	0.0	0.1	0.2	0.6

DIR: Deformable image registration

H&N: Head-and-neck

PTV: Planning target volume

Cord: Spinal cord

The simple bulk density override approach employed herein, with an override of the whole body with water equivalent HU, proved to be the poorest performing of the investigated methods. On the other hand, the tissue specific HU-to-density curve generation technique was the most accurate in all test cases, with the DIR based density mapping technique revealing similar accuracy.

5.2.3 Discussion

Four approaches to allow dose calculation on CBCT scans, or a representation of the anatomy in online acquired CBCT images, have been investigated. All techniques revealed acceptable results in the test cases examined, with the tissue specific HU-to-density curve generation revealing the highest accuracy. Even though limited test cases were employed for this study, the results agree with the general conclusions of previous studies which showed that certain density override and tissue specific HU-to-density curves can provide more accurate dose calculations [61; 64].

Tissue specific HU-to-density curve generation techniques, and especially the example investigated herein (i.e. the manual procedure offered by the RayStation system), can be subjective and user dependent, subject to the choice of each tissue grey level ranges. For example, bony structures typically have high density regions only on the periphery with lower density in the center, while differences between adipose and tissue are occasionally difficult to distinguish. The impact of different tissue density thresholds was investigated by gradually varying these values using the sliding bars shown in Fig. 5.7, p. 86, while maintaining visually plausible tissue segmentation results, as seen in Fig. 5.8, p. 86. Dose calculations were performed on the resulting images and compared, exposing differences of up to $\pm 1.5\%$. Therefore, to minimise this uncertainty, the optimum parameters were identified before this methodology was evaluated and compared to alternative techniques.

The DIR based density mapping technique showed slightly inferior accuracy to the tissue specific HU-to-density curve generation technique. However, provided the DIR algorithm can be proven accurate and robust and with a practical quality assurance procedure put in place, this method has two important advantages: (a) it can be highly automated, and (b) DIR between the reference CT and the daily CBCT would need to be performed anyway, as part of inter-fractional treatment monitoring by deformable dose accumulation.

In terms of clinical workload, phantom based HU-to-density curve generation may be considered the most efficient as generation of these curves can be created once, for every CBCT scanning protocol, with quality assurance procedures performed thereafter. Other techniques can be potentially automated but, with the exception of bulk density override, their result would require review by an appropriately trained member of staff.

5.2.4 Conclusions

Four techniques to allow calculation of delivered dose at a radiotherapy fraction based on the online acquired CBCT were evaluated, revealing acceptable accuracy. The generation tissue specific HU-to-density curves, following manual grey level threshold segmentation of the CBCT images, revealed the most accurate results but was found to be subjective and user dependent. DIR based density mapping, which revealed similarly good results, has the potential to be a good alternative.

Chapter 6

Validation of Dose Warping

This chapter demonstrates a workflow for the evaluation of dose warping following deformable image registration for inter-fractional dose accumulation. This study −and the majority of this chapter's content− has been accepted for publication by the British Journal of Radiology [96].

Robust and clinically relevant validation methodologies of procedures such as dose warping for inter-fraction dose accumulation are of great importance to inform physicists and clinicians on the accuracy of DIR algorithms for this task and the uncertainties involved.

6.1 Validation of dose warping in inter-fractional dose accumulation

In this study, a workflow for the robust validation of DIR and dose warping is presented, using patient images artificially deformed with clinically realistic deformations and clinically optimised Monte Carlo dose calculations of IMRT plans, quantifying both the spatial errors in the deformable registration and their dosimetric impact when applied to dose

accumulation. In contrast to previously proposed evaluation procedures (discussed in Section 1.4.4, p. 20), this method examines and reports dose warping uncertainties under clinically relevant scenarios. Although this validation workflow is applicable for different DIRs and clinical indications, it was here employed specifically for the evaluation of a commercial software (OnQ rts, Oncology Systems Limited, Shrewsbury, UK) in head and neck cancer (H&N) patients.

6.1.1 Methods

Data selection

A total of twelve H&N patient datasets, consisting of a DICOM CT dataset, associated structure set (RTS), 6 MV IMRT plan (RTP) and corresponding dose distribution (RTD), were selected for this retrospective simulation study. Of the twelve patients used, four were treated for uni-lateral and eight for bi-lateral H&N cancer, while all treatment plans were created using the Monte Carlo dose calculation algorithm in Monaco (Elekta AB, Sweden, v3.20) treatment planning system (TPS). Treatment plans ranged in complexity with 1-3 target volumes (1 PTV: $n = 5$, 2 PTVs: $n = 3$, 3 PTVs: $n = 4$) while the prescribed dose per fraction ranged from 2 to 3 Gy.

Artificial deformations

For each patient, the planning CT dataset (pCT) and structures (pRTS) were transferred to ImSimQA software, where four clinically realistic artificial deformations were manually introduced to create four 'CTn' and 'RTSn' datasets (i.e. n referring to the n^{th} deformation in an assumed inter-fractional dose accumulation workflow), in a process previously described in Section 4.1.1 (p. 56). Backward (Def1) and forward (Def2) neck flexion, weight loss (Def3) and upward shoulder movement (Def4) were simulated as shown in Fig. 5.1 (p. 70). Artificial deformations were based on actual observations during image

guidance at our institution (Queen Elizabeth Hospital, Birmingham, UK) and visually inspected by a specialist consultant and a specialist radiographer for clinical relevance. Volume conservation in critical structures was quantitatively investigated by comparing the original and deformed volumes.

Deformable image registration and dose warping validation

The performance of the DIR algorithm has been evaluated with the workflow shown in Fig. 6.1. Following the application of clinically artificial deformations on the planning CT images (pCT), the new deformed images (CTn) were sent to Monaco TPS where the original treatment plan was applied as a 'QA plan' to ensure identical conditions (i.e. beam arrangement, isocentre, segment positions, monitor units and monitor units per segment). The new dose calculated (**Dose_True**) was considered the 'true' dose since this would be the distribution delivered to the patient if the original plan was to be applied on this anatomy. CTn with the associated structure set (RTSn) and the calculated dose distribution (Dose_True) were then loaded to OnQ rts together with the pCT, where rigid followed by DIR was performed. pCT was treated as the 'reference' and each CTn as the 'moving' image (i.e. deforming back to the original anatomy). The CTn, RTSn and associated Dose_True were deformed accordingly by applying the deformation matrix calculated during DIR (dynamic vector field). This resulted in warped image (dCTn), structure set (dRTSn) and dose distribution (**Dose_Warp**), which were then also copied to pCT.

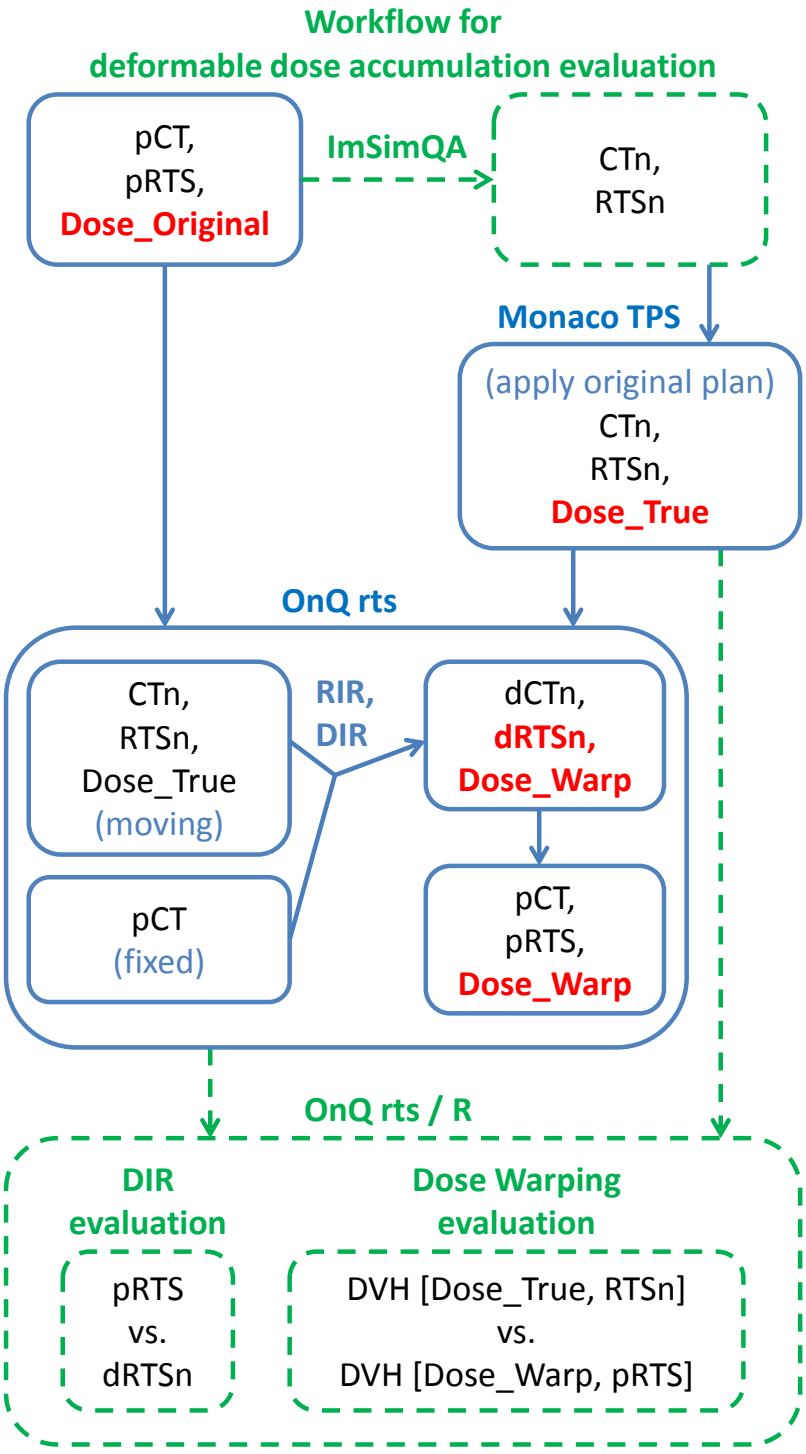

Figure 6.1: Flow chart summarizing the dose warping evaluation workflow.

A 'perfect' DIR algorithm would be able to bring the artificially deformed images and structures back to their original configuration. The registration result was evaluated following the procedure described in Chapter 4, using the conformity index (CI; Eq. 5.1, p. 80) and mean distance to conformity (MDC; definition in p. 31) metrics.

Furthermore, for the above situations in which the mass of organs under investigation is conserved, the same hypothetical 'perfect' DIR algorithm would result in agreement between the dose-volume histograms (DVH) analyses of [Dose_Warp, pRTS] and [Dose_True, dRTSn]. Differences in these values can therefore be attributed to DIR inaccuracies and are evaluated by both visual comparison of the DVHs and by quantitative differences in clinically relevant dose metrics. The two dose distributions were also compared with the original plan (**Dose_Original**) in order to expose the dose delivery errors owing to the introduction of artificial deformation on the original anatomy and the ability of dose warping to estimate the effect. Evaluation was herein performed utilizing the DVHs of brain, brainstem, larynx, spinal cord, contralateral parotid and the clinical target volume (CTV). Spatial evaluation was performed by dose subtraction and 3D gamma analysis with $2\%/1mm$ criteria calculating the passing ratio for all voxels receiving $> 20\%$ of the maximum dose, for Dose_Original, Dose_True and Dose_Warp.

Statistical analysis

The non-parametric two-sided Wilcoxon signed-rank test was used for statistical analysis of dosimetric results, comparing the mean absolute differences in mean, median, minimum or maximum dose within certain structures as calculated by the Dose_True, Dose_Warp and Dose_Original. This method was chosen since the data did not follow a normal distribution. Statistical analysis was performed using R programming language (R Core Team, Vienna, Austria; www.r-project.org) and was carried out for each type of artificial deformation individually in order to retain statistical independence.

6.1.2 Results

Deformable image registration evaluation

The evaluation of the DIR algorithm's performance under the four artificial test conditions is shown in Fig. 6.2. Figure 6.2a exposes the respective mean values of MDC over the 12 patient data sets used. With the exception of chiasm, the analysis of all volumes revealed an MDC of 1.2-2.0 mm. Considering the voxel size of the CT scans used for this study (1.2 mm in the x and y directions, and 2.0 mm in the z-direction), it is observed that the average registration result is accurate to within 1-2 voxels.

Figure 6.2b shows the mean CI value for each artificial deformation for the data sets employed. Brainstem, contralateral parotid and mandible revealed mean CI values of 0.7 while cord, brain, body and CTV a CI of $\geqslant 0.8$. Note that even though CTV is not an anatomically definable structure, it was incorporated in the DIR evaluation process as it would be used for the dosimetric analysis. Chiasm was the organ that revealed the lowest CI values (0.4-0.2). As chiasm only covers 2-3 slices, an average inaccuracy of 1-2 voxels in the z-direction results in very low CI.

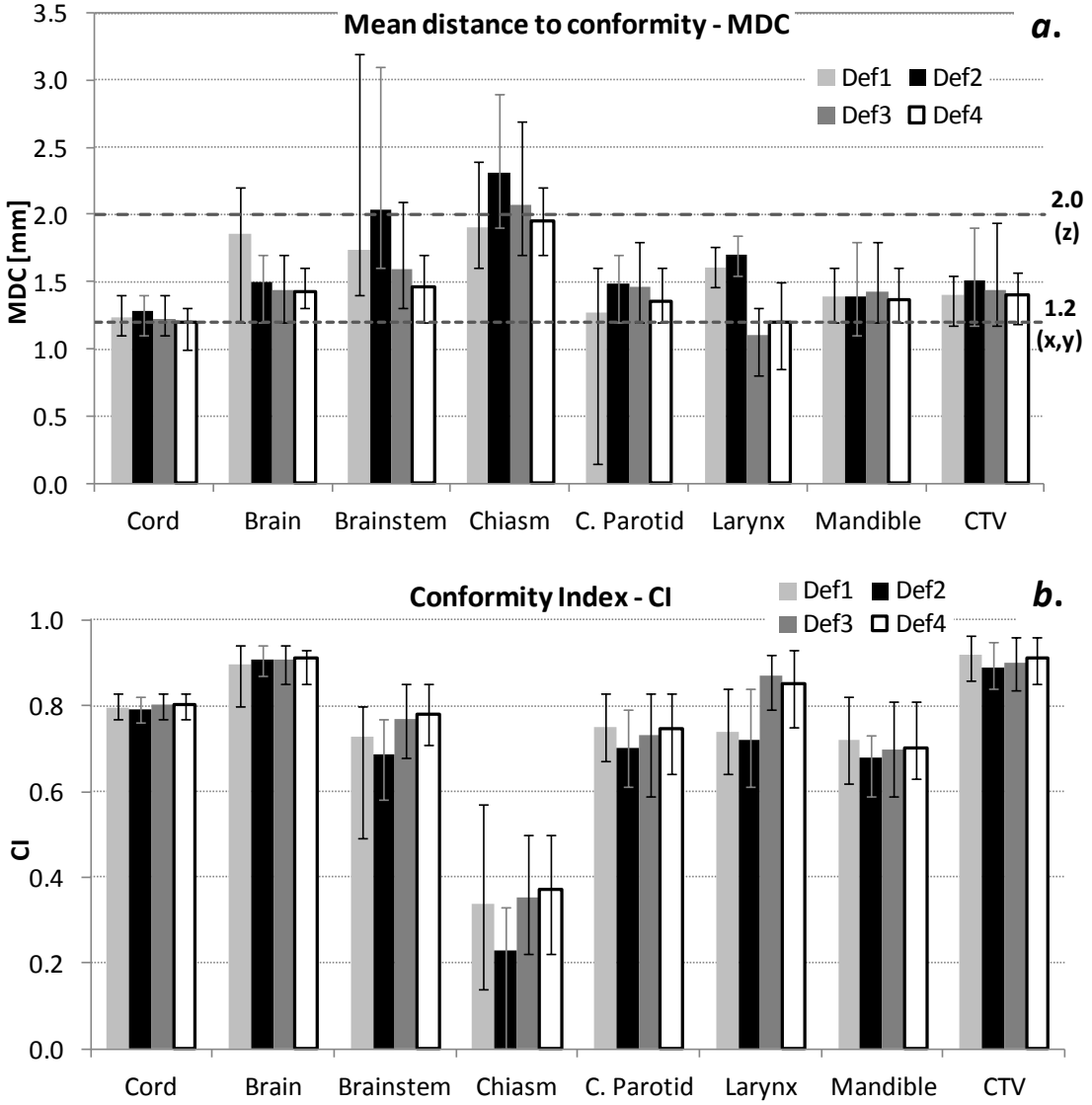

Figure 6.2: Evaluation analysis of deformable image registration (DIR) in the 12 patient cases under investigation, for the 4 artificial deformations applied (Def1-4). ***a.*** Average mean distance to conformity (MDC); ***b.*** average conformity index (CI), for spinal cord, brain, brainstem, chiasm, contralateral parotid (C. Parotid), larynx, mandible and clinical target volume (CTV). The error bars in both plots represent the range of values observed, while the horizontal dashed lines in ***a.*** represent the voxel size of the CT scans used in x, y and z direction.

Dose warping validation

Examples of single fraction DVHs for the Dose_Original, Dose_True and Dose_Warp for a typical patient in the four artificial deformations are shown in Fig 6.3. The differences observed between the Dose_Original and Dose_True curves clearly demonstrate the expected dose delivery errors in the presence of the artificial clinically realistic deformations. The warped dose distribution revealed a generally good agreement with the Dose_True. However, regions receiving low dose (i.e. $< 20\%$ of prescribed dose) were occasionally underestimated, as observed in brain, brainstem and spinal cord DVHs in Fig. 6.3

Figure 6.4 shows a comparison of the actual dose delivery change owing to the presence of deformations (i.e. $|\text{Dose\_True} - \text{Dose\_Original}|$), against the estimated change after warping back to the original anatomy (i.e. $|\text{Dose\_Warp} - \text{Dose\_Original}|$), which is subject to errors owing to DIR uncertainties. For the perfect DIR algorithm and when mass is conserved in the regions under investigation, these values would be identical in all cases. This comparison revealed good agreement (0.02 ± 0.03 Gy) with no statistically significant differences ($0.6 \geqslant P \geqslant 0.2$ in all cases). It is observed that the minimum dose received by the CTV is occasionally, but not significantly (i.e. $0.4 \geqslant P \geqslant 0.3$ in all cases), underestimated after dose warping as small spatial uncertainties have bigger effect in regions with steep dose gradients.

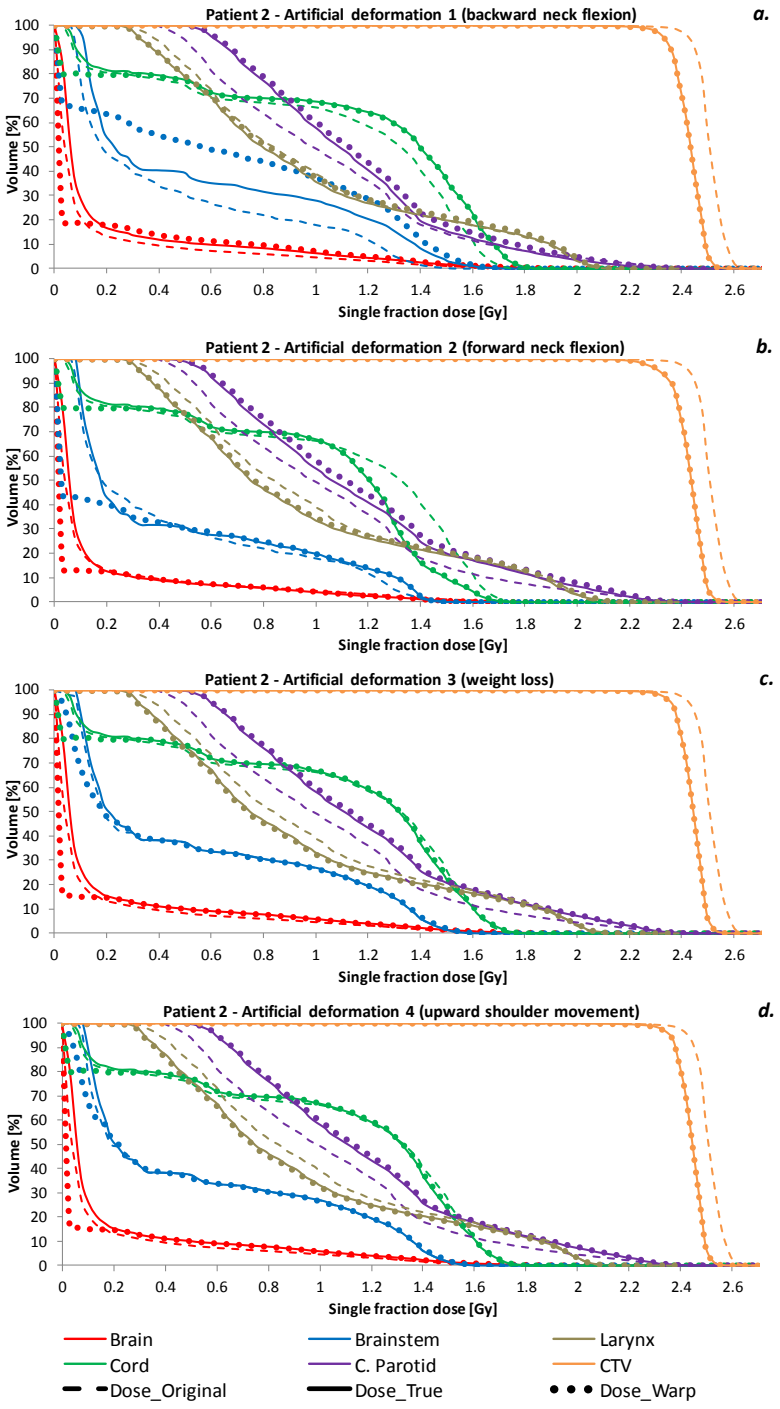

Figure 6.3: Dose-volume histograms comparing the clinically prescribed [Dose_Original, pRTS], recalculated on the artificial image [Dose_True, dRTSn], and warped [Dose_Warp, pRTS] dose distributions, of artificial deformations 1-4 (***a.-d.***) for a single typical patient. In situations where the mass of these organs is conserved, a 'perfect' deformable image registration would show agreement between [Dose_True, dRTSn] and [Dose_Warp, pRTS]. C. Parotid: contralateral parotid; CTV: clinical target volume.

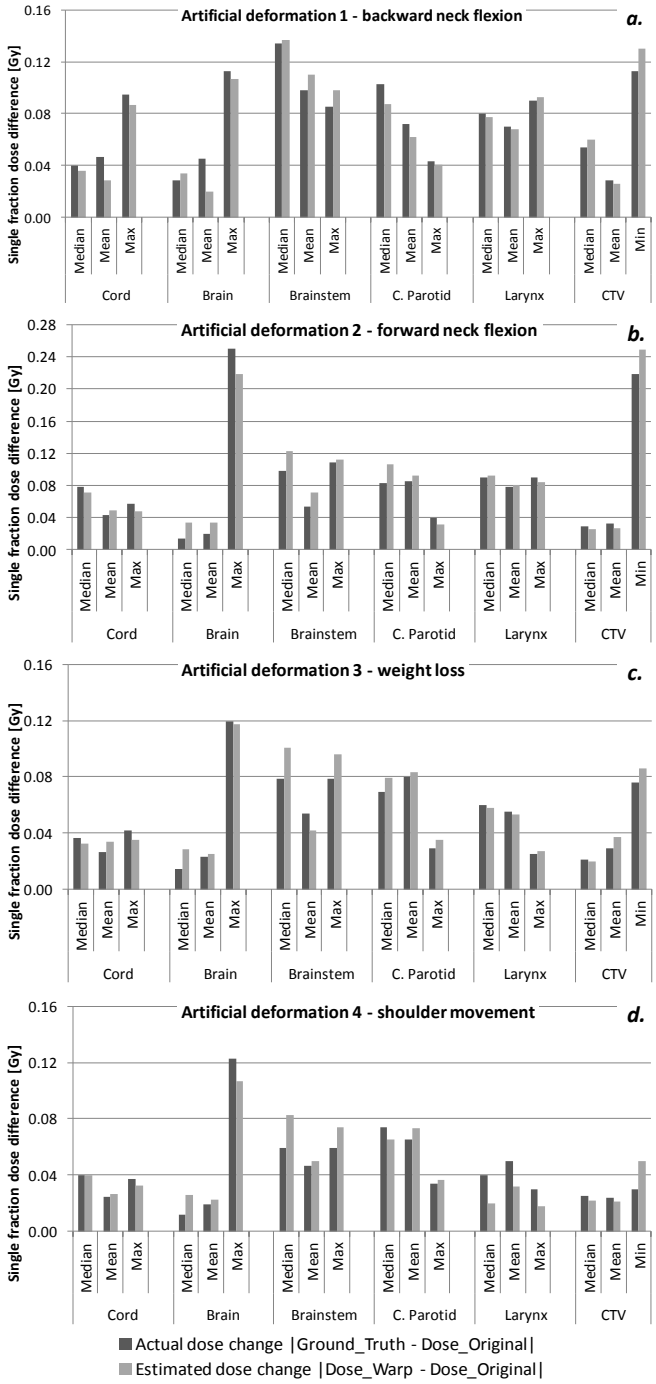

Figure 6.4: The mean actual dose delivery change introduced by anatomical deformations, |Dose_True − Dose_Original| against the mean estimated dose change by deformable image registration-based dose warping |Dose_Warp − Dose_Original|, for the 12 patient cases investigated and the 4 artificial deformations applied (***a.-d.***), in spinal cord, brain, brainstem, contralateral parotid (C. Parotid), larynx and clinical target volume (CTV). In situations where the mass of these organs is conserved, a 'perfect' deformable image registration would result in these values beaing the same for all situations. Max: maximum; min: minimum.

To further quantify the clinically relevant accuracy of the dose warping process, Dose_Warp was compared against Dose_True in terms of median, mean and maximum or minimum dose to the brain, brainstem, spinal cord, contralateral parotid and CTV, with the observed differences revealing no statistical significance (i.e. $0.5 \geqslant P \geqslant 0.2$ in all cases).

The importance of the DVH parameter utilised depends on the structure and type of each underlying organ or tissue. Organs with serial architecture (e.g. spinal cord) typically have a maximum dose threshold beyond which the whole organ may be damaged and, therefore, maximum dose would be the most relevant metric. On the other hand, organs with parallel architecture (e.g. parotid gland) may be able to tolerate a higher dose to a small region without affecting the function of the organ as a whole. In this case the mean or median dose would be more relevant metrics.

It should be emphasised that small uncertainties can be critical when the original treatment plan is created to the OAR's or target region's tolerance. Therefore, particularly thorough evaluation of the DIR result consideration of dose warping uncertainties should be carried out in such circumstances.

Even though non-statistically significant throughout a total of 48 cases investigated, a number of substantial DIR and dose-warping errors have been observed, as, for example, in Fig. 6.3a; in this instance, Dose_Warp reveals good agreement with Dose_True for all organs except the brainstem, for which a difference of 0.35 and 0.25 Gy in the median and maximum dose estimation was observed, respectively, being the result of an outlier MDC error of 3.2 mm in DIR.

Figure 6.5 shows examples of gamma maps ($\gamma_{2\%/1mm}$) between Dose_Original, Dose_True and Dose_Warp. Gamma analysis of Dose_Original vs Dose_True (Fig. 6.5a) illustrates the differences in dose distribution owing to the applied anatomical deformation. Dose_Original vs Dose_Warp (Fig. 6.5b) shows the same differences after warping of Dose_True back to

the original anatomy, while Dose_True vs Dose_Warp (Fig. 6.5c) illustrates the effect of warping the recalculated dose distribution to the reference anatomy. Subsequent review of the 3D gamma maps confirmed that the regions of largest disagreement are situated in regions that combine dose gradient and displacement, as would be expected, which also demonstrated in the dose difference map (Fig. 6.5d). Gamma analysis of Dose_True vs Dose_Warp in the forward neck flexion simulation (Fig. 6.5c) revealed greater discrepancies at the chin and neck area, which experienced the greatest displacement, while small differences were observed at the inner body region where anatomical displacement was smaller.

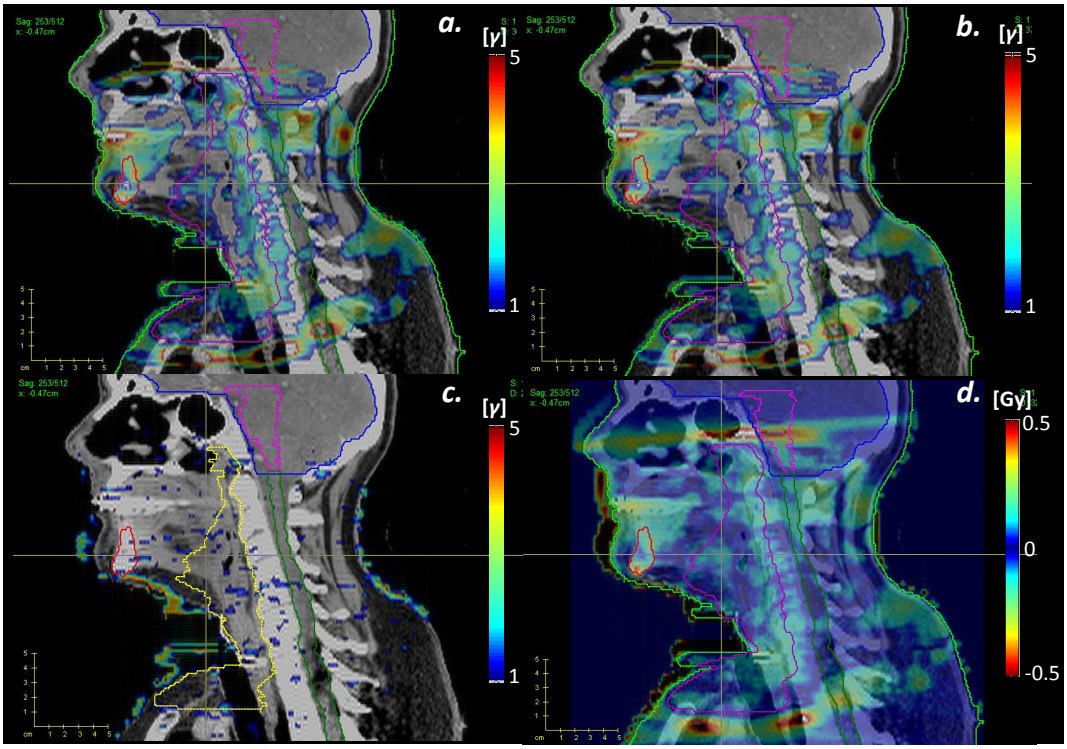

Figure 6.5: Three-dimensional gamma analysis (2%/1mm criteria) of ***a.*** Dose_Original vs Dose_True; ***b.*** Dose_Original vs Dose_Warp; ***c.*** Dose_True vs Dose_Warp, and ***d.*** dose subtraction Dose_Original – Dose_True

6.1.3 Discussion

This study demonstrates a DIR algorithm validation workflow for image registration and dose warping throughout fractionated radiotherapy, while overcoming the three limitations of recent studies that employed physical phantoms for the evaluation of dose warping [49; 50; 52; 71; 72] that were noted earlier (Section 1.4.4, p. 20): limited image complexity, limited dose measurement accuracy and inability to directly transfer the findings to clinical scenarios. The workflow has been used for the validation of a commercial algorithm which, largely, demonstrated accurate predictions of the actual dose distributions under four clinically realistic deformation scenarios. All analysis was performed for single fraction situations in order to simulate the scenario of daily treatment monitoring that would be most sensitive to errors, as it excludes averaging effects from daily anatomical variations. Recalculated dose distributions were successfully warped to the reference anatomy [Dose_Warp, pRTS] and revealed good agreement to the ground truth [Dose_True, dRTSn], with the observed differences having no statistical significance (Wilcoxon test, $0.5 \geqslant P \geqslant 0.2$). However, considerable registration and dose-warping errors have been observed in a small number of cases, a finding that illustrates the importance of such validation work as a means of highlighting and understanding the nature and extent of errors in dose accumulation processes and the need for visual inspection of DIR results. These registration and dose-warping errors were primarily observed in low contrast regions where intensity-based-only DIR algorithms are known to have inferior performance.

There are three limitations to the current work. First, deformable registration and dose warping between two planning CT quality scans, as herein, is perhaps more robust than would be encountered clinically; even though some radiotherapy centres use CT-on-rails for daily imaging, typical IGRT procedures employ a range of alternative modalities,

including CBCT, MVCT or mega-voltage CBCT. These scans have lower image quality and smaller field-of-view than does CT, which may further compromise DIR performance. The results of the present study could represent a 'best case' scenario for the use of this algorithm in H&N cases.

Second, as the proposed workflow employs DVH analysis, the mass and volume of organs under investigation must be conserved. As previously observed, the volume of tumours and certain organs occasionally decreases during the course of fractionated H&N radiotherapy [97], a situation that was not simulated herein. Nevertheless, this limitation is also present in other dose warping evaluation workflows, such as the ones employing deformable dosimetric gel phantoms, while mass conservation is an underlying assumption in many DIR algorithms.

The third limitation of the present work is the difficulty in quantitatively determining the spatial position of observed differences using, for instance, gamma analysis as an evaluation test, since the warped (Dose_Warp) and calculated (Dose_True) dose distributions are associated with different anatomies. A gamma analysis between them includes effects of both real anatomical change and errors in deformable registration, the separation of which is challenging. DVH analysis is applicable however, which is one of the main examination tools used by physicians and physicists in a clinical setting. This could be construed as an advantage of the current evaluation workflow, providing clinically meaningful and organ-specific uncertainty measures, a feature absent from most validation work in the area.

As discussed previously, a number of studies have investigated techniques for the evaluation of dose warping using deformable dosimetric and non-dosimetric gel phantoms [49; 50; 52; 71; 72]. Even though such evaluation methodologies offer valuable advantages, such as the ability to perform quantitative spatial dosimetric evaluation, their

ability to offer comprehensive conclusions under clinical conditions is hindered by limited anatomical complexity, dosimetric and dose-reading uncertainties and inability of clinical interpretation of results. These techniques could therefore be employed as an initial dose warping evaluation workflow, while the procedure described herein could provide further insight and interpretation of the results under clinical scenarios.

6.1.4 Conclusions

This study proposed and demonstrated a workflow for the validation of DIR and dose-warping performance of any DIR algorithm in cases of mass-conserving deformations. Using this workflow with H&N patient images artificially deformed with clinically realistic deformations, the successful propagation of actual delivered dose to the reference anatomy following dose warping through CT-to-CT DIR, by OnQ rts was confirmed. Larger errors were occasionally observed, however, highlighting that DIR performance should always be evaluated and approved before proceeding to dose warping and accumulation in clinical setting.

Chapter 7

Plan Robustness to Anatomical Changes

This chapter investigates the robustness of several radiotherapy techniques to anatomical changes. Implementation of treatment monitoring workflows in routine clinical setting require considerable time and capital investments, while the potential treatment adaptations translate to requirements for further labour. Furthermore, certain anatomical changes may not necessarily result in clinically critical dosimetric changes, while the same anatomical change can have a different effect to individual treatment techniques. It is therefore important to perform prospective and retrospective studies to identify: (a) which treatment techniques are more sensitive to anatomical changes, and (b) which patient cohorts or anatomical changes can result in clinically significant dosimetric changes.

The first section of this chapter describes a retrospective study evaluating the robustness of three modern radiotherapy techniques for the treatment of nasopharyngeal carcinoma, to clinically realistic artificial anatomical changes. The treatment planning part of the study has been performed by Trevor Williams as part of an MSc dissertation. This section will, therefore, only concentrate on the plan robustness simulation and evaluation. The results of the planning study have been presented at ICHNO 2015 conference [98],

while, the preliminary results of the plan robustness study have been presented at the 3rd ESTRO Forum [99] and UKRO 2015 conferences.

The second section describes a similar retrospective study which evaluated the robustness of VMAT plans for the treatment of high risk prostate cancer, to several clinically observable anatomical changes.

7.1 Comparison of intensity modulated proton therapy versus rotational IMRT for nasopharyngeal carcinoma: robustness to anatomical changes

Studies to identify the robustness of treatment plans are important to inform physicians, physicists and radiotherapy technologist of the therapy techniques more sensitive to certain anatomical changes, as well as which anatomical changes are more likely to result in clinically critical dosimetric changes.

Previous studies in the literature investigated the effects of single anatomical change scenarios to one or two treatment techniques, employing online acquired images or repeat CT scans. Muller et al. [100] retrospectively investigated the effect of weight loss and tumour shrinkage to IMPT and IMRT plans of five H&N patients previously treated with TomoTherapy. New plans were created on the first online acquired MVCT of each patient and recalculated on weekly MVCT scans. Fukumitsu et al. [101] retrospectively examined the changes in delivered dose distributions due to changes in aeration of nasal cavities throughout proton beam therapy in patients treated for nasal cavity or para-nasal sinus cancers. The original plans were recalculated on repeat CT scans.

This study examines the dosimetric effects of various anatomical changes occasionally observed during the course of nasopharyngeal carcinoma (NPC) radiotherapy, to three modern treatment techniques: volumetric modulated arc therapy (VMAT), TomoTherapy, and intensity modulated proton therapy (IMPT). Clinically realistic artificial changes were retrospectively introduced to the original planning CT scans of ten NPC patients in order to avoid uncertainties due to dose calculations on online-acquired CBCT scans and isolate the dosimetric changes observed to the introduced anatomical change only.

7.1.1 Methods

Patient cohort and treatment planning

Ten nasopharingeal carcinoma (NPC) patients previously treated at our institution (Queen Elizabeth Hospital, Birmingham) were retrospectively selected. For each case a new VMAT, TomoTherapy and IMPT plan was created. All patients were planned to receive 65, 60 and 54 Gray equivalents (GyE) in 30 fractions to the PTVs of primary tumour (PTV65), intermediate risk (PTV60) and low risk (PTV54) nodal regions, respectively. Relative biological effectiveness (RBE) was assumed to be 1.0 for photon and 1.1 for proton beams.

- ***VMAT:*** VMAT plans were created using Monaco treatment planning system assuming a 6MV photon delivery on an Elekta Precise Treatment System^{TM} equipped with MLCi2 multi-leaf collimator and employing 3 dynamic arcs.

- ***TomoTherapy:*** TomoTherapy plans were designed on a version 5 planning station for a 6MV delivery on a TomoHD^{TM} machine.

- ***IMPT:*** IMPT was planned on an XiO system employing a spot-scanning technique, 3 beams, 0.4 cm pencil beam size and 0.6 cm spot spacing.

Simulation of clinically realistic anatomical changes

Three anatomical changes previously observed in patients with similar disease were applied on each patient's planning CT image using ImSimQA simulation software:

- ***Partial nasal cavity filling,*** was simulated by filling 50% of the empty volume, as shown in Fig. 7.1*a,b*. The HU used was the mean HU for sinus observed in the patient cohort employed in this study.

- ***Weight loss,*** was simulated by shrinking the body contour by 5 mm in the neck area. The 5 mm thick region between the boundaries of the two contours −green region in Fig. 7.1*c*− was then filled with HU = −1000, to represent air.

- ***Upward shoulder movement,*** moving the shoulders by 10-15 mm, as shown in Fig. 5.1*b*.

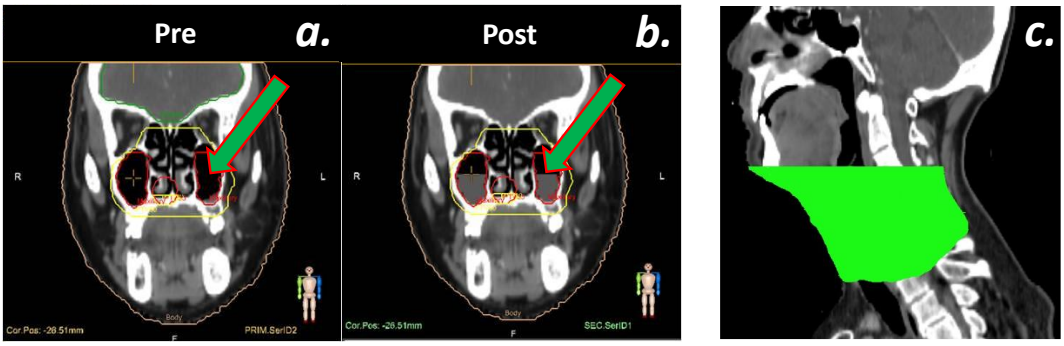

Figure 7.1: Artificial changes introduced for plan robustness evaluation. ***a.*** Pre and ***b.*** post partial nasal cavity filling simulation. ***c.*** Area of 5 mm shrinkage of body contour for simplified weight loss simulation.

Dose calculations and analysis

The artificially modified images were sent back to the three treatment planning systems, where plans were re-calculated retaining identical conditions to the original plans (i.e. beam arrangement, isocentre, segment positions, monitor units).

The original plans from the three treatment techniques were firstly compared in terms of PTV coverage, mean and maximum dose to normal structures. Then, the dosimetric effect of each artificial anatomical change has been estimated for each treatment technique. This analysis was based on the comparison of clinically important DVH metrics: minimum dose to 98%, 95% and 50% of the volume (D98%, D95% and D50%, respectively), maximum dose to 5% and 2% of the volume (D5% and D2%, respectively) and mean dose (D_{mean}).

To estimate the potential clinical relevance of the observed differences in dose distributions following the applied anatomical changes, radiobiological analysis has been performed using the BioSuite software [102]. The LKB model (Equation 1.1, p. 4) was employed for NTCP and the Poisson model (Equation 1.4, p. 5) for TCP. NTCP was calculated using the parameters and endpoints listed in Table 7.1, for the parotid glands [103], larynx [104] and spinal cord [105]. TCP was calculated under the following assumptions: $\alpha = 0.4$ Gy^{-1}, $\sigma = 0.009$ Gy^{-1}, $\alpha/\beta = 10$ Gy, homogeneous clonogenic cell density $= 10^7$ cells cm^{-3} [106].

Table 7.1: Endpoints and parameters used for normal tissue complication probability (NTCP) modelling.

Organ	Endpoint	m	n	TD_{50} [Gy]	α/β [Gy]
Larynx [104]	Grade 2 edema within 15 months	0.160	0.45	46.3	3
Parotid Glands [103]	Salivary flow 25% of pre-treatment flow after 1 year	0.180	1.00	28.4	3
Spinal Cord [105]	Myelitis necrosis	0.175	0.05	66.5	3

7.1.2 Results

Evaluation and comparison of original plans

Clinically acceptable plans were produced with all three modalities, achieving dose constraints for all target regions and critical organs at risk. IMPT revealed significantly better PTV conformity (Wilcoxon test, $P = 0.008$) and achieved significant reductions in mean dose to certain structures including the larynx ($P = 0.002$), brain ($P = 0.006$) and oral cavity ($P = 0.004$) [99]. No statistically significant differences were observed between VMAT and TomoTherapy plans.

Figure 7.2 shows a representative example of the three generated plans for the same patient. One of the treatment planning aims was the preservation of a low dose 'central corridor' for sparing of the spinal cord. In the coronal and transverse plane images, it is observed that IMPT plans achieved very low dose to this region while both photon plans (VMAT, TomoTherapy) revealed similar results with a higher dose bath. Similarly, greater dose sparing of oral mucosa was achieved by IMPT plans.

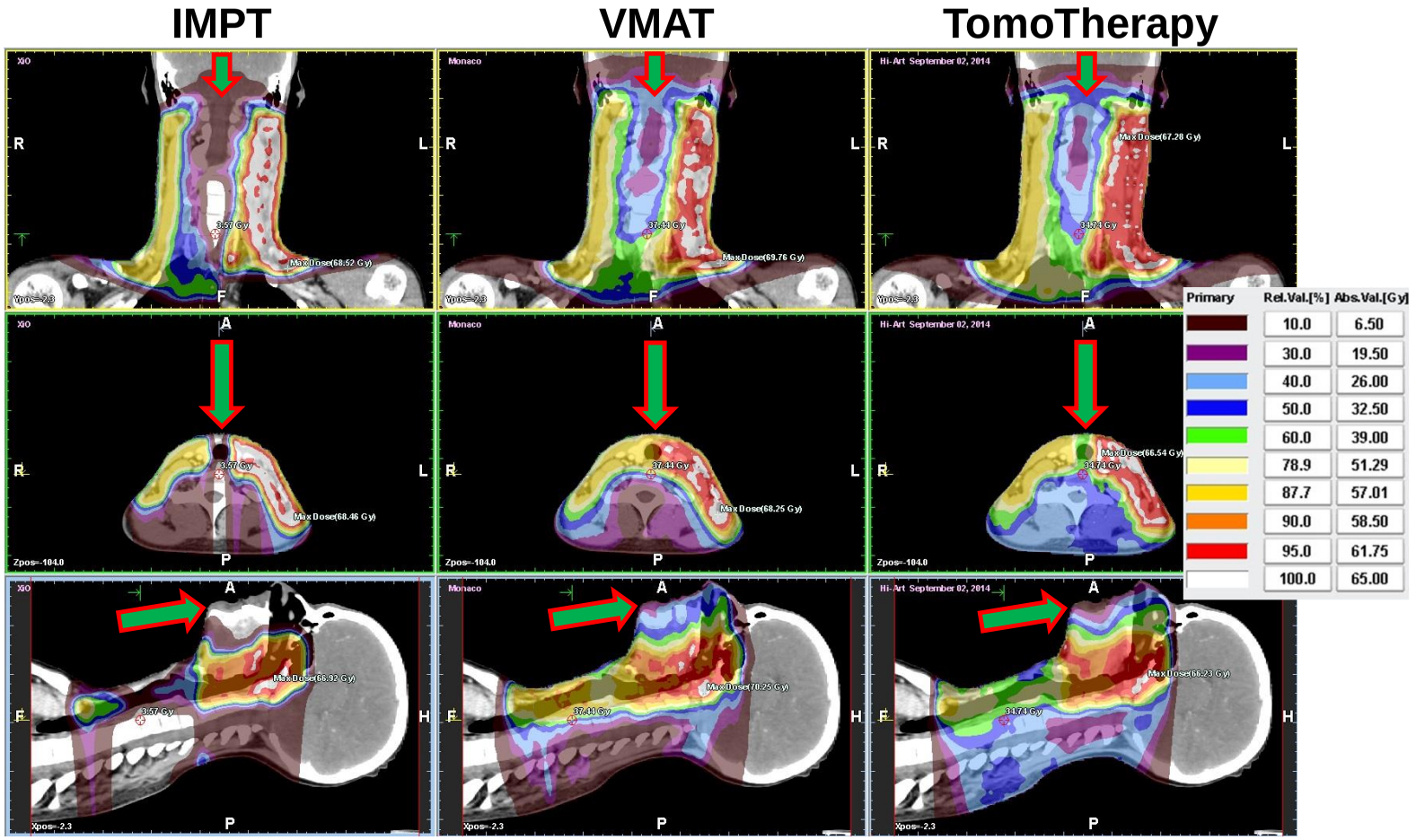

Figure 7.2: Examples of planned dose distributions achieved using the three radiotherapy modalities (IMPT: intensity modulates proton therapy; VMAT: volumetric modulated arc therapy; and TomoTherapy) for the same patient. The arrows point out to regions of interest were sparing was aimed during planning.

Robustness to anatomical changes

Partial nasal cavity filling simulation

Figure 7.3 shows a representative example of the effect of nasal cavity filling simulation for the three generated plans. Substantial alteration of the dose distribution in the transverse plane slice where nasal cavity was artificially filled by sinus is observed for the IMPT plan, with smaller changes exposed for both photon plans (VMAT and TomoTherapy).

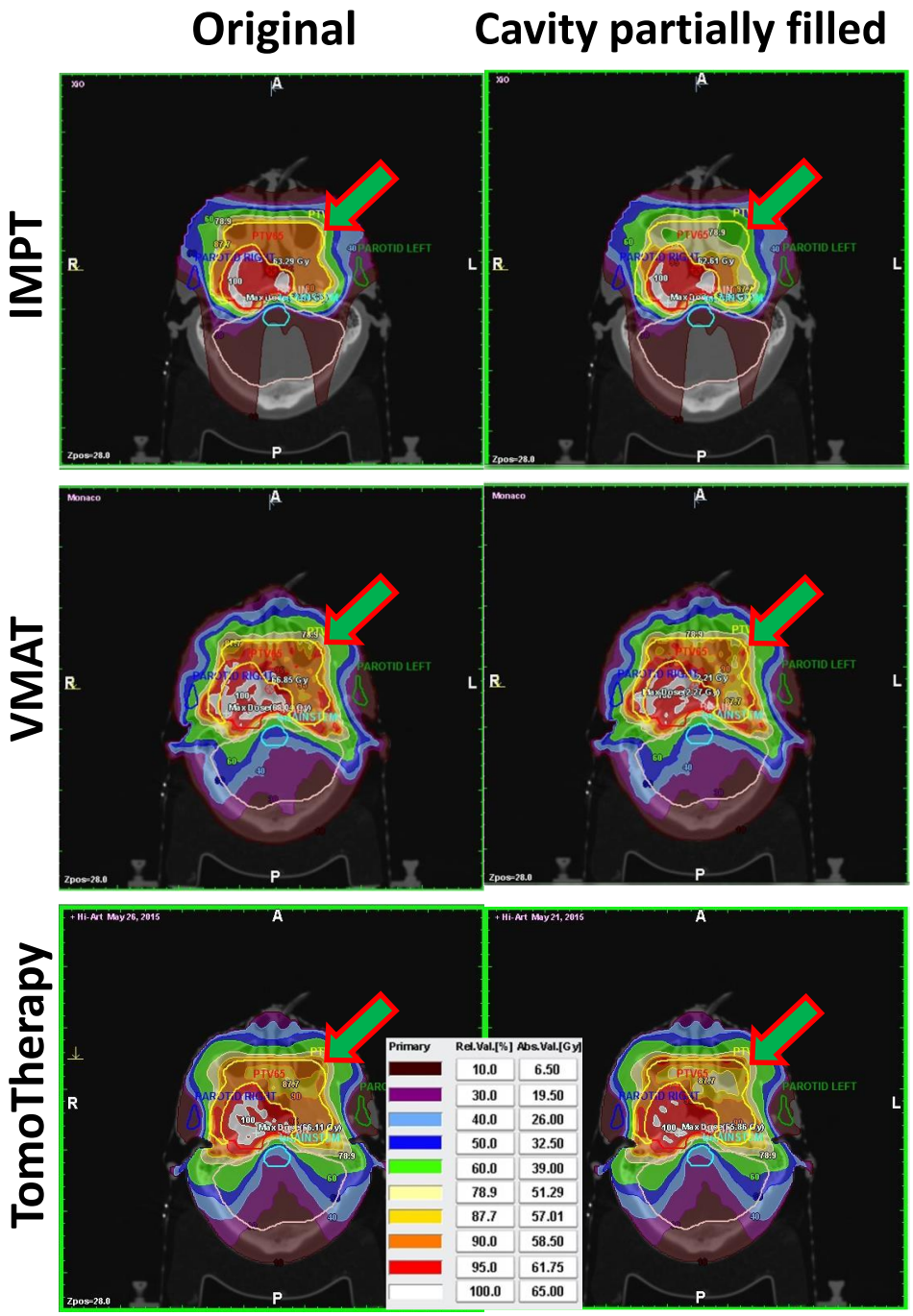

Figure 7.3: Effect of partial nasal cavity filling simulation to the three generated plans. On the left, transverse plane images show the original plans with the arrows showing the dose distribution within the PTV60 (yellow contour). On the right, the change of dose distribution following nasal cavity filling is shown.

PTV60: planning target volume with a prescription of 60 Gy;

IMPT: intensity modulated proton therapy;

VMAT: volumetric modulated arc therapy.

The mean absolute differences in important DVH metrics of the target volumes are shown in Fig. 7.4. It is observed that dosimetric changes for IMPT plans were larger in the higher dose target regions (CTV65, CTV60, PTV65, PTV60) as, due to the nature of the disease in the patient cohort studied, these regions were immediately behind the nasal cavities. Both photon plans revealed smaller differences, with TomoTherapy plans exposing slightly larger differences in median dose, as compared to IMRT. This may be attributed to the helical nature of TomoTherapy delivery.

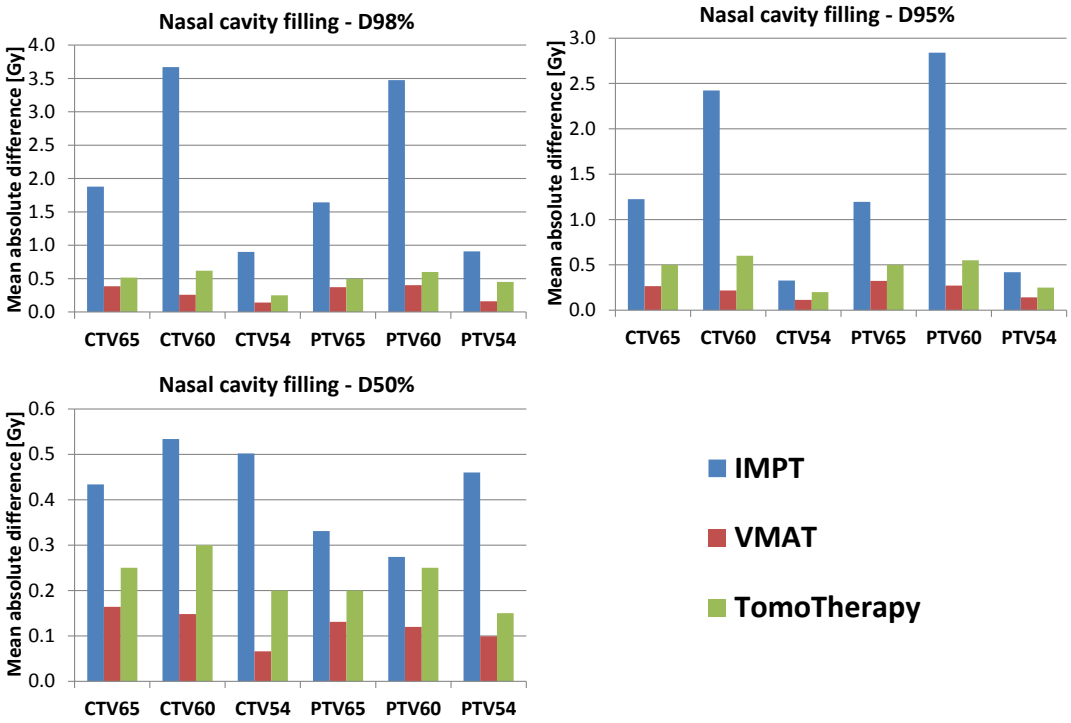

Figure 7.4: Mean absolute difference between original plan and dose distribution following partial nasal cavity filling simulation, in minimum dose received by 98%, 95% and 50% (D98%, D95%, D50%, respectively) of clinical target volumes (CTV) and planning target volumes (PTV), for the three plans generated.
IMPT: intensity modulated proton therapy;
VMAT: volumetric modulated arc therapy.

Upward shoulder movement simulation

Dosimetric differences following upward shoulder movement simulation are shown in Fig. 7.5. This comparison is concentrated in medium and low risk nodal regions (CTV60, CTV54, PTV60, PTV54) which, in the patient cohort used in this study, were positioned near the shoulders region. IMPT plans revealed substantial changes in the minimum dose to target regions, while both photon plans experienced smaller changes. TomoTherapy plans exposed larger differences in median dose, with similarly large differences for VMAT plans is some occasions. This may be attributed to the dynamic delivery of these plans, leading to more beamlets being affected by this simulated shoulder movement.

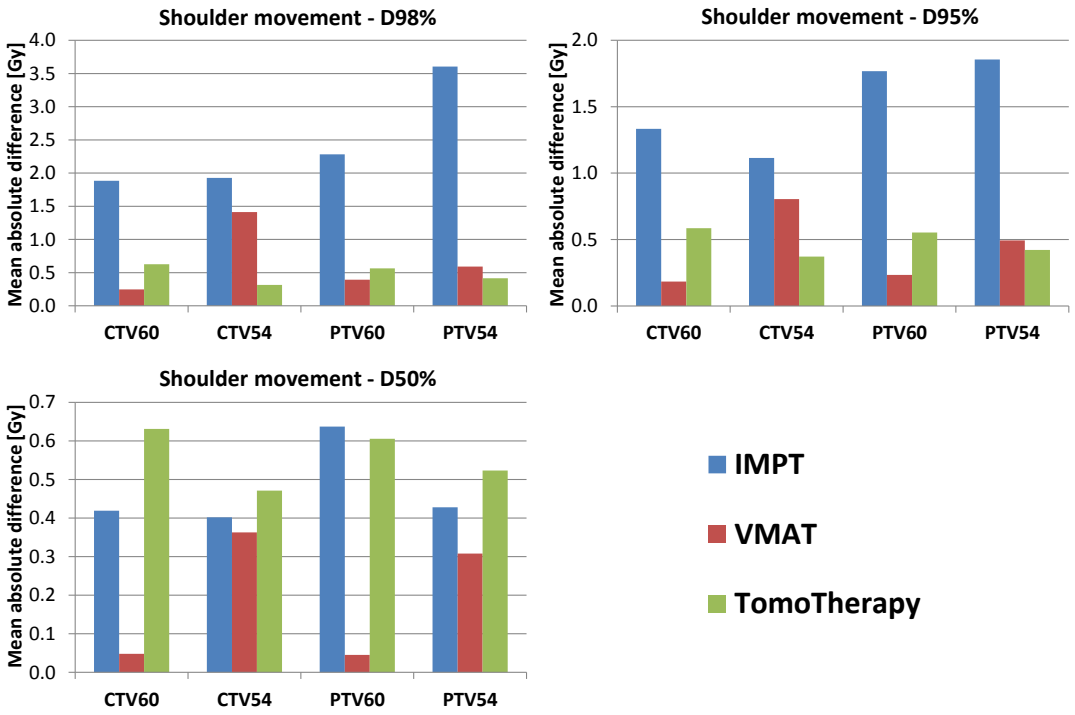

Figure 7.5: Mean absolute difference between original plan and dose distribution following upward shoulder movement simulation, in minimum dose received by 98%, 95% and 50% (D98%, D95%, D50%, respectively) of clinical target volumes (CTV) and planning target volumes (PTV), for the three plans generated.
IMPT: intensity modulated proton therapy;
VMAT: volumetric modulated arc therapy.

Weight loss simulation

Figure 7.6 shows the dosimetric consequences of the simplified weight loss simulation in target regions and organs at risk. Again, substantial changes were exposed in IMPT plans with generally considerably lower differences for both photon plans. Of special note is the large difference in median (D50%) and high dose (D2%) to the larynx for IMPT plans. VMAT plans revealed higher differences in minimum dose (D98%) to target regions, compared to TomoTherapy plans, caused by higher proportion of dose delivered through the front part of the neck region.

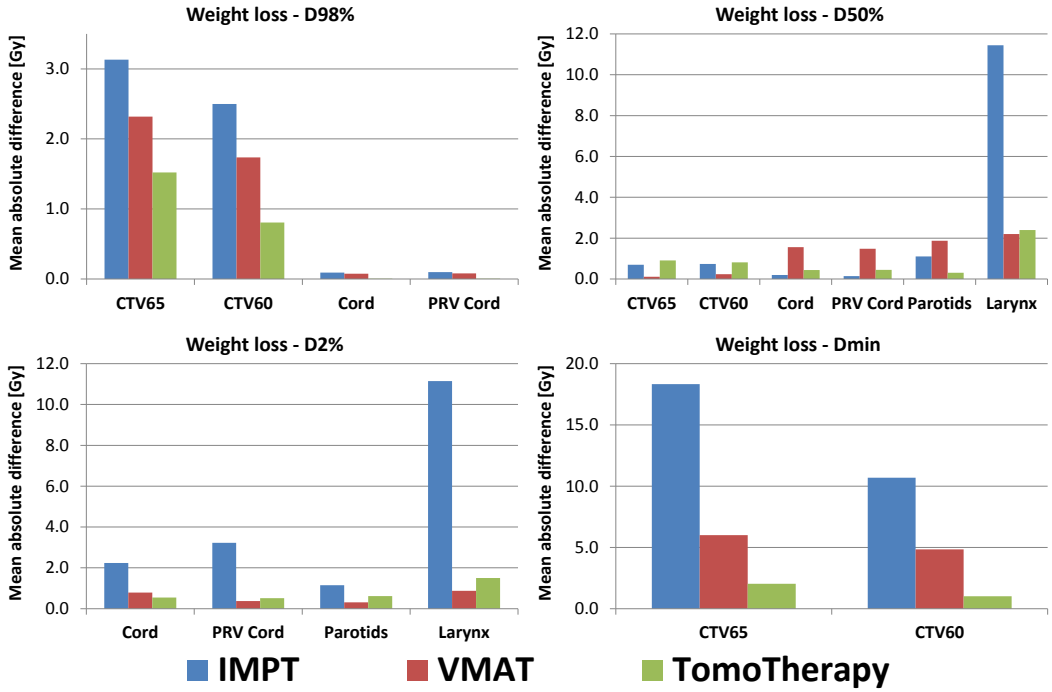

Figure 7.6: Mean absolute difference between original plan and dose distribution following weight loss simulation, in minimum dose received by 98%, 50% (D98%, D50%, respectively); maximum dose received by at least 2% (D2%); and minimum dose (D_{min}) of regions-of-interest.
CTV: clinical target volume;
PTV: planning target volume;
PRV: planning organ-at-risk volume;
IMPT: intensity modulated proton therapy;
VMAT: volumetric modulated arc therapy.

Table 7.2 summarises the mean TCP calculations, throughout the 10 patients employed in this study, comparing TCP calculation for the original treatment plans against the TCP following partial nasal cavity filling and weight loss simulations. It should be emphasised that TCP was calculated for the primary target volume only (i.e. CTV65) as the assumptions in the underlying model presume, while it was not calculated following the shoulder movement simulation scenario as the main dosimetric changes would be away from the main tumour site. These calculations reveal important changes to the TCP for IMPT treatments under both nasal cavity filling and weight loss scenarios, with lower and similar changes for VMAT and TomoTherapy plans.

Table 7.2: Mean tumour control probability (TCP) −as calculated for the primary clinical target volume (CTV65)− before (pre) and following (post) partial nasal cavity filling (D1) and weight loss (D3) simulation, for the three treatment plans generated.

	Mean TCP pre [%]			Mean TCP post [%]			Mean TCP change [%]		
	IMPT	VMAT	Tomo	IMPT	VMAT	Tomo	IMPT	VMAT	Tomo
D1	90.0	89.1	89.5	82.5	88.7	88.7	7.5	0.3	0.8
D3				86.7	87.2	88.4	3.1	1.8	1.1

IMPT: Intensity modulated proton therapy

VMAT: Volumetric modulated arc therapy

Tomo: TomoTherapy

Table 7.3 lists the NTCP values as calculated for the spinal cord, parotid glands and larynx in the original plans and following weight loss simulation. Due to the position of the other two artificial changes (i.e. nasal cavities and shoulder region), only minor dosimetric changes resulting in minimal NTCP changes were observed and are therefore not listed for clarity. NTCP modelling revealed that original IMPT plans were less likely

to induce late effects to the parotid glands or the larynx, as compared to both photon plans. However, a substantial change in probability of inducing grade 2 edema within 15 months after treatment following weight loss was exposed for IMPT plans.

Table 7.3: Mean normal tissue complication probability (NTCP) of organs-at-risk before (pre) and following (post) weight loss simulation, for the three treatment plans generated.

	Mean NTCP pre [%]			Mean NTCP post [%]			Mean NTCP change [%]		
	IMPT	VMAT	Tomo	IMPT	VMAT	Tomo	IMPT	VMAT	Tomo
Cord	0.0	0.1	0.1	0.0	0.2	0.1	0.0	0.1	0.0
Parotids	37.6	45.4	42.2	38.6	46.7	43.0	1.0	1.3	0.8
Larynx	5.1	9.0	9.6	32.6	12.5	12.8	27.5	3.5	3.2

IMPT: Intensity modulated proton therapy

VMAT: Volumetric modulated arc therapy

Tomo: TomoTherapy

7.1.3 Discussion

This study investigated the robustness of three modern radiotherapy techniques for nasopharyngeal carcinoma, namely spot-scanning IMPT, VMAT and TomoTherapy, under three clinically realistic artificial anatomical changes. Even though all techniques achieved high quality plans, IMPT achieved significantly better target conformity and sparing of surrounding healthy tissue. Under the investigated anatomical change scenarios, IMPT exposed important discrepancies to the intended dose distributions. Photon plans were generally proven more robust, and especially to the nasal cavity filling scenario.

The results of this study were largely as expected, considering the underlying physical differences between the examined treatment techniques. Taking advantage of the Bragg peak, proton plans were able to deposit the required dose to the target region while greatly sparing surrounding normal tissue. This distinct advantage compared to photon treatments, however, is also what makes these plans more vulnerable to anatomical changes. Photon beam plans were more robust not only due to the physical principles of their percentage depth dose curve formation, but also due to the fact that VMAT and TomoTherapy treatment are dynamically delivered from many (or all) angles around the body, making them less sensitive to localised anatomical changes.

The three introduced artificial anatomical changes were simplified simulations of actual situations observed during image guidance at Queen Elizabeth Hospital, Birmingham. Partial nasal cavity filling was simulated by 'filling' 50% of the empty volume inferiorly to superiorly. Under the assumption that the present sinus has very high viscosity, its position would not substantially change during the course of the treatment. If, on the other hand, the viscosity is low the same simulation would need to be represented by filling all slices by 50%. This would result in a smaller dosimetric change but would affect twice as many slices. The weight loss simulation ignores any movement or volumetric change of internal organs, such as the parotids which have been observed to shrink and move inwards and downwards during the course of treatment [97]. Such a simulation was attempted but the results were considered poor and unrealistic utilising the available software tools, therefore the simplified weight loss with the removal of a thin slap of tissue was employed.

It could be argued that plan quality is highly dependent on the departments and especially the planner's experience with the planning system, delivery hardware and underlying technique. This is partly true for the plans involved in this study, as the planner has no clinical experience in proton treatments while being more experienced in TomoTherapy

than VMAT planning. Provided that the same plan quality, robustness and dosimetric constraint considerations have been employed for the generation of all plans, the impact on robustness to anatomical changes would be diluted.

7.1.4 Conclusions

The robustness of IMPT, VMAT and TomoTherapy treatment plans for nasopharyngeal carcinoma treatment to clinically realistic anatomical changes was investigated. High quality plans were achieved with all three techniques. Spot-scanning IMPT revealed significantly better target conformity while significantly reducing dose to peripheral structures, however, it was severely affected by anatomical changes with possible dramatic changes to delivered dose distributions and clinical outcomes. On the other hand, both rotational IMRT techniques exposed similar differences between them and proved considerably more robust than proton treatments under the investigated scenarios.

The results suggest that regular monitoring of proton treatment is required in order to observe any potential anatomical changes and adapt the treatment to retain the original dosimetric targets. In the case of both photon treatments, nasal cavity filling variations are less likely to produce clinically critical dosimetric changes, while planning strategies to refrain beams from entering through or just above the shoulders can minimise the risk of suboptimal treatments due to shoulder movements. Weight loss, and especially considering the parotid gland movements and tumour shrinkage typically concurrently observed in clinical setting may result in sub-optimum photon treatments.

7.2 Volumetric modulated arc therapy for prostate cancer: robustness to anatomical changes

A similar worklfow was performed for the evaluation of prostate cancer VMAT plan robustness to certain clinically observable anatomical changes.

Previous studies investigated the dosimetric changes during radiotherapy for prostate cancer caused by inter-fraction anatomical changes. Godley et al. [53] used daily in-room CT scans (CT-on-rails) on which the daily dose was directly recalculated using the original plan, to assess the effects of anatomical changes to IMRT treatments. Thor et al. [107] utilised an average of 9 offline repeat CT scans in order to estimate the 'motion-inclusive' actual delivered dose to bladder and rectum of 38 prostate cancer patients treated with IMRT. The original planned and motion-inclusive dose distributions were then retrospectively associated with late effect outcomes. Varadhan et al. [108] employed online CBCT and MVCT scans to directly re-calculate the daily delivered dose distributions and assess dosimetric effects of anatomical change to IMRT and TomoTherapy plans.

This study examines the dosimetric effects of various clinically observable anatomical changes to prostate cancer patients treated with VMAT. This is achieved by employing realistic artificial deformations, thus eliminating uncertainties involved during dose calculations on on-line acquired scans while enabling the isolation and association of observed dosimetric changes to individual and controlled anatomical changes.

7.2.1 Methods

Patient cohort and treatment planning

Five high risk prostate cancer patients previously treated at our institution were randomly retrospectively selected and new single arc VMAT plans were created on RayStation TPS, assuming a delivery on an Elekta Versa HD™ with an Agility™ multi-leaf collimator and flattening filter free mode. The prescribed dose was 74 Gy in 37 fractions with PTV margins and dose constraints as listed in Table 7.4.

Table 7.4: Planning target volume (PTV) margins and dose volume constraints employed for high risk prostate cancer treatment planning.

	Margins	Constraints [Gy]
PTV1	Prostate & SV + 10 mm margin in all directions	min>56.2
PTV2	Prostate & SV + 10 mm margin except 5 mm towards rectum	min>67.3
PTV3	Prostate only + 5 mm margin except 0 mm towards rectum	min>70.3 73.3<median<74.7

SV: seminal vesicles

min: minimum

Simulation of clinically realistic anatomical changes

Three anatomical changes previously observed in patients with similar disease were applied on each patient's planning CT scan, using ImSimQA simulation software (as previously outlined in Section 4.1.1, p.56):

- ***Rectal expansion,*** was simulated by introducing an artificial rectal gas pocket and radially expanding the rectum by 5-10 mm.
- ***Bladder expansion,*** was simulated by 10-15 mm superior expansion and 2-5 mm expansion towards the prostate.
- ***Weight loss,*** was simulated by shrinking the body contour by 5 mm and filling the region between the original and shrunk body contours with HU $= -1000$, to represent air.

Dose calculations and analysis

The artificially modified images were sent back to RayStation where the original plans were re-calculated, with identical conditions (i.e. beam arrangement, isocentre, segment positions, monitor units). This process simulated the clinical image guidance protocol performed at our institution, whereby daily positioning is based on soft tissue matching (i.e. the prostate). Therefore, the prostate (target) will always be at the isocentre in the presence of small volumetric changes of bladder or rectum. The original and new dose distributions were then compared to expose the effect of each of the introduced anatomical changes. As two of the artificial changes involved the expansion of rectum and bladder, direct comparison of DVHs for these organs would be misleading. Similarly, dose-volume constraints and radiobiological analysis using these values would be vague. The actual volume of each organ receiving a specific, clinically important amount of dose was therefore evaluated.

7.2.2 Results

Rectal expansion simulation

Figure 7.7 shows the mean volume of rectum receiving high dose before and after the simulation of rectal expansion. Considerable differences are observed, with the mean rectal volume receiving 74 Gy increasing by almost a factor of 3 while the volume receiving 70 Gy increasing by 30%. Considering the QUANTEC recommendations [109], which suggest that particular effort should be made to minimise these volumes as they can significantly impact the complication probability, these results may be translated as clinically critical.

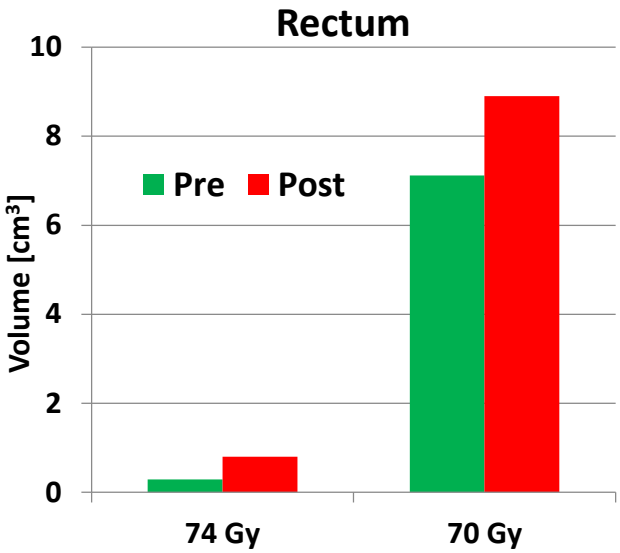

Figure 7.7: Rectal volume (mean across 5 cases) receiving a dose of 74 and 70 Gy before (Pre) and after (Post) rectal expansion simulation.

Bladder expansion simulation

Figure 7.8 shows the change in mean bladder volume receiving a dose of 74, 60 and 50 Gy following bladder expansion simulation. In this case, the bladder volume receiving 74 Gy has not increased noticeably. However, the volumes receiving 60 Gy and 50 Gy increased by 13% and 36%, respectively.

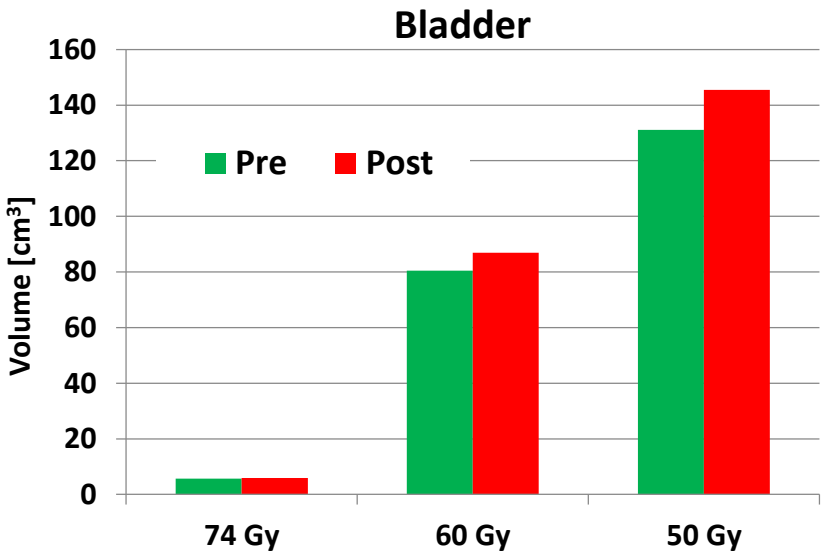

Figure 7.8: Bladder volume (mean across 5 cases) receiving a dose of 74, 60 and 50 Gy before (Pre) and after (Post) bladder expansion simulation.

Weight loss simulation

Figure 7.9 shows the mean dosimetric changes observed in rectum, bladder and PTV3 following weight loss simulation. Of note is the potential of hot and cold spot formation within the target region in this scenario, even if daily image guidance is based on soft tissue matching.

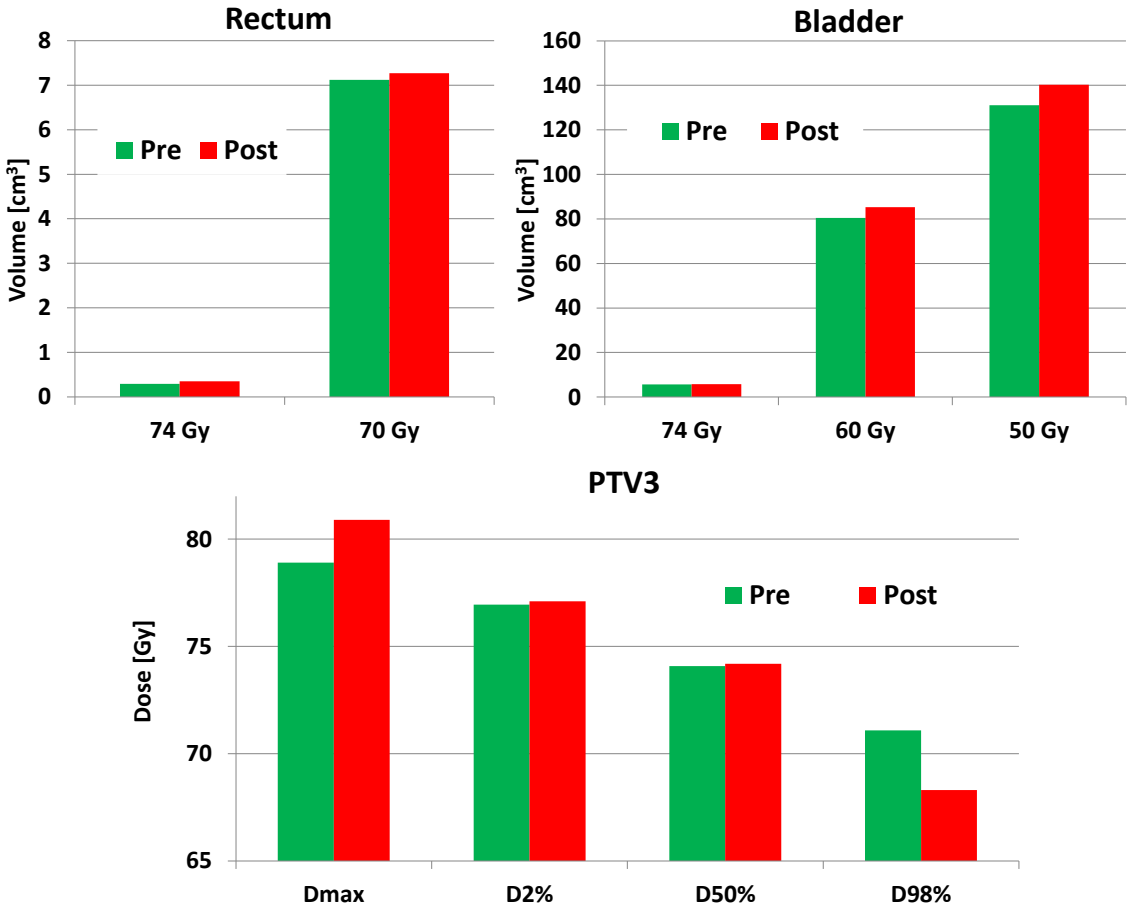

Figure 7.9: Rectal and bladder volume receiving a dose of 74, 70, 60 or 50 Gy; maximum dose (D_{max}), maximum dose received by at least 2% (D2%), minimum dose received by 50%, 98% (D50%, D98%, respectively) of the primary planning target volume (PTV3) before (Pre) and after (Post) weight loss simulation (mean across 5 cases).

7.2.3 Discussion

This study investigated the robustness of VMAT plans for prostate cancer under three clinically realistic artificial anatomical changes. Considerable dosimetric changes were exposed following the three simulations which may translate in alterations to the otherwise expected critical clinical outcomes.

A limitation of this study is the simplicity of anatomical changes introduced. As observed by Crevoisier et al. [110], bladder and especially rectal volume changes can cause different displacement of the prostate and seminal vesicles as large as 8.4 mm and 15.6, respectively. Using the available tools, these changes could not be simulated. Furthermore, combinations of the individual anatomical changes applied in this study may be observed in clinical situations, which can lead to more sever dosimetric changes to those observed in this analysis.

7.2.4 Conclusions

The robustness of prostate cancer VMAT plans to clinically observable anatomical changes was assessed. Rectal and bladder expansion revealed important dosimetric changes to the respective organs at risk, while small changes were observed following weight loss simulation. Important changes to the minimum dose received by the primary PTV following weight loss simulation was exposed.

Chapter 8

Adaptive Radiotherapy Workflows

This chapter outlines the evaluation of several offline treatment adaptation workflows in prostate cancer patient cases receiving VMAT treatments. The outcomes of the work described in previous chapters are combined in order to perform a complete offline ART workflow.

8.1 Adaptive radiotherapy techniques in prostate cancer treatments

Current IMRT and VMAT techniques in prostate cancer treatments achieve highly conformal plans to the target region with steep dose gradients to minimize dose to neighbouring healthy tissues. Inter-fraction target (prostate or prostate and seminal vesicles) displacements relative to the treatment beam and deformation of bladder and rectum [111; 112] may, therefore, result in a sub-optimal treatment or incorrect assumptions about the dose

being received. On-line image guidance, especially the use volumetric imaging which provides sufficient soft tissue contrast, was proven efficient in ensuring target coverage and normal tissue sparing [113]. However, some anatomical changes can not be compensated with simple patient positioning adjustments following on-line image guidance.

Adaptive radiotherapy techniques have been investigated in an attempt to compensate for these inter-fraction anatomical changes. Yan et al. [114] used on-line portal images and off-line CT scans acquired during the first 4 days of treatment, calculated the systematic and random set-up errors, target motion and deformation for each patient and created a new patient-specific PTV which was subsequently used for treatment adaptation. On the other hand, both prostate and rectum variations were monitored and considered during treatment plan adaptation in two further studies, using off-line repeat CT scans [115] in one case and weekly CBCT scans in the other [116]. Liu et al. [66] evaluated treatment adaptation techniques by incorporating dose feedback based on dose calculated on repeat CT scans, when cumulative dose to the target exceeded the pre-defined tolerance.

This study aimed to investigate several treatment plan adaptation workflows in prostate cancer patient cases, in an attempt to identify the most efficient method for:

A. Dose accumulation method,

B. Dosimetric monitoring frequency,

C. Treatment plan adaptation method,

D. Treatment plan adaptation timing;

while aiming for resource viable and widely applicable implementations.

8.1.1 Methods

Patient cohort and treatment planning

Ten prostate cancer patients previously treated at our institution with daily CBCT-guided IMRT were retrospectively selected from a pool of cases that required a plan evaluation during the course of treatment due to a noticeable anatomical change (e.g. weight loss/gain, large deviations in bladder/rectum volume). The planning CT, structure sets and all daily CBCT scans were transferred to RayStation TPS where new single arc VMAT plans created assuming a delivery on an Elekta Versa HD™ with an Agility™ multi-leaf collimator, 6 MV beams and a prescribed dose of 74 Gy in 37 fractions to the primary PTV (and the dose-volume constraints listed in Table 7.4, p. 124).

A. Dose accumulation method for dosimetric treatment monitoring

Dosimetric treatment monitoring was performed by directly calculating the original plan on CBCT images, using the tissue specific HU-to-density curve generation method which revealed the highest accuracy during evaluation study outlined in Section 5.2, p. p. 84.

Dose calculated on each CBCT was then accumulated on the patient's planning CT following DIR between the two images, employing RayStation's hybrid DIR with the use of controlling ROIs (prostate and rectum), which was shown to be the most accurate when evaluated in Section 4.2 (p. 64). The same algorithm without the use of controlling ROIs was also used in order to quantify the dosimetric consequence of potential DIR accuracy differences. The accumulated dose distributions using the two DIR methods were then compared.

B. Dosimetric treatment monitoring frequency

Daily dosimetric treatment monitoring was compared against weekly monitoring. Weekly dose distributions were estimated by calculating the original plan on the first CBCT of

each week into treatment, giving it a 'weight' of 5 fractions, and then accumulating on the planning CT image following DIR (RayStation's Hybrid with controlling ROIs).

C. Treatment plan adaptation method

Two methods for off-line plan adaptation were investigated:

i. Simple re-plan

A new plan was created based on the new anatomy and the original dosimetric targets, ignoring the 'actual delivered dose to date' estimated by dosimetric treatment monitoring and dose accumulation.

ii. Re-optimisation (i.e. addressing cumulative dose to date)

The accumulated dose was taken into account and therefore the new plan aimed to address potential over-dose of organs at risk or under-dose of target regions.

The new plan was created on the original planning CT, modified to match the anatomy in the respective CBCT. The modification was performed by copying the outer body contour from the appropriate CBCT onto the planning CT and compensating for excess tissue by introducing HU = 0 (i.e. water) when the patient gained weight or HU = −1000 (i.e. air) when tissue was lost. The rest of the image and contours were kept unchanged. This was considered sufficient as it enabled the compensation of the gradual anatomical change of weight loss or gain. On the other hand, changes in bladder and rectal volume are random −if the patient does not fully comply with the preparation instructions− and therefore the original contours would suffice. In all cases, the original plan was calculated on all CBCTs up to the point of adaptation and the adapted plan calculated on the remaining CBCTs. Recalculated dose distributions were then accumulated on the planning CT images, following DIR, to allow for dosimetric comparison of the two workflows.

D. Treatment plan adaptation timing

Three treatment adaptation time points were simulated for all patients and for each of the two adaptation methodologies. These were either immediately after fraction 20, 25 or 30 (out of a total of 37 fractions). Fixed time points were chosen in order for the adaptation results to be directly comparable and to examine whether the investigated adaptation methodologies are affected by this choice. Immediately after fraction 20 was chosen as the first time point since it would be unlikely for a treatment adaptation to be dosimetrically justified earlier based solely on inter-fractional anatomical changes. On the other hand, the last time point was chosen to be after fraction 30 so as to allow for a few remaining fractions for potential dose compensation.

Analysis

3D gamma analysis with 1%/1mm passing criteria ($\gamma_{1\%/1mm}$) was used for the comparison of dose accumulation methodologies. This analysis was chosen in order to identify and localise small deviations between the two methods. Since previous results suggested that RayStation's Hybrid algorithm with the use of controlling ROIs showed the most accurate and consistent performance, it was here considered to 'gold-standard' to which the same algorithm without the use of controlling ROIs was compared to.

DVH evaluations were employed for qualitative investigation and comparison of dosimetric monitoring frequency, treatment plan adaptation method and timing workflows. Moreover, analysis included the comparison of clinically important DVH metrics: minimum dose to 98%, 95% and 50% of the volume (D98%, D95% and D50%, respectively), and maximum dose to 2% of the volume (D2%).

Biologically effective dose (BED) (Eq. 1.5, p. 5) metric was employed −as a qualitative measurement and as an attempt to give clinical context to the observed results− for the comparison of these methodologies for changes in dose to the rectum and, specifically,

the region intersecting with the PTV2 therefore receiving highest dose. The following assumptions were made for the calculation of BED in this region:

- Intersection between PTV2 and rectum ≈15% of rectal volume.

  Rectum and PTVs were delineated according to the CHHiP trial protocol (register number ISRCTN97182923) [117], i.e. outlined as solid organ by defining the outer wall starting from the anus to the recto-sigmoid junction, while PTV2 included the prostate −or prostate and seminal vescicles− with a 10 mm margin except 5 mm towards the rectum. This assumption was made in accordance to QUANTEC [109], which indicates that higher dose ($\geqslant$ 70 Gy) regions as the driver of rectal toxicity.

- $\alpha/\beta = 3$

  Value employed in QUANTEC [109].

- Dose delivered to this volume relatively uniform, approximated by median dose.

Furthermore, the resulting high dose regions following treatment adaptations were compared against the estimated rectal tolerance, assuming a tolerance of 70 Gy in 15% of the rectal volume (CHHiP trial). The BED of a rectal region uniformly irradiated with 70 Gy in 37 fractions was calculated and compared against the BED following each treatment adaptation method.

8.1.2 Results and Discussion

A. Dose accumulation method

Figure 8.1 shows an example of a gamma map ($\gamma_{1\%/1mm}$) comparing the accumulated dose distributions by dose warping following DIR by RayStation's hybrid DIR with and without controlling ROIs. It is observed that most of the 'failing' voxels are concentrated

in and around the rectum, area where the use of controlling ROIs showed enhanced registration accuracy. This result highlights the dosimetric uncertainties introduced by image registration uncertainties, previously observed and discussed in Chapter 4.

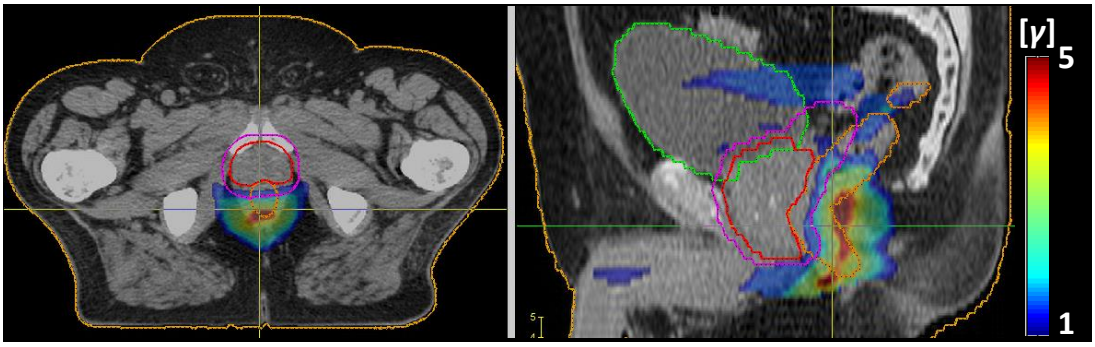

Figure 8.1: Three-dimensional gamma analysis (1%/1mm criteria) comparing accumulated dose distributions using RayStation's Hybrid deformable image registration algorithm with and without controlling regions-of-interest (prostate and rectum). [Green: bladder; orange: rectum; red: PTV3; pink: PTV1; PTV: planning target volume.]

The use of RayStation's hybrid algorithm without using controlling ROIs revealed important differences as compared to the same algorithm along with the use of controlling ROIs, especially in the rectum area. This is partly due to poor performance of the DIR algorithm in the rectal region —which has relatively low contrast compared to the surrounding tissues, while new features (e.g. rectal gas) may appear in certain daily images— but also due to the steep dose gradients present between the PTVs and the rectum whereby even small registration inaccuracies can lead to substantial dosimetric uncertainties.

The Hybrid DIR algorithm with controlling ROIs was therefore used for the remaining of the study.

B. Dosimetric monitoring frequency

Figure 8.2 shows DVHs for a single representative −median− case, comparing daily against weekly dose monitoring. Minor differences are observed in dose accumulation at the target (PTV3) while noticeable differences are exposed in bladder and especially rectum. This observation can be attributed to the systematic and random nature of volume and shape change experienced by these organs, the full extend of which can be best captured by daily monitoring.

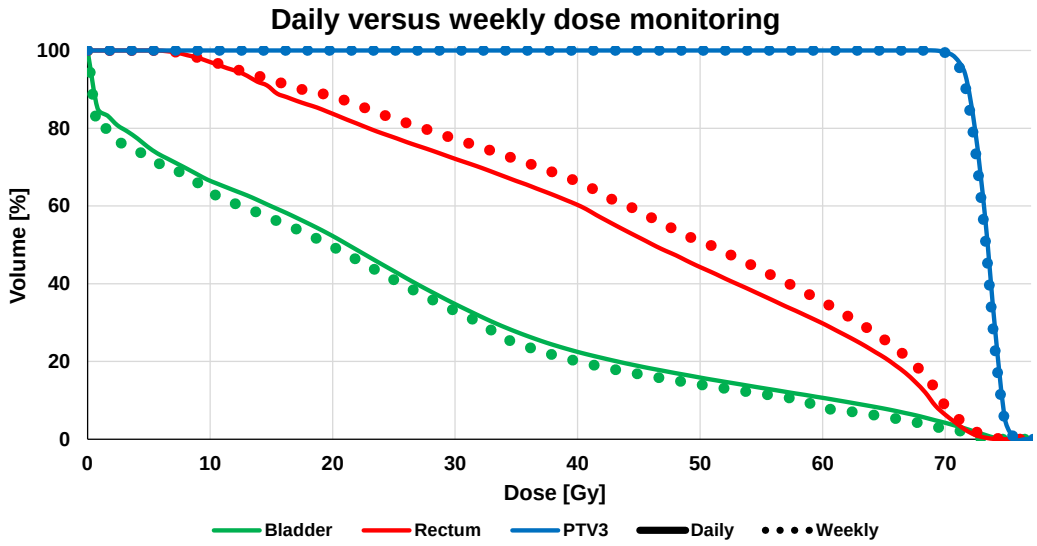

Figure 8.2: Dose volume histograms of a single representative patient case showing accumulated dose distributions following daily and weekly treatment monitoring. PTV: planning target volume.

Even though the example shown in Fig. 8.2 was a median case in terms of the extent of observed differences, this result was systematic throughout the test cases, highlighting the importance of daily dosimetric monitoring in anatomical regions were random changes are expected (e.g. pelvis).

Daily monitoring was therefore employed for the rest of the study.

C. Treatment plan adaptation method

Figure 8.3 shows DVHs for a single typical patient case (who experienced weight gain during radiotherapy treatment), comparing the original plan intention against the actual delivered dose up to fraction 25. A small under-dose of PTV3 is observed, along with a considerable under-dose of PTV2 and PTV1. Of note is the lower dose received by the rectum compared to the original treatment plan.

Figure 8.4 shows DVHs comparing the two adaptation methods against the original plan intention and the result of the original plan without adaptation. Dotted lines show the intention of each method for the remaining of the treatment and the intention of the original plan for the whole treatment, while straight lines show the actual accumulated dose at the end of the treatment for each case. As discussed earlier, simple re-plan method created a new plan with the original dosimetric aims and constraints, therefore, its intention looks similar to the original plan (green dotted line). Since it does not account for the cumulative dose to date, the PTV2 and PTV1 under-dose observed in Fig. 8.3 have only been restrained rather than compensated (green line). It is also observed that the final dose distribution following simple re-plan is slightly 'hotter' than the intended dose distribution in all PTVs, rectum and bladder, which can be attributed to anatomical changes after treatment adaptation. On the other hand, re-optimisation method aims to compensate for the under-dose observed in Fig. 8.3 by delivering additional dose to the regions within the PTVs which have been under-dosed; resulting in 'hotter', non-uniform dose distributions (red dotted line). It is observed that this task has been mostly achieved since the minimum dose to the PTVs matches or exceeds the original intention (red line). Regrettably though, due to the presence of further anatomical changes following the adaptation, the median and maximum dose to all PTVs and maximum dose to the rectum exceeded the original intentions.

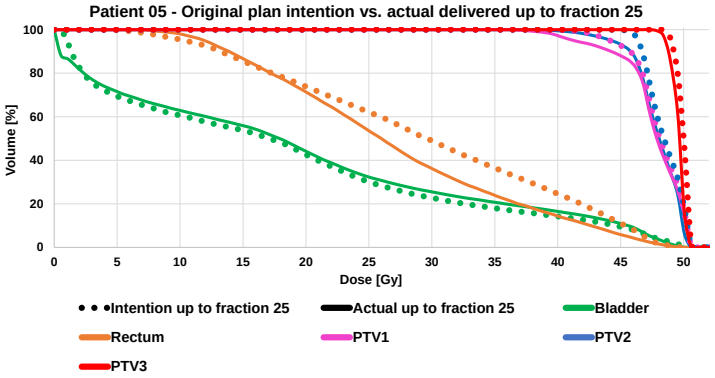

Figure 8.3: Dose volume histograms for a single patient, showing the original plan intention and actual delivered dose up to fraction 25.

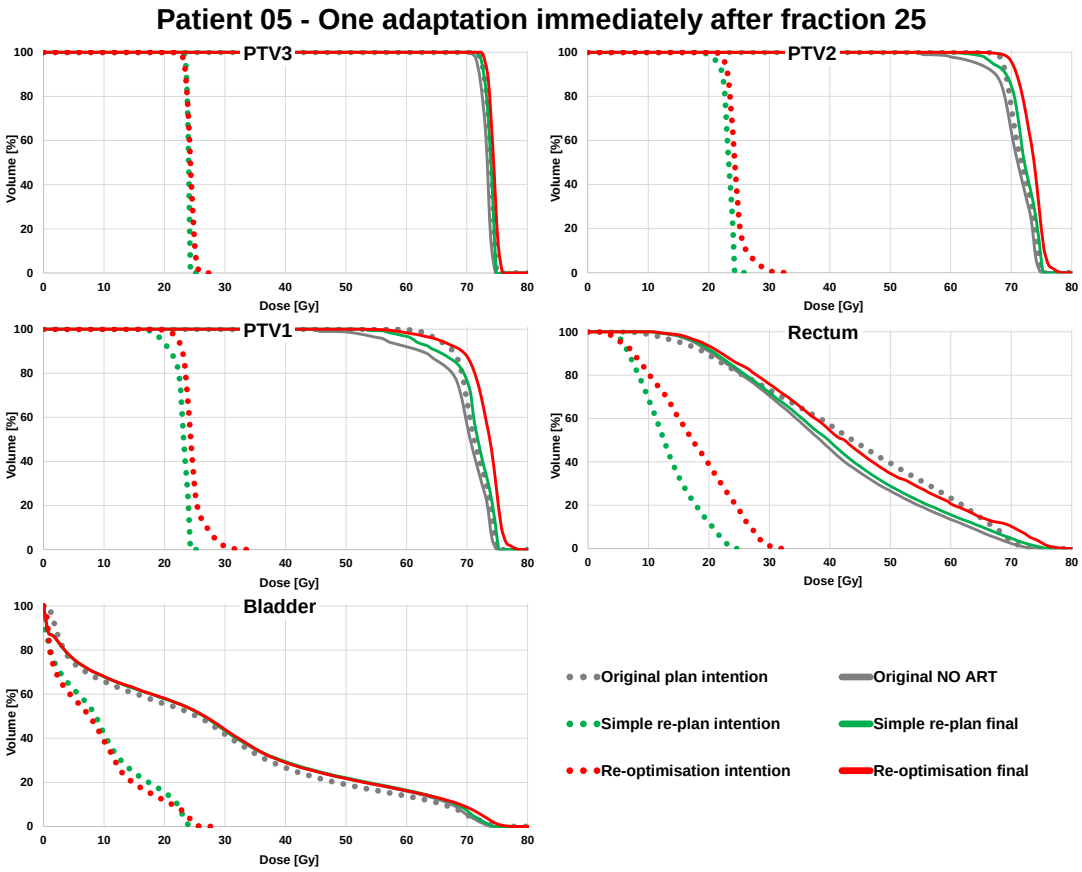

Figure 8.4: Dose volume histograms for a single patient comparing two adaptation methods immediately after fraction 25 (intention for remaining of treatment and actual accumulated dose at the end of treatment) against the original plan intention and the actual delivered dose without adaptation.

As observed in Fig. 8.4, the rectal volume receiving 70 Gy in the cumulative DVH following adaptation by re-optimisation (red straight line) is 9.8%, being lower than the assumed maximum tolerance of 15%. Figure 8.5 shows BED as calculated for the rectum/PTV2 intersection region in the same patient case, comparing the two adaptation methods immediately after fraction 25, against the estimated rectal tolerance. Even though the cumulative dose following adaptation by re-optimisation seemed well in tolerance through DVH analysis, BED calculations suggest that the region of rectum intersecting with PTV2 exceeded tolerance. This is caused due to the presence of hot-spots in this region during the last 12 fraction, following re-optimisation.

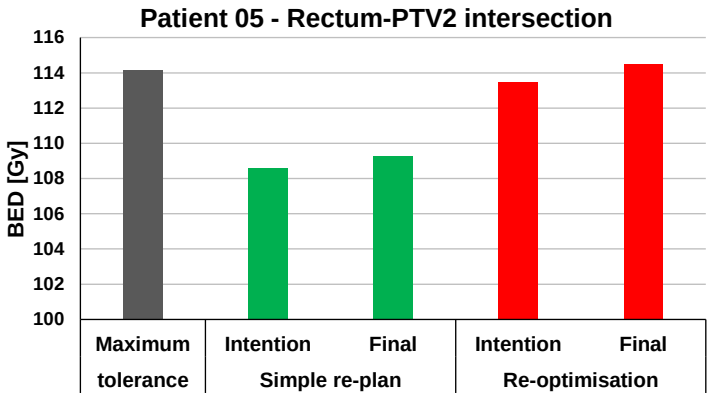

Figure 8.5: Biologically effective dose (BED) for a single patient, as calculated for the region of the rectum intersecting with planning target volume 2 (PTV2), comparing two adaptation methods immediately after fraction 25 against the estimated maximum rectal tolerance.

Figure 8.6 shows the dosimetric differences for another patient case, comparing the two adaptation methods and the result of the original plan without adaptation against the original plan intention. Substantial differences are observed in the two PTV regions between the original plan intention and its result without adaptation, while increased median dose to the rectum was also observed. Simple re-plan only managed to limit the PTV underdose and excess dose to the rectum. Re-optimisation achieved a compensation

for the underdose of the PTVs up to fraction 25 with only a small excess dose to the target regions, rectum and bladder.

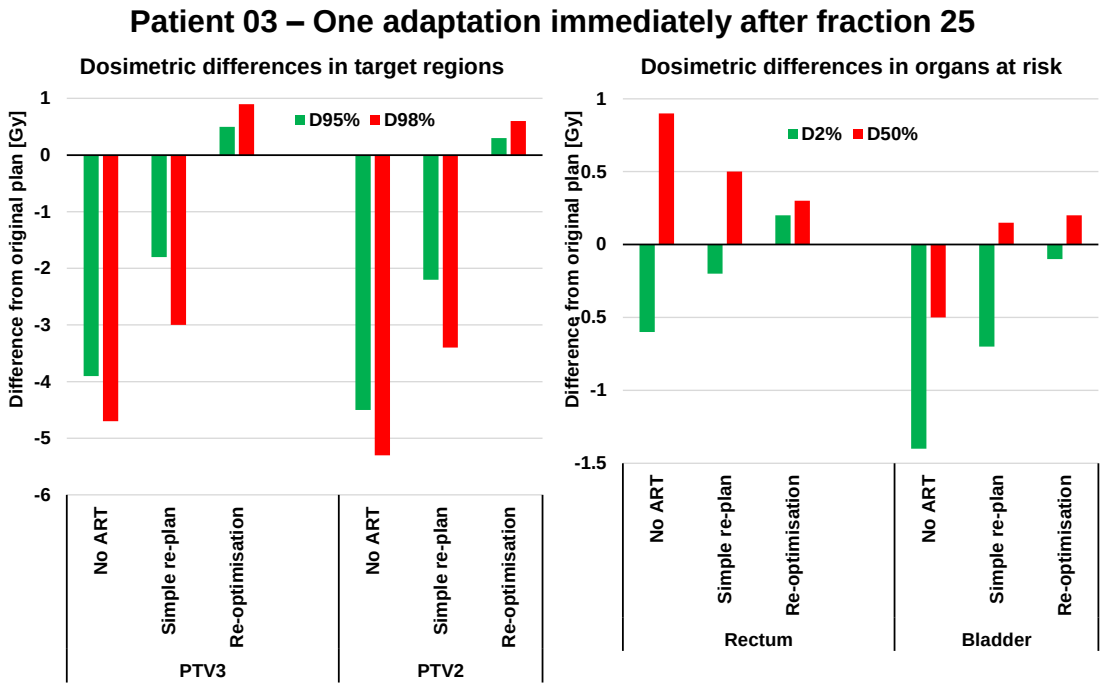

Figure 8.6: Dosimetric differences for a single patient comparing two adaptation methods immediately after fraction 25 and the actual delivered dose without adaptation against the original plan intention.
D95%: minimum dose received by at least 95% of volume;
D98%: minimum dose received by at least 98% of volume;
D2%: maximum dose received by at least 2% of volume;
D50%: maximum dose received by at least 50% of volume;
ART: adaptive radiotherapy;
PTV: planning target volume;

The 2 examples of results discussed here are representative of the 2 main observations throughout the test cases employed. Both examples demonstrate the inability of simple re-plan adaptation method to compensate for under-dose of target regions up to the point of plan adaptation, while, the dose-per-fraction and the cumulative dose to organs at risk remain similar to the original plan intentions. Therefore, provided that before treatment

adaptation the observed target underdose is small or the reason of adaptation is excess dose to organs at risk, simple re-plan method may result in the desired dose distribution.

Re-optimisation method achieved a compensation of PTV underdose in both examples, as intended. However, in the first case (Fig. 8.8) excess dose was delivered by the end of the treatment, in both the target regions and organs at risk. This was the result of further anatomical changes after plan adaptation. Moreover, the additional dose intended to be delivered by the adapted plan in an attempt to compensate for underdosage in target regions, may seem within tolerance when assessed using cumulative DVHs but could exceed tolerance when assessed radiobiologically (Fig. 8.5). On the other hand, the second example (Fig. 8.6) demonstrates that if minimal anatomical changes occur after plan adaptation, re-optimisation can compensate for underdosage of target regions while keeping within tolerance or even reducing dose to organs at risk.

To summarise adaptation methodology findings, for the 10 patient cases investigated in this study:

Simple re-plan

- Can not compensate for under/over dose.
- Does not introduce 'hot-spots' in OARs nor non-uniform dose distributions in target regions.
- Conventional planning and plan assessment protocols are applicable.

Re-optimisation

- Can compensate for under/over dose.
- May generate non-uniform dose distributions in target regions and 'hot-spots' in target regions or organs at risk.
- Further anatomical changes after adaptation may have a greater effect due to excess dose and non uniformity.
- Cumulative DVHs may not be adequate in generating and assessing the adapted plan. Radiobiological analysis may be required to estimate the effect of new dose distribution to target regions and organs at risk.

D. Treatment plan adaptation timing

Figure 8.7 shows DVHs for another typical patient case, comparing the original plan intention against the actual delivered dose up to fraction 20, followed by DVHs comparing the two treatment adaptation methods immediately after fraction 20 against the original plan intention and the actual delivered dose without adaptation, in certain regions of interest. Figures 8.8 and 8.9 show similar plots for the same patient case for adaptation immediately after fraction 25 and 30, respectively. Concentrating on the actual accumulated dose following simple re-plan (green straight line) it is observed that coverage of PTV2 is closer than the original intention when plan was adapted immediately after fraction 25, compared to when adapted after fraction 20 or 30. PTV3 coverage, on the other hand, is 'colder' when plan was adapted after fraction 30 compared to adaptations after fraction 20 and 25. Focusing on the re-optimisation intention (red dotted line), it is seen that as the plan adaptation is delayed this intended dose distribution gets 'hotter', especially in PTV2 and bladder. The actual accumulated dose following re-optimisation is closer to the original intention when re-optimisation was performed after fraction 30, compared to adpatations performed after fraction 20 and 25.

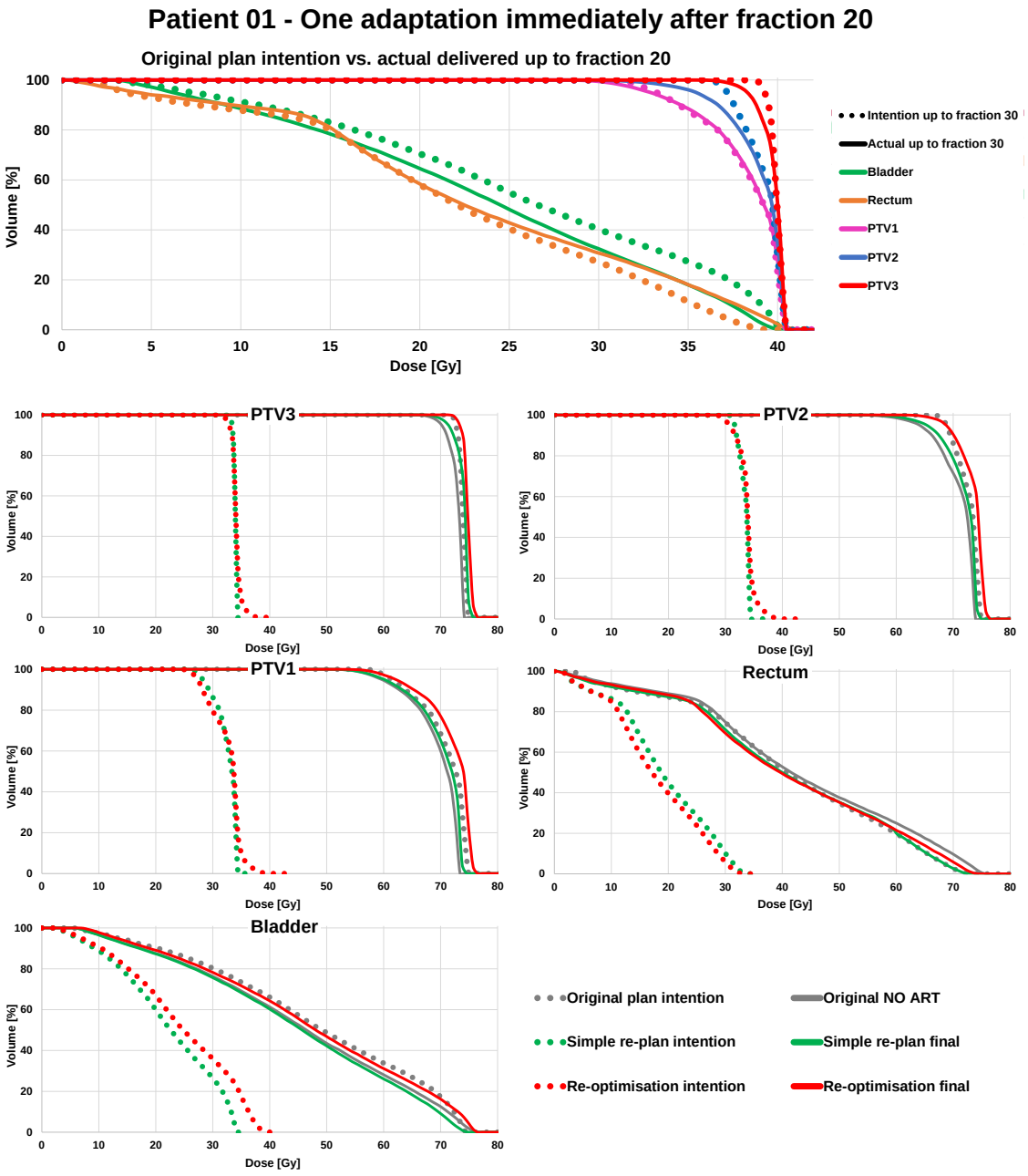

Figure 8.7: Dose volume histograms (DVHs) for a single patient showing the original plan intention and the actual delivered dose up to fraction 20, followed by DVHs comparing two adaptation methods immediately after fraction 20 against the original plan intention and the actual delivered dose without adaptation.

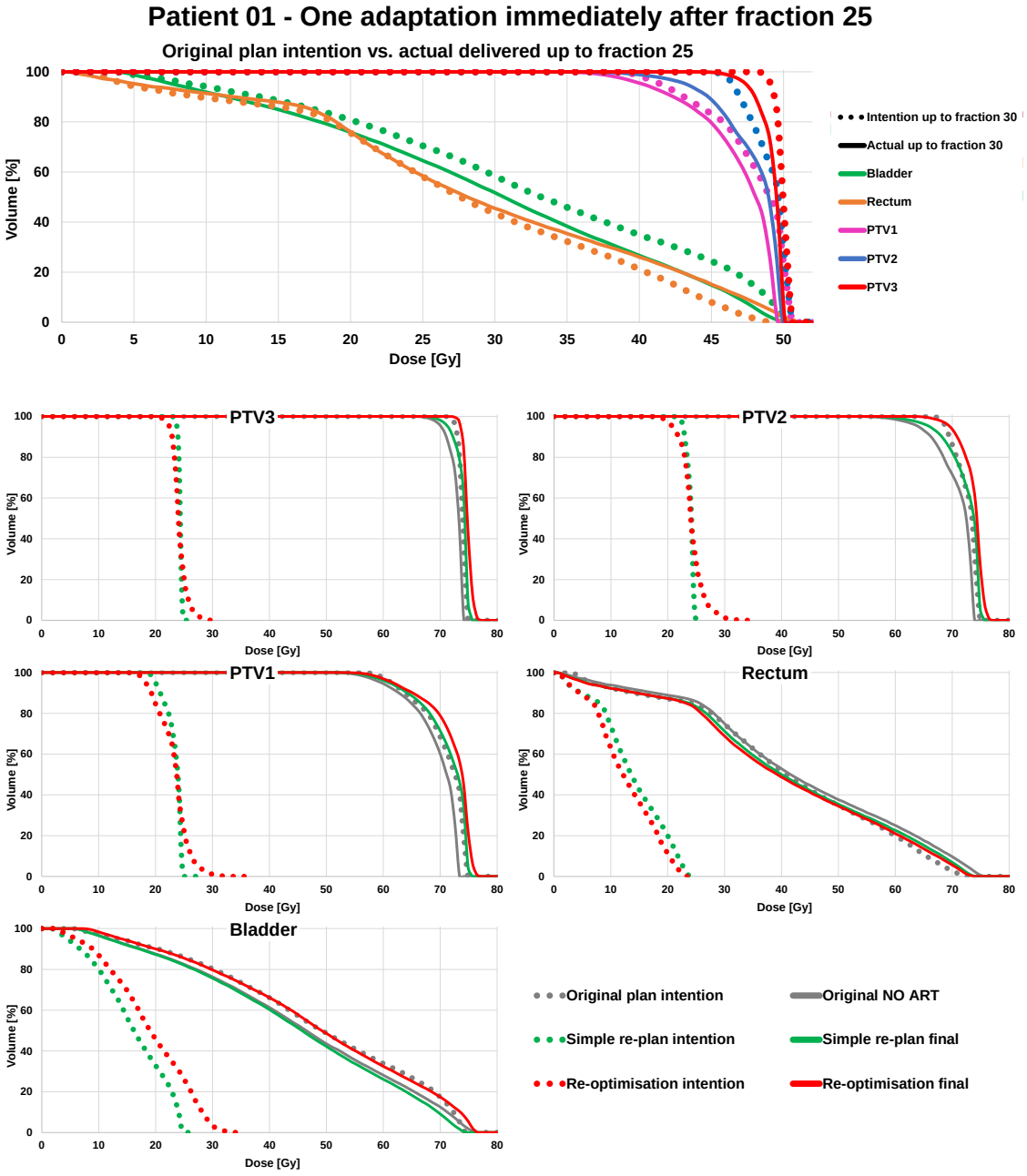

Figure 8.8: Dose volume histograms (DVHs) for a single patient showing the original plan intention and the actual delivered dose up to fraction 25, followed by DVHs comparing two adaptation methods immediately after fraction 25 against the original plan intention and the actual delivered dose without adaptation.

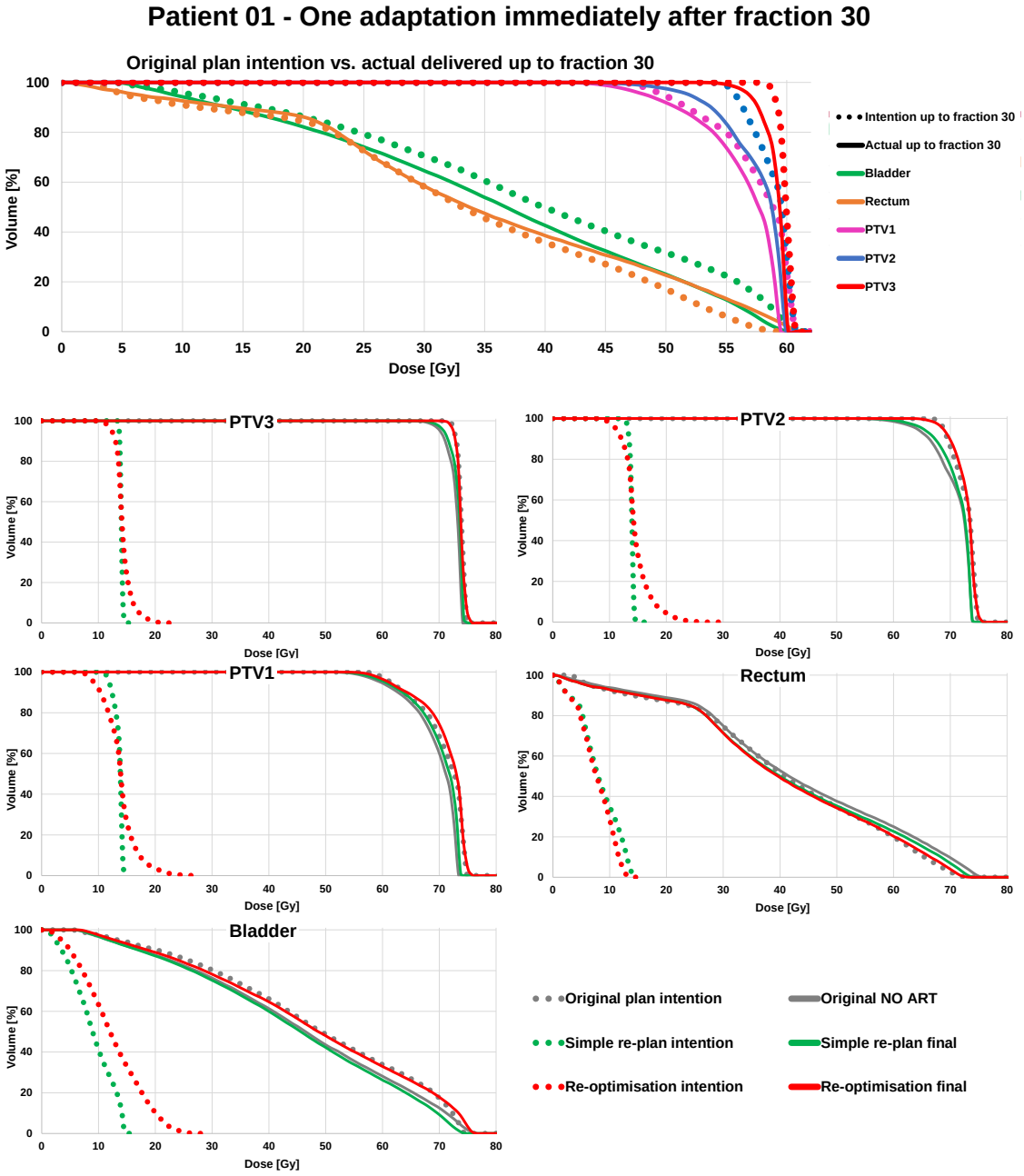

Figure 8.9: Dose volume histograms (DVHs) for a single patient showing the original plan intention and the actual delivered dose up to fraction 30, followed by DVHs comparing two adaptation methods immediately after fraction 30 against the original plan intention and the actual delivered dose without adaptation.

Figure 8.10 shows the mean dosimetric differences across all 10 patient test cases, comparing the two adaptation methods applied immediately after fraction 20, 25 or 30 and the actual delivered dose without adaptation against the original plan intention. Simple re-plan revealed better agreement with the original plan intention when applied immediately after fraction 25 compared to fraction 20 −in both target regions and organs at risk− while when applied after fraction 30 showed the largest differences. Re-optimisation exposed good agreement with intended dose when applied following fraction 25 and 30, compared to its application after fraction 20, in target regions. Cumulative dose to organs at risk was kept the closest to the intended dose distribution when re-optimisation was applied after fraction 30.

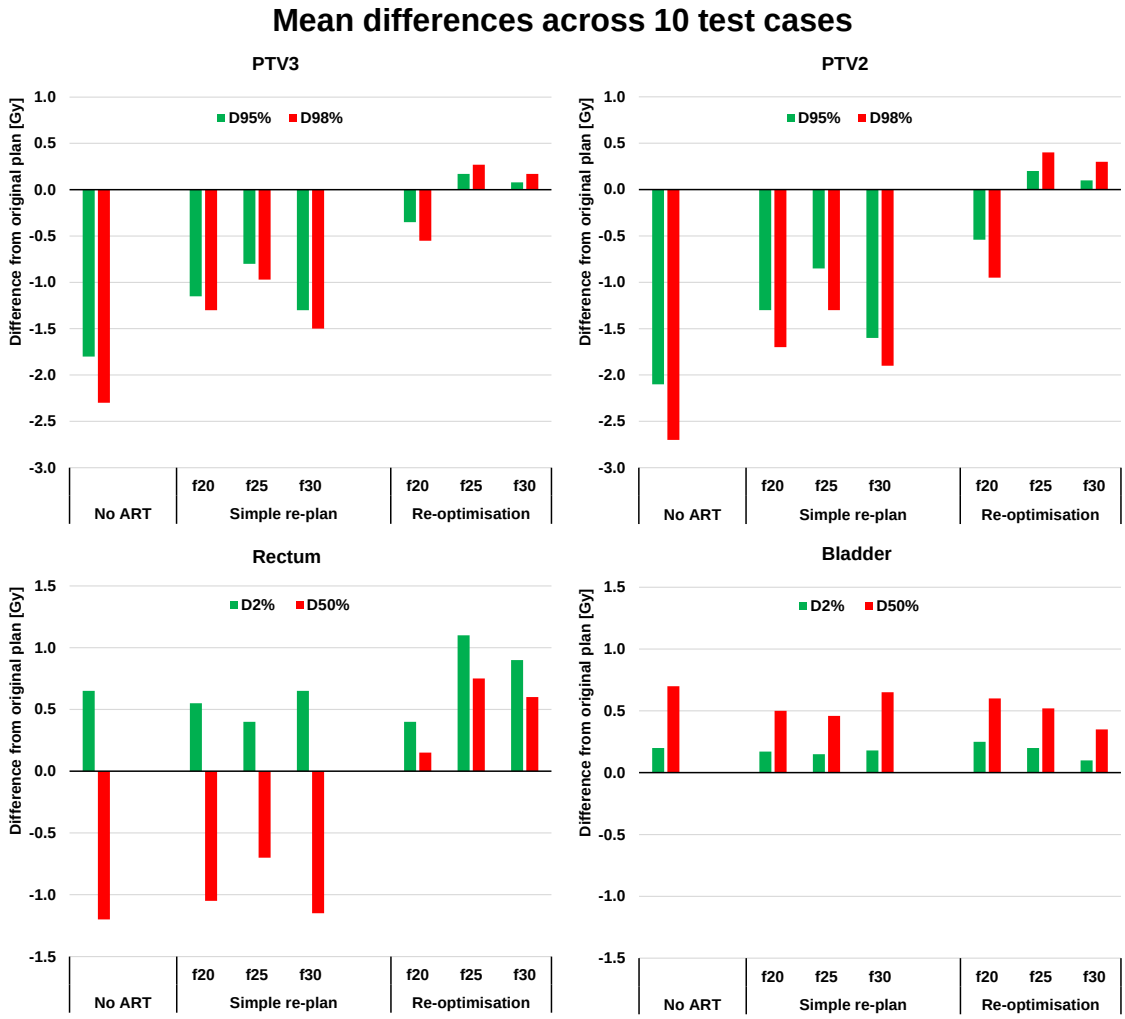

Figure 8.10: Mean dosimetric differences across 10 patient cases comparing the two adaptation methods immediately after fraction 20 (f20), 25 (f25) and 30 (f30) and the actual delivered dose without adaptation against the original plan intention.

PTV: planning target volume;

D95%: minimum dose received by at least 95% of volume;

D98%: minimum dose received by at least 98% of volume;

D2%: maximum dose received by at least 2% of volume;

D50%: maximum dose received by at least 50% of volume;

ART: adaptive radiotherapy.

To summarise adaptation timing comparison findings, for the 10 patient cases investigated in this study:

Simple re-plan timing

- Adaptation after fraction 20 was too early.

  Since the patients involved in the study showed relatively poor bladder and rectum preparation compliance after the simulated treatment plan adaptation and/or experienced further anatomical changes (e.g. weight loss/gain).

- Adaptation after fraction 30 was too late.

  Since all the patient cases employed experienced substantial anatomical changes throughout the treatment course, relatively high dosimetric changes occurred up to fraction 30. Adapting for the last 7 fractions resulted in only a small benefit.

- Adaptation after fraction 25 revealed closest result to the original intention.

  Since further anatomical changes have, on average, less of an effect compared to adaptation after fraction 20; while more fractions benefit from adapted plan compared to adaptation after fraction 30.

Re-optimisation timing

- Adaptation after fraction 20 was too early.

  As mentioned above, further anatomical changes were observed during the last 17 fractions in most cases, affecting the adapted plan. However, the final cumulative dose distribution following re-optimisation after fraction 20 was, on average, closer to the originally intended than any of the simple re-plan strategies.

- Adaptation after fraction 30 revealed closest result to the original intention.

  However, the intended dose distributions for the remaining of the treatment typically

include substantial 'hot-spots', in an attempt to compensate for target underdose in just 7 fraction, occasionally found near or within organs at risk.

- Adaptation after fraction 25 revealed slightly inferior results to adaptation after fraction 30.

  Target coverage was, on average, very similar. Organs at risk received slightly higher dose due to further anatomical changes.

Alternative method: re-optimisation using dose constraints per fraction

An additional variable may be used during re-optimisation method: the plan can be re-optimised based on accumulated dose using dose constraints for the final distribution (as before), while also introducing dose constraints for the remaining fractions (e.g. considering the rectal constraint of maximum 15% of volume receiving 70 Gy, for adaptation immediately after fraction 25, this constraint would be adjusted so as a maximum of 15% of volume to receive 22.7 Gy for the remaining of the treatment). This could −theoretically− allow the flexibility to, say, compensate for target region underdose while preserving low dose per fraction for organs at risk.

The practicality of this method was explored employing 5 of the patient cases used for the previous parts of this study. Three new plans were created taking into account the accumulated dose, while, for the remaining of the treatment 1 of the following constraints was used:

(i) Original dose constraints for both target regions and organs at risk;

(ii) Original dose constraints for target regions only;

(iii) Original dose constraints for organs at risk only.

No figures were included for the evaluation results of this method, to retain clarity, as these were similar to previous findings. The results can be summarised as follows:

- In scenario (i) the results were very similar to simple re-plan.
  Re-optimisation would attempt to compensate for, say, underdose in PTV2 which covers part of the rectum. To achieve this, additional dose would need to be delivered to the rectum which is prevented by the dose constraints.

- In scenario (ii) the results were very similar to re-optimisation.

  Considering the example above, no constraints related to rectal volume would prevent compensation of PTV2 underdose.

- In scenario (iii) the results were very similar to simple re-plan.

  (Same as first example)

No benefit in dose distributions was observed during re-optimisation method when additional constraints for the remainder of the treatment were included. However, these constraints may be too conservative for the specific scenarios (e.g. certain organs may be able to tolerate high doses during the last part of the treatment while others may not). Further investigations towards various tissue tolerances under such circumstances could allow for additional flexibility during treatment re-optimisation and potentially greater benefit.

Limitations

There are 3 limitations to this study. First, the small number of patient cases employed ($n = 10$) restricts confidence in any conclusions. Second, the application of a single re-plan since some cases could have benefited from more adaptations. This choice was made for simplicity and to retain similarity between the compared methods. Of course daily re-optimisation would allow superior target coverage while maintaining organ at risk constraints. Third, the fact that patient cases employed were chosen from a pool of patients that experienced substantial anatomical changes throughout the treatment course may represent the worst case scenario and not the average clinical practice. This choice was made for two reasons: to evaluate ART workflows in cases that would benefit the most (i.e. choosing random patients would mean that some of them may not have required treatment adaptation in the first place); and, amplify the dosimetric differences (i.e. large anatomical changes lead to considerable dosimetric changes and difference in performance between different methods would likely be noticeable).

Furthermore, it could be argued that using daily treatment monitoring and employing a voxel-to-voxel matching DIR algorithm, the dosimetric analysis could have been performed based on the actual tumour region (or the CTV) rather than the PTV. The choice of analysis based on PTV was based on the fact that the treatment adaptation would be designed using the PTVs, as well as to account for the uncertainties involved during manual delineating of contours used to drive the hybrid DIR algorithm and those involved during DIR (investigated in Chapter 4, p. 54).

Radiobiology-based adaptation

As shown earlier −Figures 8.4 and 8.5− radiobiological analysis of treatment adaptation could suggest that certain tolerances may be exceeded while cumulative DVH analysis

suggests the opposite. This finding was observed during re-optimisation method, during which an additional dose may be delivered in target regions that have been underdosed but may simultaneously mean that additional dose is subsequently delivered to sensitive normal tissues. Employing additional dose constraints was investigated, in an attempt to control this effect but was shown to have minimal benefit when conventional constraints are used. Alternatively, radiobiology-based constraints could be explored in an attempt to control re-optimisation in a more flexible and potentially clinically meaningful manner.

8.1.3 Conclusions

This retrospective study investigated several workflows towards offline adaptive radiotherapy in prostate cancer cases receiving VMAT treatments.

Dose accumulation: Out of the investigated DIR algorithms evaluated, Raystation's hybrid DIR algorithm with the use of controlling ROIs revealed the best results due to the additional advantage of being driven by these contours in areas of low image contrast and areas were new features appeared or disappeared.

Dosimetric treatment monitoring frequency: Weekly dosimetric treatment monitoring was found insufficient when compared to daily monitoring, primarily due to the random nature of the main anatomical changes observed in the pelvis region during prostate cancer radiotherapy. In such cases where random anatomical changes are expected, daily monitoring would be recommended, while in areas were mainly systematic changes are observed (e.g. cancer in the head and neck region) weekly monitoring may suffice.

Treatment adaptation method: The inability of simple re-planning during treatment adaptation to compensate for under/over-dose up to the point of adaptation was exposed. On the contrary re-optimisation method allows such dose compensations. How-

ever, in order to achieve dose compensation, 'hot-spots' may be generated near or within sensitive tissues. Such dose distributions may seem within tolerance when assessed using cumulative DVHs but radiobiological analysis can expose the opposite. Additionally, depending on the tumour type, extend and area the underdose occurred, simple 'boost' of dose delivery for the remaining of the treatment may not suffice. Therefore, conventional treatment planning protocols applicable to simple re-planning may not be adequate for re-optimisation procedures, for which alternative planning techniques −such as radiobiology-based optimisation− would need to be further investigated.

Treatment adaptation timing: Since simple re-planning can not compensate for under/over-dose, a late treatment adaptation (after fraction 30 out of 37) was not beneficial in the cases investigated herein. On the other hand, if treatment is adapted midway through treatment, further anatomical changes after the adaptation may hinder its success. Adaptation by simple re-planning immediately after fraction 25 was found to give the best results in this study. Similar issues were experienced during adaptation by re-optimisation, when this was performed midway through treatment. The best results for this method were observed when applied immediately after fraction 30, however, this coincides with the creation of the most sever 'hot-spots' when target region underdose occurred.

8.1.4 ART in head-and-neck cancer treatments

Radiotherapy treatment for cancers in the head and neck (H&N) region is particularly challenging due to the large number of sensitive structures typically surrounding the target volumes. Advanced techniques, such as IMRT, VMAT, TomoTherapy and others, facilitate highly conformal treatment plans with good coverage of PTV regions while enabling acceptable sparing of critical organs at risk [118]. However, anatomical changes such as tumour shrinkage, changes in normal glands and mucosa, or weight loss occasionally observed during the course of treatment [97; 119–121], can lead to suboptimal dose distributions [74; 122–125]. Such anatomical changes and deformations can not be compensated using simple patient re-positioning during online image guidance.

Adaptive radiotherapy (ART) techniques have been retrospectively and prospectively investigated, in an attempt to compensate for these anatomical changes [54; 59; 74; 126; 127]. Due to the gradual nature of the main changes observed during H&N radiotherapy, as well as the technical challenges involved and time restrictions in online ART workflows, offline ART has been suggested as the most practical option for this patient cohort in routine clinical practice [127]. Wu et al. [74] retrospectively simulated an offline ART workflow, employing six weekly helical CT scans for treatment monitoring and re-planning. Mid-course, bi-weekly and weekly adaptations were investigated, revealing 3%, 5% and 6% reductions in mean dose to the parotids, respectively (assuming 1 week delay between re-scan and actual implementation of adapted plan). Schwartz et al. [54] performed a prospective clinical trial during which re-planning was performed based on daily CT-on-rails scans. Significant reductions of the mean dose to parotids were observed following both one or two adaptations during the course of treatment. To conclude, the main H&N ART studies presented in the literature utilised high quality CT scans for the dosimetric monitoring, offline ART was performed, while only the 'simple re-planning'

technique was considered.

In this thesis, ART workflows for H&N cancer cases were not investigated for the following reasons:

- CT-to-CBCT DIR was considered unreliable

  The performance of available DIR algorithms for CT-to-CBCT registration in the H&N region was not considered reliable nor could be reliably evaluated, particularly in the low contrast regions (e.g parotids, tumour and nodal boundaries, optic chiasm, optic nerve). RayStation's Hybrid algorithm with the use of controlling ROIs could potentially assist in reliable and accurate registration results, however, this would require contouring of soft tissue ROIs which could not be reliably performed due to relatively low image quality.

- Low CBCT image quality

  CBCT scans employed during IGRT in our institution are sufficiently good for patient positioning based on 3D bone matching. However, soft tissues have low contrast and can not be accurately identified. This is true for both parotid glands and tumour boundaries. Manual contouring of these regions would, therefore, be very time consuming and potentially inaccurate, especially considering the lack of contouring expertise in this type of images.

- CBCT auto-contouring considered unreliable

  Likewise, CBCT auto-contouring was considered unreliable, mainly due to the lack of 'ground truth' which would need to be acquired through manual contouring. The introduction of artificial changes to represent tumour shrinkage/weight loss and concurrent change of parotid volume and position was not feasible using the available software tools.

Improvements in CBCT image quality could improve the prospects of offline ART following the workflows investigated in this thesis. Immediate improvements, to a certain extent, could be achieved by altering the imaging parameters and subsequently increasing the imaging dose. However, it would be difficult to justify such action lacking palpable indications of potential clinical benefit. Two potential ways to overcome this issue were hypothesised, under the assumptions that weekly treatment monitoring would suffice for this patient cohort (considering the gradual nature of changes typically observed in H&N patients), and that higher quality (and dose) CBCT scans would provide sufficient low contrast visibility: (a) perform low-dose CBCT scans 4 times a week and a high-dose CBCT once a week; or (b) perform planar imaging 4 times a week and a high-dose CBCT once a week, during IGRT.

Considering the finding of the prostate cancer ART investigations performed in Section 8.1 (p. 130) the likely similarities and differences applicable to H&N ART under similar conditions are summarised below:

Dose accumulation: Assuming manual contouring of at least critical soft tissue structures (e.g. parotids, GTV/CTV, optic structures) could be performed on online CBCT scans, RayStation's Hybrid DIR algorithm with the use of controlling ROIs would be expected to produce the best registration results. Due to the relative rigidity of the head and neck region and the large number of bony structures present, most DIR algorithms would be expected to produce good results, except in low contrast regions.

Dosimetric treatment monitoring frequency: Weekly dosimetric treatment monitoring would be expected to be sufficient, compared to daily monitoring, since most of the anatomical changes observed in this patient cohort are typically gradual (exceptions include nasal cavity filling and swelling variations).

Treatment adaptation method: Simple re-planning is expected to be the most appropriate option for H&N ART. This is not only because of the relatively small and infrequent expected changes of target dose, but also due to the risks associated with the potential introduction of hot-/cold-spots for compensation of under-/over-dose of certain regions. As observed in the previous investigation, anatomical changes occurring after plan adaptation by re-optimisation may result in sever over-dose of organs close to the treatment target, which in the case of H&N patients might have substantial clinical consequences.

Treatment adaptation timing: As Schwartz et al. [54] observed in their prospective ART trial, typically one and occasionally two treatment plan adaptation is expected to be required. In their study, the average fraction the first replan was performed was 15.5 out of 30-33 fractions in total (i.e. midway through treatment), while the mean fraction for the second replan was 20.5.

Chapter 9

Summary

This chapter summarises the main results of the work outlined in this thesis and discusses possible future directions for a wider implementation of adaptive radiotherapy workflows in routine clinical practice.

9.1 Discussion and Conclusions

Auto-segmentation

Several auto-segmentation algorithms have been evaluated on CT datasets of head-and-neck and prostate cancer patients (Chapter 2, p. 27). The results revealed generally greater uncertainties to inter-observer variability. At this stage, none of the evaluated algorithms would be adequate for incorporation into a semi-automated ART workflow without or with minimal manual corrections. Auto-segmentation on CBCT datasets was considered unreliable, primarily due to considerable uncertainties in the potential evaluation of the results, arising from relatively low soft tissue contrast and limited specialist experience of manual contouring in such datasets.

Deformable image registration

DIR is an important tool for adaptive radiotherapy procedures as it can provide voxel-to-voxel correspondence between two image datasets, thus enabling contour and dose mapping for treatment monitoring. Three DIR algorithms have been evaluated in Chapter 4 (p. 54) for CT-to-CT and CT-to-CBCT registration. The use of controlling ROIs (contours drawn in both images) to drive the registration, in addition to the image intensity based algorithm offered by RayStation's hybrid implementation proved the most accurate and robust method, especially in areas of low contrast. Regrettably though, the superior performance of this hybrid implementation comes with the additional requirement of accurate manual or automated segmentation of regions of interest in the CT and CBCT datasets, which may be time consuming or even practically impossible due to poor image quality.

Dosimetric treatment monitoring

Dosimetric treatment monitoring is an integral step for certain offline ART workflows, allowing the assessment of dosimetric impact caused by observed anatomical changes on online or repeat scans. Chapter 5 (p. 68) initially investigated a controversy in the literature, in a methodology to calculate actual delivered dose in a radiotherapy fraction or different stages of a 4D-CT scan. This investigation confirmed that direct dose deformation following DIR between the original and a repeat CT scan, without a dose re-calculation step, (herein 'Dose_Deform') is not appropriate. It was also shown that, dose recalculated on the original image as deformed to match the new anatomy (herein 'Dose_Recalc'), can reveal accurate results. This later method can be useful when direct dose calculation on the online-acquired image is unreliable, and provided DIR performance is accurate.

The accuracy of various methods for direct dose calculation on CBCT images acquired during image guidance was then inspected. Highest accuracy was revealed using the 'tissue based HU-to-density curve generation' approach, during which the CBCT image is segmented into various pre-defined tissue types, the mass density of which is then used during dose calculations. Due to its high accuracy, this method was afterwards chosen during the investigation of ART workflows. However, this method is subjective and manual, and therefore not ideal for routine clinical practice.

Deformable dose accumulation

The ability to accumulate delivered dose from any given fraction on a single frame of reference is of great importance, as the cumulative effects of random or gradual changes can be assessed, enabling accurate clinical decisions for treatment adaptations. Furthermore, accumulated dose can be fed back to the treatment planning system in order re-optimise the new treatment plan based on delivered dose to date. This process can be facilitated using DIR algorithms, following the registration of each daily image to the reference anatomy and the subsequent warping of re-calculated dose distributions. A new method for the evaluation of this process has been presented in Chapter 6 (p. 92), and subsequently applied for the evaluation of a commercial algorithm.

Plan robustness

Investigation of treatment plan robustness to frequently observable anatomical changes is an important process prior to the potential decision for clinical implementation of dosimetric treatment monitoring and ART, as it enables the identification of cases that would benefit from treatment interventions. Chapter 7 (p. 107) investigates the robustness of several treatment techniques in nasopharyngeal carcinoma and prostate cancer cases. In

the nasopharyngeal carcinoma cases, proton (IMPT) and photon (VMAT, TomoTherapy) plans were compared under clinically realistic artificial anatomical changes. IMPT plans were severely affected by anatomical changes with possible dramatic changes to delivered dose distributions and clinical outcomes. In contrast, both rotational IMRT techniques revealed smaller but occasionally considerable dosimetric alternations. In the prostate cancer cases, the robustness of VMAT plans was investigated under clinically realistic scenarios, revealing considerable dosimetric changes in some cases.

Adaptive radiotherapy workflows

Several techniques towards a clinically justifiable, resource viable and widely applicable ART workflow were investigated in Chapter 8, p. 130. The investigation was concentrated on prostate cancer patient cases treated with daily CBCT-guided VMAT, and incorporated the findings and techniques outlined in previous chapters to identify optimum: dose accumulation method, dosimetric monitoring frequency, treatment adaptation technique and adaptation timing. For the particular patient cohort and treatment technique daily dosimetric treatment monitoring by direct dose calculation on CBCT scans and accumulation using RayStation's hybrid DIR with the use of controlling ROIs is recommended. In terms of adaptation method, simple re-plan exposed the inability to compensate for potential under-/over-dose up to the point of adaptation contrary to re-optimisation that allows dose compensation. Re-optimisation, however, can produce hot-spots in an attempt to achieve these dose compensations which can induce unexpected radiobiological outcomes. In the scenario of single plan adaptation, execution within the third quarter of treatment fractions was found favourable for both techniques.

9.2 Future directions

Improvement in auto-segmentation and DIR algorithms

Further improvements in auto-segmentation and DIR algorithms would be required before their implementation in automated or semi-automated ART workflows with no or with minimal manual interventions. Auto-segmentation for CBCT or repeat CT images can be replaced by 'auto-re-contouring' following DIR (or contour mapping) between the planning CT and a CBCT. However, the evaluation of either auto-segmentation or DIR involving a CBCT scan is hindered by the low soft tissue contrast occasionally seen in these images. Improvements in CBCT image quality would therefore enhance the performance of these processes.

Optimisation of planning software to facilitate ART workflows

As discussed throughout this thesis, an efficient offline ART workflow would benefit from a robust auto-segmentation algorithm which would minimise manual contouring time; an accurate and robust DIR algorithm to facilitate dose accumulation and contour mapping; functions for practical dose calculations on daily online acquired scans (e.g. dose calculation on CBCTs, compensation for limited field-of-view); and convenient re-planning and re-optimisation capabilities with enhanced capabilities (e.g. radiobiological optimisation and assessment). Currently not many commercial treatment planning systems (TPS) incorporate the full range of these processes, meaning that the users would need to purchase individual software resulting in additional technical and data handling concerns. Thus, TPS incorporating the above processes are crucial for wider implementation of ART in routine clinical setting.

Design of clinical ART protocols

Clinical ART protocols would need further development with inter-disciplinary input, including: evaluation and quality assurance procedures for new software tools and workflows; specialized training; pre-defined adaptation triggers and adaptation techniques for individual patient cohorts.

Investigation and understanding of radiobiology following adaptations

Radiobiological considerations would need to be integrated into the ART decision-making process, and during the acceptance of adapted plans. New radiobiological or dose-volume constraints may be required in order to allow treatment re-optimisation by dose compensations and consequent non-uniform dose distributions in target regions. Application of dosimetric treatment monitoring and correlation of dose-effect relationships to actual accumulated dose distributions, rather than estimates based on original plans, is expected to provide better correlations and allow improvements in radiobiological modelling.

Investigate radiobiology-based plan optimisations

As mentioned above, radiobiological consideration may by vital during plan adaptation, and especially when re-optimisation is performed. Further development of radiobiology-based plan optimisation may allow safer and reliable implementation of this technique.

MRI-guided ART

Drawbacks associated with the quality of online acquired x-ray based images, such as CBCT and MVCT, may be overcome with the clinical implementation of MR-linacs [128–131] or MRI-equipped cobalt radiotherapy [132] systems. The very good soft tis-

sue contrast in MR images may allow for improved auto-segmentation and DIR algorithm performance, as well as easier assessment of their results [133].

Online ART

Online ART can efficiently adapt to random daily changes, in contrast to offline ART workflows. A new plan can be created −or the existing re-optimised− based on the daily anatomy, following the online scan, and while the patient remains immobilized on the treatment couch [134; 135]. The whole adaptation process, therefore, needs to be performed very quickly, including potential clinical decision making and plan quality assurance tests. Dedicated planning protocols and plan acceptance should be developed while additional pressure and responsibilities are put towards radiotherapy radiographers, physicists and clinicians tasked with the process.

Plan-of-the-day techniques, whereby several plans are created for a given patient with the most appropriate one chosen following the assessment of an online-acquired scan [136; 137] are less technically challenging and are currently finding their way in routine clinical practice. The implementation of this online ART technique requires specialised training of radiotherapy radiographers assigned for choosing the appropriate plan [138; 139], some authors investigated algorithms for automatic selection of the appropriate plan-of-the-day [140]. Plan-of-the-day ART techniques are particularly useful in cases with predictable, frequent and substantial anatomical changes, such as bladder cancer treatments. RAIDER [141] is a phase II clinical trial currently underway investigating the feasibility and long term outcomes of this ART technique in bladder cancer treatments.

ART based on tumour response

The main ART objective discussed throughout this thesis has been the adaptation against anatomical changes. However, certain researchers are proposing the application of the same principles for adaptation of the treatment plan according to tumour response. This response could be either observed through repeat CT, MRI or functional imaging techniques such as PET [142–145].

Clinical trials to investigate actual clinical benefit from ART

Once efficient and robust ART workflows are established prospective clinical trials would need to investigate their actual clinical benefit. This would be especially important for offline adaptive radiotherapy techniques utilising plan re-optimisation method were, as discussed in this theses, non-uniform dose distributions can be generated in the target regions and hot-spots in certain healthy tissues.

Appendix A

Cherenkov Emission Imaging in Radiotherapy

A.1 Real-time Cherenkov emission portal imaging during radiotherapy: a proof of concept study

This appendix outlines the investigation of real-time portal imaging during radiotherapy using an optical camera, utilising the Cherenkov emission effect. This technique can be especially useful for radiotherapy equipment that lack alternative portal imaging options, such as CyberKnife®. This work, and the content of this appendix, has been accepted for publication by the 'Physics in Medicine and Biology' journal [5].

A.1.1 Introduction

Cherenkov radiation is emitted when charged particles travel faster then the speed of light in a given dielectric medium [146]. During x-ray radiotherapy, secondary electrons

are predominantly generated from Compton scattering which have sufficient energy to produce Cherenkov Emission (CE) in water or tissue. CE was shown to be measurable with an optical imaging system, given that the main wavelength of emission is within the visible spectrum and shown to be proportional to radiation doses deposited in the medium [147].

The incorporation of CE in electronic portal imaging device (EPID) detectors has recently been investigated [148], while Cherenkov radiation imaging has been utilised as a dosimetry technique during photon [149–152], electron [149; 153] and proton [154] radiotherapy. Furthermore, real-time superficial dosimetry has been demonstrated in both animal and human studies [155; 156].

EPID-based portal imaging may be employed for pre-treatment positioning or intra-fraction tumour tracking. While some commercial linear accelerators incorporate portal imaging equipment, the CyberKnife system does not offer such option. Due to the fact that this system utilises multiple non-coplanar and non-isocentric beams, a conventional portal imaging device would not be practicable.

In this study the feasibility of real-time CE portal imaging during CyberKnife radiotherapy is investigated. This would enable real-time verification of in-patient targeting and treatment delivery accuracy.

A.1.2 Methods

A medical linear accelerator (CyberKnife®, Accuray, Sunnyvale, CA) irradiated a partially filled water tank ($300 \times 300 \times 350$ mm) using a 60 mm diameter circular beam of 6 MV energy at a dose rate of 10 Gy per minute at 800 mm source-axis-distance (SAD), as shown in Fig. A.1a. A number of attenuating materials were placed on a moving platform between the beam source and the water tank, while an opaque 10 mm thick white solid

water slab (Solid Water HE, Gammex Inc, Middleton, WI) was placed at the beam exit face to ensure that only the CE due to radiation at the exit face would be imaged. A gated electron-multiplying-intensified-charged-couple device (emICCD) (PI-MAX4: 512 EM, Princeton Instruments) with a commercial lens (Cannon EF 135 mm f/2L USM) positioned at the exit face at a small incident angle (to reduce direct radiation to the camera) imaged the induced CE at the surface of the solid water slab. Image acquisition was synchronised with the beam pulses using the trigger signal from the accelerator in order to maximise the signal-to-background ratio [157].

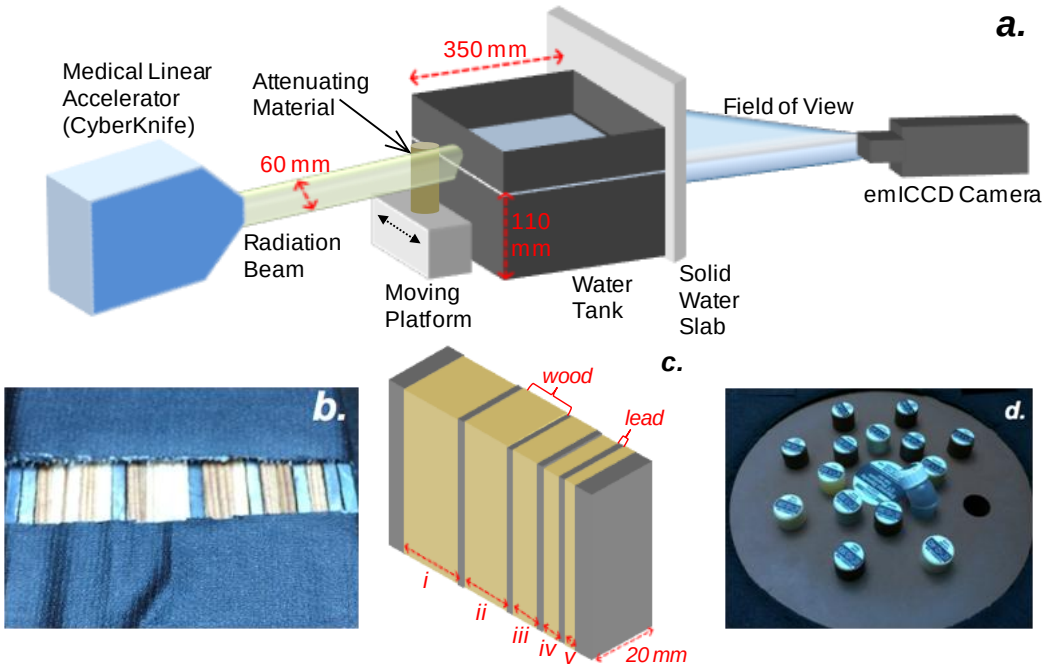

Figure A.1: ***a.*** Schematic diagram of the experimental setup; ***b.*** lead/plywood graticule with ***c.*** its corresponding schematic, and ***d.*** tissue equivalent rods (RMI phantom).

A number of attenuating materials were used to assess the resolution and contrast of the measured CE at the exit face. For assessment of the resolution a graticule was constracted from a number of lead sheets ($2 \times 60 \times 20$ mm) spaced at regular intervals using plywood separators, Fig. A.1b,c. To asseess the contrast of the measured CE a number of 28 mm

diameter tissue equivalent rods (RMI 467 phantom, Gammex Inc, Middleton, WI) were used Fig. A.1d.

The water tank was filled to a depth of 110 mm, and the phantoms were placed such that the lower 50% of the radiation beam travelled through the phantom and the water, whereas the upper 50% of the beam travelled through the phantom and air, before being measured at the exit face through the solid water slab. CE images were acquired at 30 frames per second (fps) with each frame accumulating CE for 5 radiation pulses. All data were background-subtracted using images with no radiation. The background-subtracted images (30 frames for static and 3 frames for dynamic) were then median-filtered to produce the CE images for data analysis.

CE-based static and dynamic portal images were visually inspected for an initial proof-of-concept validation. Resolution and contrast were assessed by analysing the profile plots through the two regions of the static images, facilitating quantitative evaluation. Resolution was quantified by measuring the distances of the minima in the profile plots, associated with the position of the lead sheets. To measure relative contrast, the mean percentage decrease of signal in the profile plots due to the presence of the tissue equivalent rods was calculated with respect to the normalised maximum.

The experimental configuration for contrast assessment in the static configuration was replicated on an Elekta Synergy® (Elekta AB, Stockholm, Sweden) linear accelerator with portal images being acquired using the iViewGT™ system, to allow direct comparison of CE-based portal imaging against current technology. For this experiment a 240×240 mm square beam of 6 MV nominal energy at a dose rate of 3.6 Gy per minute at 1000 mm SAD was employed, with portal images acquired for 50 monitor units (or 0.33 Gy to 1000 mm SAD) acquisitions.

A.1.3 Results

A Chereknov image using the stationary graticule is shown in Fig. A.2 along with profile plots of the measured CE intensity. The intensity profiles though both air and water are self-normalised to allow quantitative analysis of the detected resolution, corresponding to the distance between each lead sheet, Fig. A.2b,c and Table A.1. The lowest separation of 3.40 mm (spacing 'v' in Fig. A.1c) is clearly evident in the portal image (Fig. A.2a) and measured as 3.58 mm and 3.92 mm for radiation travelling through air and water, respectively. Larger errors are observed at the two edges of the detected signal, which is primarily due to data capture geometry and beam divergence. Each lead sheet with a width of 2 mm is visible, highlighting the potentials of higher resolution imaging.

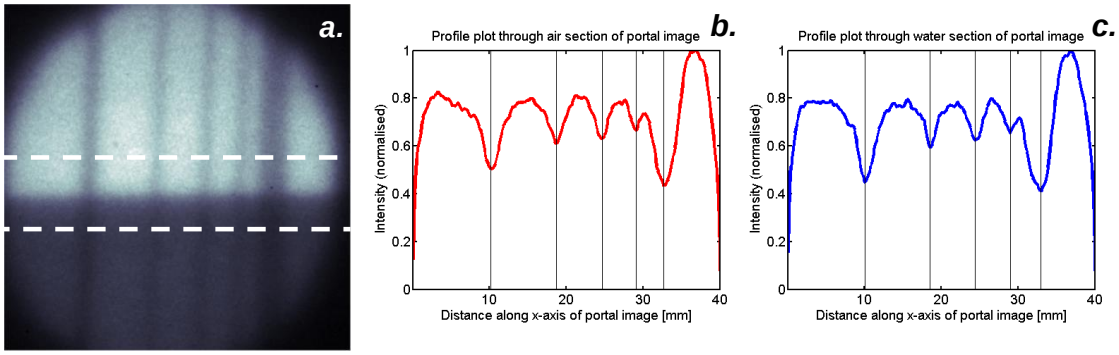

Figure A.2: ***a.*** Cherenkov emission portal image of the graticule phantom acquired from the beam exit face: the white dashed lines represent the profiles used to calculate the contrast for the radiation travelling through (***b.***) water and (***c.***) air.

Table A.1: The actual and measured distances between lead sheets from resolution graticule phantom.

	Actual [mm]	Recovered [mm]	
		Air	Water
i	10.40	10.60	10.80
ii	8.80	8.50	8.42
iii	6.00	6.00	5.92
iv	4.40	4.42	4.58
v	3.40	3.58	3.92

Portal images in the presence of tissue equivalent rods were analysed and the intensity profiles used to assess the measurable relative contrast for each phantom. Figure A.3 shows example portal images of the 'SB3 Cortical Bone' phantom as well as the corresponding self-normalised profile plots.

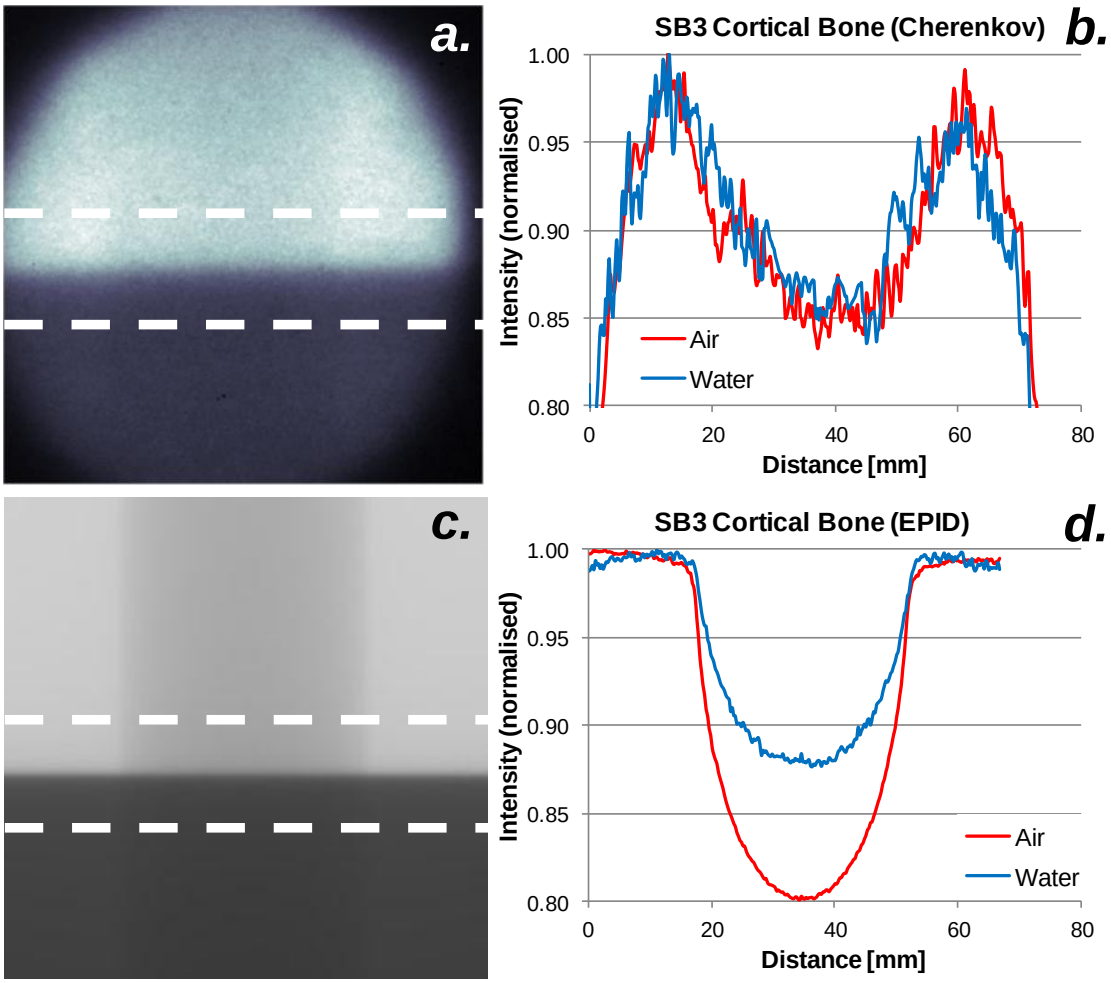

Figure A.3: ***a.*** Cherenkov emission portal image of the 'SB3 Cortical Bone' tissue equivalent rod with ***b.*** the associated normalised profile plots; ***c.*** EPID-based portal image of the same tissue equivalent rod (intensities inverted to match the CE intensities) with ***d.*** the associated normalised profile plots.

The use of tissue equivalent phantoms demonstrate that CE emission detected at the exit face are sensitive to small contrasts often seen in biological tissue. As seen in Fig. A.4 and summarised in Table A.2, for tissue equivalent rods with 28 mm diameter, a contrast in CE up to 15% was observed with an electron density of 1.69 relative to water, with similar contrast seen for beam passing through the air and water-filled part of the water tank. On the other hand, EPID-based portal images revealed higher contrast for beam travelling through the empty part of the water tank, with lower contrast for beam travelling through the water-filled part.

Real-time video (30 fps) CE imaging of a moving graticule and a tissue equivalent rod was recorded (snapshots shown in Fig. A.5, and video available in supplementary material of paper in PMB website [158]). Even though of inferior image quality due to the substantially lower number of frames in each median-filtered stack, the periodic movement (0.25 Hz) of the phantoms can be observed.

Table A.2: Properties of tissue equivalent rods together with the relative Cherenkov emission (CE) and EPID-based portal image contrast.

Name if tissue equivalent rod	Electron density relative to water	Relative CE contrast [%]		Relative EPID contrast [%]	
		Air	Water	Air	Water
SB3 Cortical Bone	1.69	15	14	19	12
CB2 - 50 CaCO3	1.47	12	11	16	10
CB2 - 30 CaCO3	1.28	10	10	15	8
B200 Bone Mineral	1.10	9	10	13	8
IB Inner Bone	1.09	8	10	13	8
BRN-SR2 Brain	1.04	9	8	12	7

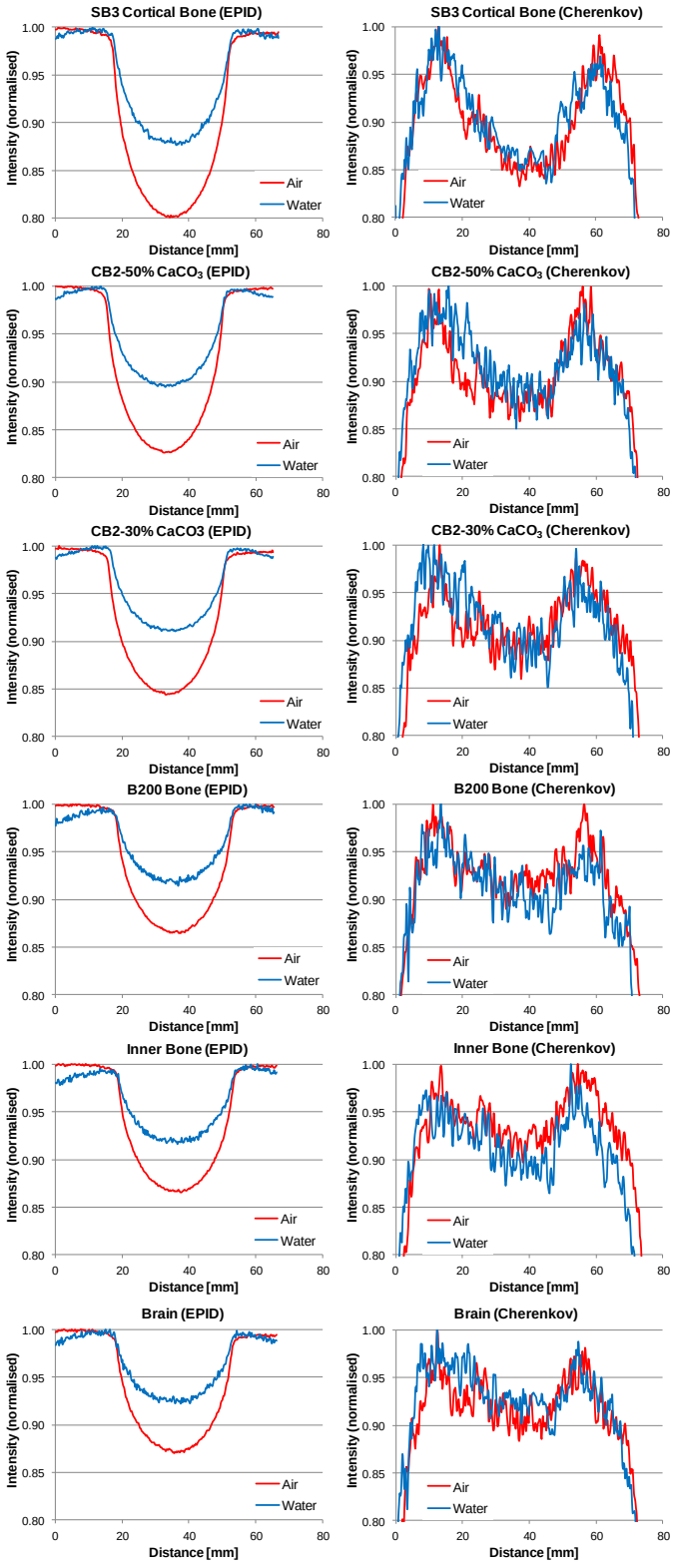

Figure A.4: Normalised profile plots of EPID-based (left) and Cherenkov emission (CE)-based portal images for a number of tissue equivalent rods.

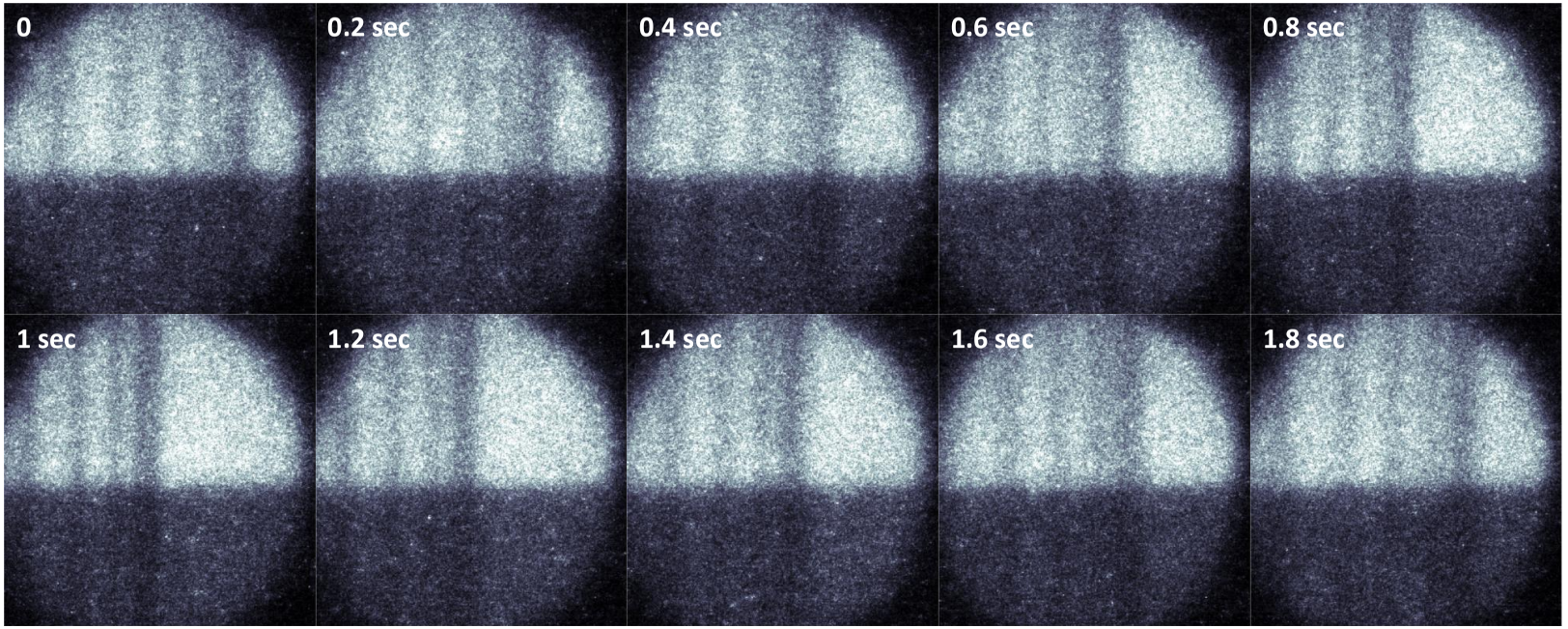

Figure A.5: Snapshots from CE video imaging of moving graticule phantom.

A.1.4 Discussion and Conclusions

The use of CE-based portal imaging was demonstrated using both a highly attenuating resolution phantom and tissue mimicking phantoms. It was shown that objects with as little as 3.4 mm separation are easily detected, while tissues with relative electron density of greater than 28% (with respect to water) demonstrate CE contrast of 8% or greater. Video rate CE portal imaging was achieved to monitor movements of the graticule and tissue equivalent phantoms in real-time. Therefore, CE-based portal imaging, a novel imaging technique during radiation therapy, can be potentially applied for systems that lack alternative options such as EPID.

Contrary to EPID-based portal imaging, CE portal imaging only utilizes visible optical photons produced at the exit surface through the CE effect while only charged particles above a threshold energy can cause CE. This minimizes the contribution of scattered radiation, resulting in higher measured contrast when radiation passes through water or tissue (as compared to EPID-based portal image), which highlights for the first time the additional benefits of CE portal imaging over EPID-based techniques.

The demonstration of CE-based portal imaging can be potentially useful in many applications. Given that the use of CE for detection of entrance and exit dose has been previously demonstrated, the current work indicates that this method could also be used to detect small deviations in patient positioning and intra-fraction anatomical movements.

Appendix B

Deformable Image Registration Algorithms

B.1 Review of DIR algorithms

This appendix provides a brief introduction to DIR algorithms and discusses their main components along with their advantages and limitations. Several reviews have been recently published in the literature with exhaustive overviews of DIR algorithms their mathematical and computational formulations [28; 159].

B.1.1 Introduction

Image registration is the process of defining correspondence between two images; in other words, transforming two images into one coordinate system. Typically two images are considered, one of which is referred as the moving image (or floating, source) and the other one the stationary image (or fixed, target, reference). In this setting, the moving image undergoes a transformation so as to 'match' the stationary image. The goal of

the image registration process is to find the optimal transformation that achieves perfect correspondence.

Rigid image registration (RIR) algorithms typically model the transformation between the two images purely rigidly, or linearly (and for the image as a whole). Six degrees of freedom exist in the transformation model, three rotational and three translational (even though a few other RIR options exist with 9 or 12 degrees of freedom).

Deformable image registration (DIR) algorithms, on the other hand, can model each image voxel independently with multiple parameters for each, resulting in hundreds to millions of degrees of freedom. For instance a commonly used $512 \times 512 \times 100$ image matrix with three parameters for each voxel would employ $78,643,200$ parameters.

DIR algorithms can generally be described by three main components:

(a) Transformation or deformation model

(b) Similarity metric or objective function

(c) Optimization method

B.1.2 Transformation or deformation model

The transformation model (also called deformation model) employed in the DIR algorithm determines the underlying relationship between the coordinates of the two images to be registered. The complexity and flexibility of the model is typically proportional to the degrees of freedom it possesses, but in general, the more complex the model the more computationally demanding the optimisation becomes. Moreover, each model is typically based on several assumptions regarding the underlying deformation to be estimated, limiting its applicability to certain scenarios.

According to Holden [23], transformation models employed by DIR algorithms can be identified in two main groups according to their underlying physical basis:

(i) those influenced by physical models,

e.g. elastic model, diffusion model, viscous fluid flow model and others.

(ii) those influenced by approximation and interpolation theories,

e.g. free form deformations, radial basis functions, elastic body splines and others.

One of the most widely used physical models is the Demons formulation which is inspired by the diffusion model and specifically Maxwell's Demons [24]. In this approach the image boundaries are seen as membranes through which the image diffuses under the influence of Demons forces. The computation of these forces can be highly efficient and robust with advanced programming techniques and has therefore attracted the interest of researchers and manufacturers of commercial products. It generally utilises image intensity-based similarity metrics and therefore possesses the advantages and limitations discussed below.

One of the most important techniques based in the interpolation theory are the thin-plate splines which lie under the radial basis functions family. Here the displacement value at a given point is a function of its distance to another given reference point. The advantage of these techniques is that the result in highly smooth while feature-based similarity metrics can be employed. In addition to the advantages and limitations of feature-based techniques discussed below, another important limitation is their inefficiency to allow for large local deformations, since the deformation field as a whole is affected by neighbouring forces.

B.1.3 Similarity metric or objective function

The similarity between the two images is computed in each iteration of the DIR algorithm in order to 'drive' the registration. Depending on the similarity metric of choice, the aim of the registration is to maximise or minimise it, meaning that (near-)perfect overlap between the images is achieved and it is therefore also referred as the objective function. These similarity metrics can be:

(i) Feature-based,

(ii) image intensity-based,

(iii) hybrid

Feature-based (or geometric) techniques utilize landmark points, lines or segmented surfaces (e.g. organ contours) to drive the registration and typically aim to minimise the distance between the corresponding features in the two images. These features must therefore be defined in both images requiring an additional step of manual or automated feature extraction. An advantage of these methods is their applicability to both intra- and inter-modality image registrations with consistent accuracy. Furthermore, they can be more robust in situations were their is lack of volume conservation between the two image instances or when new features appear in one of the images. Their limitation is the requirement of feature extraction which may occasionally be a very time consuming process, while when automated feature extraction techniques are employed potential errors will be translated into registration errors.

Image intensity-based (or iconic) techniques employ grey level similarity measures aiming to maximise the similarity between the two images. Such similarity measures include the cross correlation, sum of squared differences, mutual information and correlation ratio.

The advantage of intensity-based techniques is that they do not require prior feature extraction and are typically fully automatic. However, they are generally more robust when used for intra-modality registrations than inter-modality cases, even though advanced techniques and the use of mutual information measure can archive good inter-modality image registration results. Another limitation is the general assumption that both images share the same characteristics (i.e. one-to-one correspondence exists).

Hybrid techniques combine the characteristics and requirements of both feature-based and intensity-based techniques in an attempt to combine the advantages of each and amplify their performance. In these cases, one of the techniques could serve as the initialisation step before the other commences; or one could serve as a constraint for the other; or they may be coupled as a single algorithm.

B.1.4 Optimization method

As mentioned earlier, the aim of image registration algorithms is the identification of the optimal transformation that would perfectly align the two image sets, by maximising (or minimizing) the similarity metric. Optimisation approaches aim to drive the registration through the high-dimensional solution space while avoiding local maxima (or minima).

One of the most commonly employed techniques in medical image registration is the 'coarse-to-fine' approach. In this technique a number of registration steps are performed at coarser image levels first before moving to finer image levels.

Appendix C

3D Printed Phantom

C.1 Design and development of a patient specific 3D printed phantom: a feasibility study

This appendix outlines the procedure performed for the preparation of a patient-specific 3D printed phantom.

C.1.1 Introduction

Phantoms are tools routinely used for medical imaging and radiotherapy related commissioning and quality assurance procedures in clinical setting. As discussed in Section 1.4.2 (p. 15), physical or digital phantoms may be chosen for the evaluation of certain workflows such as DIR algorithm performance, each having its own advantages and limitations.

The aim of this project was to investigate the feasibility of creating a patient-specific 3D printed phantom for the evaluation of workflows for dose calculation on CBCT scans, DIR and deformable dose accumulation.

C.1.2 Methods

A workflow has been implemented for the construction of a patient-specific 3D printed phantom, consisting of two steps: the modification of patient datasets into 'printable' file format and the investigation of water to contrast agent concentrations to replicate realistic CT contrast. The feasibility study was concentrated on male pelvis anatomy for simplicity.

Preparation of 3D printed phantom

The planning CT and manually drawn contours (RTS) of a prostate cancer patient were randomly selected and transformed into 'printable' objects through the following steps:

- Import CT and RTS data into 3DSlicer software (v. 4.3.1, www.slicer.org, [160]) using the radiotherapy module [161], then transform the RTS contours into 2D surface layers using the model making module.
- Load 2D surface models into MeshLab software (v. 1.3.2, meshlab.sourceforge.net) and perform 'Poisson surface reconstruction' to transform the 2D surface layers to 3D triangular mesh surfaces.
- Load 3D triangular mesh of each contour to slic3r software (slic3r.org) for the generation of a 'G-code' (i.e. file format containing instructions for the 3D printer).
- Load the generated G-code to the 3D printer (RepRapPro Medel, reprappro.com) and 'print' individual hollow organs.

C.1.3 Results

Figure C.1 shows the original radiotherapy structure set (RTS) of a prostate cancer patient and their transformation into 2D surface layers, using the model making module of Slic3r software. The 2D surface layers were then transformed into hollow 3D triangular meshes surfaces as shown in Fig. C.2.

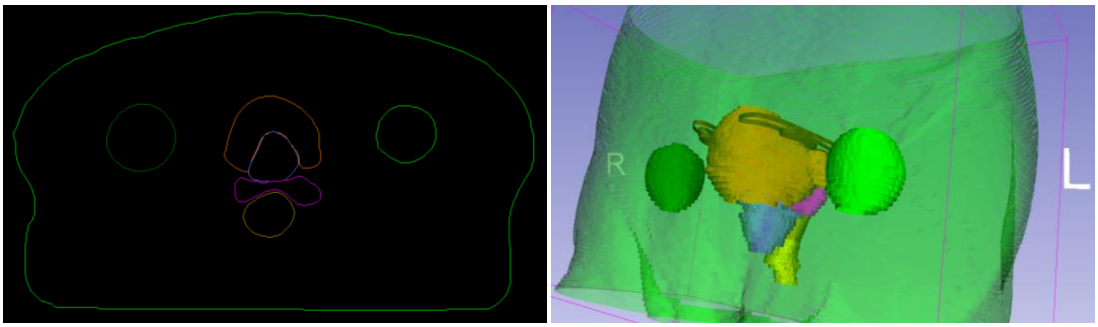

Figure C.1: *Left:* A transverse slice showing the original contours of regions of interest of a prostate cancer patient. *Right:* Original contours transformed into 2D surface layers.

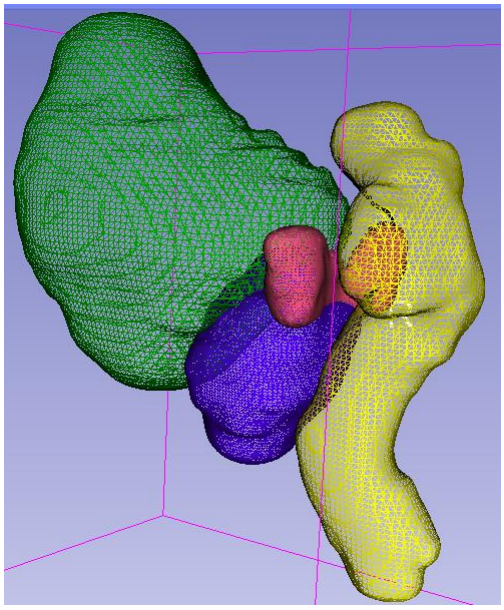

Figure C.2: The prostate patient regions of interest contours transformed into 3D triangular mesh surfaces.

Figure C.3 shows the final 3D printed plastic hollow organs, glued together as seen in the original patient CT scan.

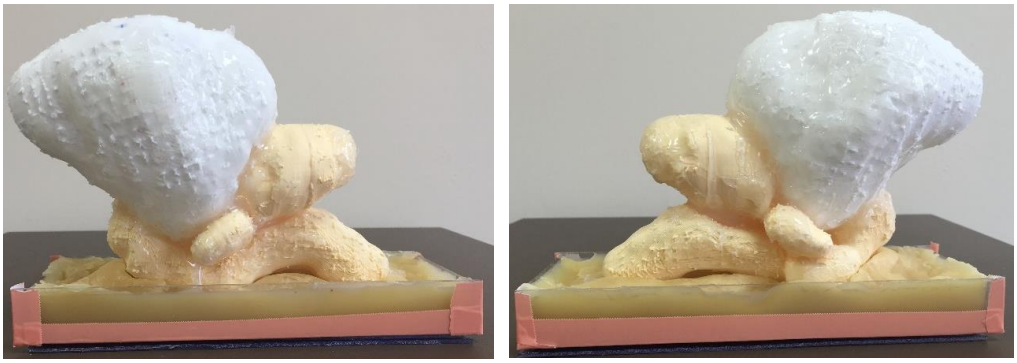

Figure C.3: The patient-specific 3D printed plastic hollow organs, glued together as seen in the original patient CT scan. The lower part shown in the image is a stand made of tissue mimicking wax in order to keep the organs in the upright position.

C.1.4 Discussion and Conclusions

This study investigated the feasibility of creating a patient specific phantom using a 3D printer. The devised workflow allowed for the replication of individual organs of a prostate cancer patient, as outlined the patient's CT scan.

The original intention involved the filling of these organs with different solutions of water to contrast agent concentrations, in an attempt to resemble a simplified patient specific CT scan. This would allow the investigation of various protocols of CT and CBCT imaging, in pursue of optimal workflows for dose calculation on CBCT scans, DIR and deformable dose accumulation. However, the achieved organ replicas were not watertight hindering the continuation of the investigation.

Further investigations could explore the use of 3D printers with different printing materials.

Bibliography

[1] ICRU Report 50-Prescribing, recording and reporting photon beam therapy. International Comminsion on Radiation Units and Measurements; 1994.

[2] ICRU Report 62-Prescribing, recording and reporting photon beam therapy (supplement to ICRU report 50). International Comminsion on Radiation Units and Measurements; 1999.

[3] ICRU Report 83-Prescribing, recording, and reporting intensity-modulated photon-beam therapy (IMRT). vol. 10. International Comminsion on Radiation Units and Measurements; 2010.

[4] De Los Santos J, Popple R, Agazaryan N, Bayouth JE, Bissonnette JP, Bucci MK, et al. Image guided radiation therapy (IGRT) technologies for radiation therapy localization and delivery. Int J Radiat Oncol Biol Phys. 2013;87(1):33–45.

[5] Roussakis YG, Zhang R, Heyes G, Webster GJ, Mason S, Green S, et al. Real-time Cherenkov emission portal imaging during CyberKnife radiotherapy. Phys Med Biol. 2015;60:N419–N425.

[6] Yan D, Vicini F, Wong J, Martinez A. Adaptive radiation therapy. Phys Med Biol. 1997;42(1):123–132.

[7] Yan D. Adaptive radiotherapy: merging principle into clinical practice. Semin Radiat Oncol. 2010;20(2):79–83.

[8] Court LE, Tishler RB, Petit J, Cormack R, Chin L. Automatic online adaptive radiation therapy techniques for targets with significant shape change: a feasibility study. Phys Med Biol. 2006;51(10):2493–501.

[9] National radiotherapy implementation group report. Image guided radiotherapy (IGRT). Guidance for implementation and Use. National Cancer Action Team; 2012.

[10] Wu QJ, Li T, Wu Q, Yin FF. Adaptive radiation therapy: technical components and clinical applications. Cancer J. 2011;17(3):182–9.

[11] Sharp G, Fritscher KD, Pekar V, Peroni M, Shusharina N, Veeraraghavan H, et al. Vision 20/20: perspectives on automated image segmentation for radiotherapy. Med Phys. 2014;41(5):050902.

[12] Hansen EK, Bucci MK, Quivey JM, Weinberg V, Xia P. Repeat CT imaging and replanning during the course of IMRT for head-and-neck cancer. Int J Radiat Oncol, Biol, Phys. 2006;64(2):355–62.

[13] Nuver TT, Hoogeman MS, Remeijer P, van Herk M, Lebesque JV. An adaptive off-line procedure for radiotherapy of prostate cancer. Int J Radiat Oncol, Biol, Phys. 2007;67(5):1559–67.

[14] Miles EA, Clark CH, Urbano MTG, Bidmead M, Dearnaley DP, Harrington KJ, et al. The impact of introducing intensity modulated radiotherapy into routine clinical practice. Radiother Oncol. 2005;77(3):241–6.

[15] Nelms BE, Tomé Wa, Robinson G, Wheeler J. Variations in the contouring of organs at risk: test case from a patient with oropharyngeal cancer. Int J Radiat Oncol, Biol, Phys. 2012;82(1):368–78.

[16] Geraghty JP, Grogan G, Ebert MA. Automatic segmentation of male pelvic anatomy on computed tomography images: a comparison with multiple observers in the context of a multicentre clinical trial. Radiat Oncol. 2013;8:106.

[17] Thomson D, Boylan C, Liptrot T, Aitkenhead A, Lee L, Yap B, et al. Evaluation of an automatic segmentation algorithm for definition of head and neck organs at risk. Radiat Oncol. 2014;9(1):173.

[18] Zhu M, Bzdusek K, Brink C, Eriksen JG, Hansen O, Jensen HA, et al. Multi-institutional quantitative evaluation and clinical validation of smart probabilistic image contouring engine (SPICE) autosegmentation of target structures and normal tissues on computer tomography images in the head and neck, thorax, liver, and male pelvis. Int J Radiat Oncol, Biol, Phys. 2013;87(4):809–16.

[19] La Macchia M, Fellin F, Amichetti M, Cianchetti M, Gianolini S, Paola V, et al. Systematic evaluation of three different commercial software solutions for automatic segmentation for adaptive therapy in head-and-neck, prostate and pleural cancer. Radiat Oncol. 2012;7(1):160.

[20] Greenham S, Dean J, Fu CKK, Goman J, Mulligan J, Tune D, et al. Evaluation of atlas-based auto-segmentation software in prostate cancer patients. J Medi Radiat Sci. 2014;61(3):151–158.

[21] Huyskens DP, Maingon P, Vanuytsel L, Remouchamps V, Roques T, Dubray B, et al. A qualitative and a quantitative analysis of an auto-segmentation module for prostate cancer. Radiother Oncol. 2009;90(3):337–45.

[22] Valentini V, Boldrini L, Damiani A, Muren LP. Recommendations on how to establish evidence from auto-segmentation software in radiotherapy. Radiother Oncol. 2014;112(3):317–320.

[23] Holden M. A review of geometric transftransform for nonrigid body registration. Medical Imaging, IEEE Transactions. 2008;27:111–128.

[24] Thirion JP. Image matching as a diffusion process: an analogy with Maxwell's demons. Med Image Anal. 1998;2(3):243–260.

[25] Christensen GE, Rabbitt RD, Miller MI. Deformable templates using large deformation kinematics. IEEE Trans Image Process. 1996;5(10):1435–47.

[26] Venugopal N, McCurdy B, Hnatov A, Dubey A. A feasibility study to investigate the use of thin-plate splines to account for prostate deformation. Phys Med Biol. 2005;50(12):2871–85.

[27] Rueckert D, Sonoda LI, Hayes C, Hill DL, Leach MO, Hawkes DJ. Nonrigid registration using free-form deformations: application to breast MR images. IEEE transactions on medical imaging. 1999;18(8):712–21.

[28] Sotiras A, Davatzikos C, Paragios N. Deformable medical image registration: a survey. IEEE Trans Med Imag. 2013;32(7):1153–90.

[29] Kirby N, Chuang C, Ueda U, Pouliot J. The need for application-based adaptation of deformable image registration. Med Phys. 2013;40(1):011702.

[30] Kashani R, Hub M, Balter JM, Kessler ML, Dong L, Zhang L, et al. Objective assessment of deformable image registration in radiotherapy: A multi-institution study. Med Phys. 2008;35(12):5944–53.

[31] Bender ET, Tomé Wa. The utilization of consistency metrics for error analysis in deformable image registration. Phys Med Biol. 2009;54(18):5561–77.

[32] Wang H, Dong L, O'Daniel J, Mohan R, Garden AS, Ang KK, et al. Validation of an accelerated 'demons' algorithm for deformable image registration in radiation therapy. Phys Med Biol. 2005;50(12):2887–905.

[33] Zhong H, Kim J, Chetty IJ. Analysis of deformable image registration accuracy using computational modeling. Med Phys. 2010;37(3):970–979.

[34] Fallone BG, Rivest DRC, Riauka TA, Murtha AD. Assessment of a commercially available automatic deformable registration system. J Appl Clin Med Phys. 2010;11(3):101–123.

[35] Kashani R, Hub M, Kessler ML, Balter JM. Technical note: A physical phantom for assessment of accuracy of deformable alignment algorithms. Med Phys. 2007;34(7):2785–88.

[36] Kirby N, Chuang C, Pouliot J. A two-dimensional deformable phantom for quantitatively verifying deformation algorithms. Med Phys. 2011;38(8):4583.

[37] Nithiananthan S, Brock KK, Daly MJ, Chan H, Irish JC, Siewerdsen JH. Demons deformable registration for CBCT-guided procedures in the head and neck: Convergence and accuracy. Med Phys. 2009;36(10):4755–64.

[38] Salguero FJ, Saleh-Sayah NK, Yan C, Siebers JV. Estimation of three-dimensional intrinsic dosimetric uncertainties resulting from using deformable image registration for dose mapping. Med Phys. 2011;38(1):343–353.

[39] Christensen G, Geng X, Kuhl J, Bruss J, Grabowski TJ, Pirwani IA, et al. Introduction to the non-rigid image registration evaluation project (NIREP). vol. 4057 of Lecture Notes in Computer Science. Pluim JPW, Likar B, Gerritsen FA, editors. Springer Berlin Heidelberg; 2006.

[40] Bender ET, Hardcastle N, Tomé WA. On the dosimetric effect and reduction of inverse consistency and transitivity errors in deformable image registration for dose accumulation. Med Phys. 2012;39(1):272–80.

[41] Wang H, Garden AS, Zhang L, Wei X, Ahamad A, Kuban DA, et al. Performance evaluation of an automatic anatomy segmentation algorithm on repeat or four-dimensional CT images using a deformable image registration method. Int J Radiat Oncol, Biol, Phys. 2009;72(1):210–219.

[42] Brock KK. Results of a multi-institution deformable registration accuracy study (MIDRAS). Int J Radiat Oncol, Biol, Phys. 2010;76(2):583–596.

[43] Mencarelli A, Beek SV, Kranen SV, Rasch C, Herk MV, Sonke JJ. Validation of deformable registration in head and neck cancer using analysis of variance. Med Phys. 2012;39(11):6879–84.

[44] Thor M, Petersen JBB, Bentzen L, Hø yer M, Muren LP. Deformable image registration for contour propagation from CT to cone-beam CT scans in radiotherapy of prostate cancer. Acta Oncol. 2011;50(6):918–925.

[45] Liu H, Wu Q. Dosimetric and geometric evaluation of a hybrid strategy of offline adaptive planning and online image guidance for prostate cancer radiotherapy. Phys Med Biol. 2011;56(15):5045–62.

[46] Varadhan R, Karangelis G, Krishnan K, Hui S. A framework for deformable image registration validation in radiotherapy clinical applications. J Appl Clin Med Phys. 2013;14(1):192–213.

[47] Hub M, Thieke C, Kessler ML, Karger CP. A stochastic approach to estimate the uncertainty of dose mapping caused by uncertainties in b-spline registration. Med Phys. 2012;39(4):2186–92.

[48] Murphy MJ, Salguero FJ, Siebers JV, Staub D, Vaman C. A method to estimate the effect of deformable image registration uncertainties on daily dose mapping. Med Image Anal. 2012;39(2):573–80.

[49] Juang T, Das S, Adamovics J. On the need for comprehensive validation of deformable image registration, investigated with a novel 3-dimensional deformable dosimeter. Int J Radiat Oncol, Biol, Phys. 2013;87(2):414–421.

[50] Yeo U, Taylor M, Supple J, Smith R. Is it sensible to "deform" dose? 3D experimental validation of dose-warping. Med Phys. 2012;39(8):5065–72.

[51] Schultheiss T, Tomé WA, Orton CG. Point/counterpoint: it is not appropriate to 'deform' dose along with deformable image registration in adaptive radiotherapy. Med Phys. 2012;39(11):6531–33.

[52] Janssens G, Orban de Xivry J, Fekkes S, Dekker A, Macq B, Lambin P, et al. Evaluation of nonrigid registration models for interfraction dose accumulation in radiotherapy. Medical Physics. 2009;36(9):4268.

[53] Godley A, Ahunbay E, Peng C, Li XA. Accumulating daily-varied dose distributions of prostate radiation therapy with soft-tissue-based kV CT guidance. J Appl Clin Med Phys. 2012;13(3):98–107.

[54] Schwartz DL, Garden AS, Shah SJ, Chronowski G, Sejpal S, Rosenthal DI, et al. Adaptive radiotherapy for head and neck cancer-dosimetric results from a prospective clinical trial. Radiother Oncol. 2013;106(1):80–84.

[55] Akino Y, Yoshioka Y, Fukuda S, Maruoka S, Takahashi Y, Yagi M, et al. Estimation of rectal dose using daily megavoltage cone-beam computed tomography and deformable image registration. Int J Radiat Oncol, Biol, Phys. 2013;87(3):602–8.

[56] Velec M, Moseley JL, Eccles CL, Craig T, Sharpe MB, Dawson La, et al. Effect of breathing motion on radiotherapy dose accumulation in the abdomen using deformable registration. Int J Radiat Oncol, Biol, Phys. 2011;80(1):265–72.

[57] Velec M, Moseley JL, Craig T, Dawson La, Brock KK. Accumulated dose in liver stereotactic body radiotherapy: positioning, breathing, and deformation effects. Int J Radiat Oncol, Biol, Phys. 2012;83(4):1132–40.

[58] Cui Y, W Piper J. Deformable dose accumulation with image guided radiotherapy for final dose evaluation in pelvic cases. J Nucl Med Radiat Ther. 2011;S3:e001.

[59] Schwartz DL, Dong L. Adaptive radiation therapy for head and neck cancer-can an old goal evolve into a new standard? J Oncol. 2011 Jan;2011.

[60] Veiga C, Mcclelland J, Lourenço A, Ricketts K, Annkah J, Souza DD, et al. Toward adaptive radiotherapy for head and neck patients: Feasibility study on using CT-to-CBCT deformable registration for 'dose of the day' calculations. Med Phys. 2014;41(3):031703.

[61] Richter A, Hu Q, Steglich D, Baier K, Wilbert J, Guckenberger M, et al. Investigation of the usability of conebeam CT data sets for dose calculation. Radiat Oncol. 2008;3:42.

[62] Guan H, Dong H. Dose calculation accuracy using cone-beam CT (CBCT) for pelvic adaptive radiotherapy. Phys Med Biol. 2009;54(20):6239–50.

[63] Elstrom UV, Wysocka Ba, Muren LP, Petersen JrBB, Grau C. Daily kV cone-beam CT and deformable image registration as a method for studying dosimetric consequences of anatomic changes in adaptive IMRT of head and neck cancer. Acta Oncol. 2010;49(7):1101–8.

[64] Fotina I, Hopfgartner J, Stock M, Steininger T, Lütgendorf-Caucig C, Georg D. Feasibility of CBCT-based dose calculation: comparative analysis of HU adjustment techniques. Radiother Oncol. 2012;104(2):249–56.

[65] Yang Y, Schreibmann E, Li T, Wang C, Xing L. Evaluation of on-board kV cone beam CT (CBCT)-based dose calculation. Phys Med Biol. 2007;52(3):685–705.

[66] Liu H, Wu Q. Evaluations of an adaptive planning technique incorporating dose feedback in image-guided radiotherapy of prostate cancer. Med Phys. 2011;38(12):6362–70.

[67] Wen N, Kumarasiri A, Nurushev T, Burmeister J, Xing L, Liu D, et al. An assessment of PTV margin based on actual accumulated dose for prostate cancer radiotherapy. Phys Med Biol. 2013;58(21):7733–44.

[68] Hunter KU, Fernandes LL, Vineberg Ka, McShan D, Antonuk AE, Cornwall C, et al. Parotid glands dose-effect relationships based on their actually delivered doses: implications for adaptive replanning in radiation therapy of head-and-neck cancer. Int J Radiat Oncol, Biol, Phys. 2013;87(4):676–82.

[69] Yan C, Hugo G, Salguero FJ, Saleh-Sayah N, Weiss E, Sleeman WC, et al. A method to evaluate dose errors introduced by dose mapping processes for mass conserving deformations. Med Phys. 2012;39(4):2119–28.

[70] Zhong H, Weiss E, Siebers J. Assessment of dose reconstruction errors in image-guided radiation therapy. Phys Med Biol. 2008;53(3):719–736.

[71] Huang TC, Liang JA, Dilling T, Wu TH, Zhang G. Four-dimensional dosimetry validation and study in lung radiotherapy using deformable image registration and Monte Carlo techniques. Radiat Oncol. 2010;5(1):45.

[72] Niu CJ, Foltz WD, Velec M, Moseley JL, Al-Mayah A, Brock KK. A novel technique to enable experimental validation of deformable dose accumulation. Med Phys. 2012;39(2):765–776.

[73] Ahn PH, Chen CC, Ahn AI, Hong L, Scripes PG, Shen J, et al. Adaptive Planning in Intensity-Modulated Radiation Therapy for Head and Neck Cancers: Single-Institution Experience and Clinical Implications. Int J Radiat Oncol Biol Phys. 2011;80(3):677–685.

[74] Wu Q, Chi Y, Chen PY, Krauss DJ, Yan D, Martinez A. Adaptive replanning strategies accounting for shrinkage in head and neck IMRT. Int J Radiat Oncol Biol Phys. 2009;75(1):924–932.

[75] Roussakis YG, McWilliam A, Hartley A, Sangera P, Benghiat H, Hickman M, et al. Evaluation of multiple auto-segmentation solutions against inter-observer variability. Radiother Oncol. 2015;115(Suppl. 1):S507–508.

[76] Pluim JPW, Maintz JBA, Viergever Ma. Mutual-information-based registration of medical images: a survey. IEEE Trans Med Imag. 2003;22(8):986–1004.

[77] Vercauteren T, Pennec X, Perchant A, Ayache N. Diffeomorphic demons: efficient non-parametric image registration. NeuroImage. 2009;45(1 Suppl):S61–72.

[78] Bookstein FL. Principal warps: thin-plate splines and the decomposition of deformations. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1989;11(6):567–585.

[79] Qazi Aa, Pekar V, Kim J, Xie J, Breen SL, Jaffray Da. Auto-segmentation of normal and target structures in head and neck CT images: a feature-driven model-based approach. Med Phys. 2011;38(11):6160–70.

[80] Vik T, Bysstrov D, Schadewaldt N, Schulz H, Peters J. A new method for robust organ positioning in CT images. In: IEEE International Symposium on Biomedical Imaging. vol. 1; 2012. p. 338–341.

[81] Pekar V, McNutt TR, Kaus MR. Automated model-based organ delineation for radiotherapy planning in prostatic region. Int J Radiat Oncol, Biol, Phys. 2004;60(3):973–80.

[82] Akbarzadeh A, Gutierrez D, Baskin A, Ay M, Ahmadian A, Riahi Alam N, et al. Evaluation of whole-body MR to CT deformable image registration. JACMP. 2013;14(4).

[83] Jena R, Kirkby NF, Burton KE, Hoole aCF, Tan LT, Burnet NG. A novel algorithm for the morphometric assessment of radiotherapy treatment planning volumes. Br J Radiol. 2010;83(985):44–51.

[84] Brooks RA, Di Chiro G. Statistical limitations in x-ray reconstructive tomography. Med Phys. 1976;3(4):237–40.

[85] Stutzel J, Oelfke U, Nill S. A quantitative image quality comparison of four different image guided radiotherapy devices. Radiother Oncol. 2008;86:20–24.

[86] Kamath S, Song W, Chvetsov A, Ozawa S, Lu H, Samant S, et al. An image quality comparison study between XVI and OBI CBCT systems. J Appl Clin Med Phys. 2011;12(2):376–89.

[87] Mail N, Moseley DJ, Siewerdsen JH, Jaffray DA. The influence of bowtie filtration on cone-beam CT image quality. Med Phys. 2009;36(1):22–32.

[88] Zhang G, Marshall N, Jacobs R, Liu Q, Bosmans H. Bowtie filtration for dedicated cone beam CT of the head and neck: a simulation study. Br J Radiol. 2013;86:20130002.

[89] Sirinivasan K, Mohammadi M, Shepherd J. Applications of linac-mounted kilovoltage cone-beam computed tomography in modern radiation therapy: a review. Pol J Radiol. 2014;79:181–193.

[90] Ping H, Kandaiya S. The influence of the patient size and geometry on cone beam-computed tomography hounsfield unit. J Med Phys. 2012;37(3):155–158.

[91] Wei Z, Guo-Tao F, Cui-Li S, Yan-Fang W, Cun-Feng W, Da-Quan C, et al. Beam hardening correction for a cone-beam CT system and its effect on spatial resolution. Chin Phys C. 2011;35(10):978.

[92] Jia X, Yan H, Cervino L, Folkerts M, Jiang SB. A GPU tool for efficient, accurate, and realistic simulation of cone beam CT projections. Medical physics. 2012;39(12):7368–78.

[93] Xu Y, Bai T, Yan H, Ouyang L, Pompos A, Wang J, et al. A practical cone-beam CT scatter correction method with optimised Monte Carlo simulations for image-guided radiation therapy. Phys Med Biol. 2015;60:3567–87.

[94] Roussakis YG, Dehghani H, Green S, Webster GJ. A framework for the validation of actual delivered dose estimation strategies. Radiother Oncol. 2014;111(Suppl. 1):S93.

[95] Deasy JO, Blanco AI, Clark VH. CERR: a computational environment for radiotherapy research. Med Phys. 2003;30(5):979–985.

[96] Roussakis YG, Dehghani H, Green S, Webster GJ. Validation of a dose warping algorithm using clclinical realistic scenrarios. Br J Radiol. 2014;88:20140691.

[97] Barker JJ, Garden AS, Kian Ang K, O'Daniel JC, Wang H, Court LE, et al. Quantification of volumetric and geometric changes occuring during fractionated radiotherapy for head-and-neck cancer using an integrated CT/linear accelerator system. Int J Radiat Oncol, Biol, Phys. 2004;59(4):960–970.

[98] Williams T, Sanghera P, Hartley A, Dumbill A, Chalkley A, Roussakis YG, et al. An evaluation of IMPT versus rotational IMRT for nasopharyngeal carcinoma. Radiother Oncol. 2015;114(Suppl. 1):48.

[99] Williams T, Roussakis YG, Sanghera P, Hartley A, Heyes G, Dumbill A, et al. Comparison of intensity modulated proton therapy versus rotational IMRT for nasopharyngeal carcinoma. Radiother Oncol. 2015;115(Suppl. 1):556.

[100] Muller BS, Duma MN, Kampfer S, Nill S, Oelfke U, Geinitz H, et al. Impact of ininterfraction changes in head and neck cancer patients on the delivered dose in intensity modulated radiotherapy with pproton and photons. Phys Med. 2015;31(3):266–272.

[101] Fukumitsu N, Ishikawa H, Ohnishi K, Terunuma T, Mizumoto M, Numajiri H, et al. Dose distribution resulting from changes in aeration of nasal cavity or paranasal sinus cancer in the proton therapy. Radiother Oncol. 2014;113:72–76.

[102] Uzan J, Nahum AE. Radiobiologically guided optimisation of the prescription dose and fractionation scheme in radiotherapy using BioSuite. Br J Radiol. 2012;85:1279–1286.

[103] Eisbruch A, Ten Haken RK, Kim HM, Marsh LH, Ship Ja. Dose, volume, and function relationships in parotid salivary glands following conformal and intensity-modulated irradiation of head and neck cancer. International Journal of Radiation Oncology Biology Physics. 1999;45:577–587.

[104] Rancati T, Schwartz M, Allen AM, Feng F, Popovtzer A, Mittal B, et al. Radiation dose-volume effects in the larynx and pharynx. Int J Radiat Oncol, Biol, Phys. 2010;76(3):S64–9.

[105] Emami B, Lyman J, Brown A, Coia L, Goitein M, Munzenrider JE, et al. Tolerance of normal tissue to therapeutic radiation. Int J Radiat Oncol, Biol, Phys. 2013;21(1):109–22.

[106] White P, Chan KC, Cheng KW, Chan KY, Chau MC. Volumetric intensity-modulated arc therapy vs conventional intensity-modulated radiation therapy in nasopharyngeal carcinoma: a dosimetric study. Journal of Radiation Research. 2013;54:532–545.

[107] Thor M, Bentzen L, Hysing LB, Ekanger C, Helle SI, Karisdottir A, et al. Prediction of rectum and bladder morbility following radiotherapy of prostate cancer based on motion-inclusive dose distributions. Radiother Oncol. 2013;107(2):147–152.

[108] Varadhan R, Hui SK, Way S, Nisi K. Assessing prostate, bladder and rectal doses during image guided radiation therapy–need for plan adaptation? J Appl Clin Med Phys. 2009;10(3):2883.

[109] Michalski JM, Gay H, Jackson A, Tucker SL, Deasy JO. Radiation dose-volume effects in radiation-induced rectal injury. Int J Radiat Oncol Biol Phys. 2010;76(Suppl. 3):S123–9.

[110] de Crevoisier R, Melancon AD, Kuban DA, Lee AK, Cheung RM, Tucker SL, et al. Changes in the pelvic anatomy after an IMRT treatment fraction of prostate cancer. Int J Radiat Oncol Biol Phys. 2007;68(5):1529–1536.

[111] van Herk M, A B, Kroes A, Shouman T, Touw A, Lebesque JV. Quantification of organ motion during conformal radiotherapy of the prostate by three dimensional image registration. Int J Radiat Oncol Biol Phys. 1995;33:1311–20.

[112] Deurloo KEI, Steenbakkers RJHM, Zijp LJ, de Bois Ja, Nowak PJCM, Rasch CRN, et al. Quantification of shape variation of prostate and seminal vesicles during external beam radiotherapy. Int J Radiat Oncol Biol Phys. 2005 Jan;61(1):228–38.

[113] Sveistrup J, Rosenschold P, Deasy J, Oh J, Pommer T, Petersen P, et al. Improvement in toxicity in high risk prostate cancer patients treated with iimage-guided internisty-modulated radiotherapy compared to 3D conformal radiotherapy without daily image guidance. Radiat Oncol. 2014;9.

[114] Yan D, Lockman D, D B, Tyburski L, Martinez A. An off-line strategy for constructing a patient-specific planning target volume in adaptive planning process of prostate cancer. Int J Radiat Oncol Biol Phys. 2000;48:289–302.

[115] Hoogens M, van Herk M, de Bois J, Lebesque JV. Strategies to reduce the systematic error due to tumour and rectum motion in radiotherapy of prostate cancer. Radiother Oncol. 2005;74:177–185.

[116] Nijkamp J, Pos FJ, Nuver TT, de Jong R, Remeijer P, Sonke JJ, et al. Adaptive radiotherapy for prostate cancer using kilovoltage cone-beam computed tomography: first clinical results. Int J Radiat Oncol Biol Phys. 2008;70(1):75–82.

[117] Dearnaley D, Syndikus I, Sumo G, Bidmead M, Bloomfield D, Clark C, et al. Conventional versus hypofractionated high-dose intensity modulated radiotherapy for prostate cancer: preliminary safety results from CHHiP randomised concontrol trial. Lancet Oncol. 2012;13(1):43–54.

[118] Lu S, Cheng J, Kuo S, Lee J, Chen L, Wu J, et al. Volumetric modulated arc therapy for nasopharyngeal carcinoma: a dosimetric comparison with TomoTherapy and step-and-shoot IMRT. Radiother Oncol. 2012;104(3):324–30.

[119] Wang ZH, Yan C, Zhang ZY, Zhang CP, Hu HS, Kirwan J, et al. Radiation-induced volume changes in parotid and submandibular glands in patients with head and neck cancer receiving postoperative radiotherapy: a longitudinal study. Laryngoscope. 2009;119(10):1966–74.

[120] Castadot P, Geets X, Lee JA, Christian N, Gregoire V. Assessment by deformable registration method of the volumetric and positional changes of target volumes and organs at risk on pharyngo-laryngeal ttumor treated with concomitant chemo-radiation. Radiother Oncol. 2010;95(2):209–17.

[121] Lee C, Langen KM, Lu W, Haimeri J, Schanarr E, Ruchala KJ, et al. Evaluation of geometric changes of parotid glands during head and neck cancer radiotherapy using daily MVCT and automatic deformable registration. Radiother Oncol. 2008;89(1):81–8.

[122] O'Daniel JC, Garden AS, Schwartz DL, Wang H, Ang KK, Ahamad A, et al. Parotid gland dose in head-and-neck intensity-modulated radiotherapy: is what you plan what you get? Int J Radiat Oncol, Biol, Phys. 2007;69(4):1290–1296.

[123] Beltran M, Ramos M, Rovira JJ, Perez-Hoyos S, Sancho M, Puertas E, et al. Dose variations in tumor volumes and organs at risk during IMRT for head-and-neck cancer. J Appl Clin Med Phys. 2012 Jan;13(6):3723.

[124] Han C, Chen YJ, Liu A, Schultheiss T, Wong JY. Actual dose variation of parotid glands and spinal cord for nasopharyngeal cancer patients during radiotherapy. Int J Radiat Oncol Biol Phys. 2008;70(4):1256–62.

[125] Bhide SA, Davies M, Burke K, McNair HA, Hansen V, Barbachano Y, et al. Weekly volume and dosimetric changes during chemoradiotherapy with intensity-modulated radiation therapy for head and neck cancer: a prospective observational study. Int J Radiat Oncol Biol Phys. 2010;76(5):1360–8.

[126] Gastadot P, Lee JA, Geets X, Gregoire V. Adaptive radiotherapy of head and neck cancer. Semin Radiat Oncol. 2010;20:84–93.

[127] Schwartz DL. Current progress in adaptive radiation therapy for head and neck cancer. Curr Oncol Rep. 2012 Apr;14(2):139–47.

[128] Lagendijk JJW, Raaymakers BW, Raaijmakers AJE, Overweg J, Brown KJ, Kerkhof EM, et al. MRI/linac integration. Radiother Oncol. 2008;86(1):25–29.

[129] Raaymakers BW, Lagendijk JJW, Overweg J, Kok JGM, Raaijmakers AJE, Kerkhof EM, et al. Integrating a 1.5 T MRI scanner with 6 MV accelerator: proof of concept. Phys Med Biol. 2009;54:N229–N237.

[130] Lagendijk JJW, Raaymakers BW, van Vulpen M. The magnetic resonance imaging-linac system. Semin Radiat Oncol. 2014;24:207–209.

[131] Keall PJ, Barton M, Crozier S. The Australian magnetic resonance imaging-linac program. Semin Radiat Oncol. 2014;24:203–206.

[132] Mutic S, Dempsey F. The ViewRay system: magnetic resonance-guided and concontrol radiotherapy. Semin Radiat Oncol. 2014;24(3):196–9.

[133] Lagendijk JJW, Raaymakers BW, van der Berg CAT, Moerland MA, Philippens ME, van Vulpen M. MR guidance in radiotherapy. Phys Med Biol. 2014;59:R349–R369.

[134] Li X, Quan EM, Li Y, Pan X, Zhou Y, Wang X, et al. A fully automated method for CT-on-rails-guided online adaptive replanning for prostate cancer intensity modulated radiotherapy. Int J Radiat Oncol Biol Phys. 2013;86(5):835–841.

[135] Li T, Wu Q, Zhang Y, Vergalasova I, Lee WR, Yin FF, et al. Strategies for automatic online treatment plan reoptimization using clinical treatment planning system: a planning parameters study. Med Phys. 2013;40(11):111711.

[136] Webster GJ, Stratford J, Rodgers J, Livsey JE, Macintosh D, Choudhury a. Comparison of adaptive radiotherapy techniques for the treatment of bladder cancer. Br J Radiol. 2013;86(1021):20120433.

[137] Heijkoop ST, Langerak TR, Quint S, Bondar L, Means JW, Heijmen BJ, et al. Clinical implementation of an online adaptive plan-of-the-day protocol for nonrigid motion management in locally advanced cervical cancer IMRT. Int J Radiat Oncol Biol Phys. 2014;90(3):673–9.

[138] McNair HA, Hafeez S, Taylor H, Lalondrelle S, McDonald F, Hansen VN, et al. Radiographer-led plan selection for bladder cancer radiotherapy: initiating a trining prprogram and maintaining competency. Br J Radiol. 2015;88(1048):20140690.

[139] Boejen A, Vestergaard A, Hoffmann L, Ellegaard MB, Rasmussen AM, Moller D, et al. A learning prprogram qualifying radiation therapists to manage daily online adaptive radiotherapy. Acta Oncol. 2015;54(9):1697–1701.

[140] Langerak T, Heijkoop S, Quint S, Mens JW, Heijmen B, Hoogeman M. Toward automatic plan selection for radiotherapy of cervical cancer by fast automatic segmentation of cone beam CT scans. Med Image Comput Comput Assist Interv. 2014;17(Pt 1):528–35.

[141] Lewis R, Hall E, Griffin C, Hafeez S, Huddart R. Current UK practice in organ sparing treatment of muscle invasive bladder cancer (MIBC) and impact on design of the RAIDER image guided radiotherapy (IGRT) trial; Presented at: 10th National Cancer Research Institute Cancer Conference. 2014. Available from: http://conference.ncri.org.uk/abstracts/2014/abstracts/B036.html.

[142] Berwouts D, Olteanu LA, Duprez F, Vercauteren T, De Gersem W, De Neve W, et al. Three-phase adaptive dose-painting-by-numbers for head-and-neck cancer: initial results of the phase I clinical trial. Radiother Oncol. 2013;107(3):310–6.

[143] Olteanu LA, Berwouts D, Madani I, De Gersem W, Vercauteren T, Duprez F, et al. Comparative dosimetry of three-phase adaptive and non-adaptive dose-painting IMRT for head-and-neck cancer. Radiother Oncol. 2014;111(3):348–53.

[144] Feng M, Kong FM, Gross M, Fernando S, Hayman JA, Ten Haken RK. Using fluorodeoxyglucose positron emission tomography to assess tumor volume during radiotherapy for non-small-cell lung cancer and its potential impact on adaptive dose escalation and normal tissue sparing. Int J Radiat Oncol Biol Phys. 2009;73(4):1228–34.

[145] Servagi-Vernat S, Differding S, Sterpin E, Hanin FX, Labar D, Bol A, et al. Hypoxia-guided adaptive radiation dose escalation in head and neck carcinoma: a planning study. Acta Oncol. 2015;54(7):1008–16.

[146] Cerenkov PA. Visible emission of clean liquids be action of gamma radiation. C R Dokl Acad Sci SSSR. 1934;2:451–4.

[147] Glaser AK, Zhang R, Gladstone DJ. Optical dosimetry of radiotherapy beams using Cherenkov radiation: the relationship between light emission and dose. Phys Med Biol. 2014;59:3789–3811.

[148] Mei X, Rowlands JA, Pang G. Electronic portal imaging based on Cerenkov radiation: a new approach and its feasibility. Med Phys. 2006;33(11):4258–70.

[149] Zhang R, Fox CJ, Glaser AK. Superficial dosimetry imaging of Cerenkov emission in electron beam radiotherapy of phantoms. Phys Med Biol. 2013;58:5477–5493.

[150] Glaser AK, Davis SC, Mcclatchy DM, Zhang R, Pogue BW, Gladstone DJ. Projection imaging of photon beams by the Cerenkov effect. Med Phys. 2013;40(1):012101.

[151] Glaser AK, Voigt WHA, Davis SC, Zhang R, Gladstone DJ, Pogue BW. Three-dimensional Cerenkov tomography of energy deposition from ionizing radiation beams. Opt Lett. 2013;38(5):634–636.

[152] Jang KW, Yoo WJ, Moon J, Han KT, Park Jy, Lee B. Measurements of relative depth doses and Cerenkov light using a scintillating fiber-optic dosimeter with Co-60 radiotherapy source. Applied Radiation and Isotopes. 2012;70:274–277.

[153] Helo Y, Rosenberg I, Souza DD, Macdonald L, Speller R, Royle G, et al. Imaging Cerenkov emission as a quality assurance tool in electron radiotherapy. Phys Med Biol. 2014;59:1963–78.

[154] Helo Y, Kacperek A, Rosenberg I, Royle G, Gibson A. The physics of Cerenkov light production during proton therapy. Phys Med Biol. 2014;59(23):7107–7123.

[155] Zhang R, Gladstone DJ, Jarvis La, Strawbridge RR, Jack Hoopes P, Friedman OD, et al. Real-time in vivo Cherenkoscopy imaging during external beam radiation therapy. J Biomed Opt. 2013;18(11):110504.

[156] Jarvis LA, Zhang R, Gladstone DJ, Jiang S, Hitchcock W, Friedman OD, et al. Cherenkov video imaging allows for the first visualization of radiation therapy in real time. Int J Radiat Oncol, Biol, Phys. 2014;89(3):615–622.

[157] Glaser AK, Zhang R, Davis SC, Gladstone DJ, Pogue BW. Time-gated Cherenkov emission spectroscopy from linear accelerator irradiation of tissue phantoms. Opt Lett. 2012;37(7):1193–1195.

[158] Roussakis YG, Zhang R, Heyes G, Webster GJ, Suzannah M, Green S, et al.. Supplementary material for: Real-time Cherenkov emission portal imaging during CyberKnife radiotherapy; 2015. Available from: http://iopscience.iop.org/0031-9155/60/22/N419/media/PMB520655_lead_mov.mp4.

[159] Sarrut D. Deformable registration for image-guided radiation therapy. Z Med Phys. 2006;16(4):285–97.

[160] Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, et al. 3D Slicer as an image computing platform for quantitative imaging network. Magn Reson Imag. 2012;30(9):1323–41.

[161] Pinter C, Lasso A, Wang A, Jaffray D, Fichtinger G. SlicerRT: radiation therapy research tootool for 3D Slicer. Med Phys. 2012;39(10):6332–38.