Intensity-Modulated Radiotherapy for the Head and Neck: Improvements in Planning and Delivery

A thesis submitted to the University of Manchester for the degree of

PhD

in the Faculty of Medical and Human Sciences

2009

Gareth John Webster

School of Medicine, Division of Cancer Studies

Table of Contents

Word count = approximately 42,000 words

Submission of thesis in an alternative format

Due to the full-time professional clinical responsibilities of the author, permission to submit this thesis in an alternative format has been sought and attained.

As a result, the structure of this thesis is such that all work submitted for publication in peer-reviewed journals (as listed on p7) are contained within Chapter 2 in the format in which they were submitted (edited to conform to the submission guidelines specified by the University of Manchester). Detailed methodology and results are provided within the publications themselves and so, to avoid repetition, only the key aspects of each have been described in Chapters 1 and 5 respectively.

List of Figures & Tables

Figure 3.3. Flowchart showing corrections applied to the in-air fluence by the TPS to calculate dose distribution. Steps highlighted in grey are not applied in 'intensity modulation' mode

Figure 3.4. Profile overlays of dosemap calculation at d_{max} in solid water slab phantom of 10x10cm flat in-air fluence with (black) and without (red) corrections for a) jaw and MLC transmission, b) penumbral blurring due to focal spot size

Figure 3.5. Profile overlays of dosemap calculation at d_{max} in solid water slab phantom of 10x10cm flat in-air fluence with (black) and without (red) corrections for a) electron contamination, b) flattening filter attenuation and c) head scatter

Figure 3.6. Central-axis profile of a) calculated dosemap (red) resulting from flat 10x10cm in-air fluence (black). b) shows the deviation of the resulting profile from the intended function

Figure 3.7. a) gradient mask over which correction is applied, b) assignation of intensity within gradient regions based on relative proximity to high and low intensity plateaus

Figure 3.8. a) a simple step-function b) is convolved and then deconvolved using the same Gaussian kernel

Figure 3.9. gamma analyses (2%/2mm, grey regions indicate regions exceeding tolerance) and comparison of profiles along red lines for recalculated dosemap (red) with original measured film fluence (black) based on optimised in-air fluence. These images highlight discrepancies in penumbral regions.

Figure 4.1. Flowchart showing function of risk calculation code

Figure 4.2. Dose-response curves used in this study: Model A assumes that cell-killing quickly becomes the dominant process; Model B assumes a higher impact from cell recovery and repopulation effects, resulting in a plateau of risk; Model C assumes that cell-killing is a less important process and assumes increasing risk with dose, albeit at a lower rate than the <4Gy region. A rounded DDREF factor of 2 is used

Figure 4.3. Variation of a) age factor A and b) gender factor G as a function of age. Data taken from BEIR VII

Figure 4.4. Relative organ sensitivity variation with age for a) male and b&c) females (Figure 4.4c is at a different scale to better visualise the trends for less sensitive tissues). Data taken from BEIR VII and modified as illustrated in Table 4.1

Figure 4.5. Absolute risk values for IMRT and 3D-conformal treatment techniques for the different dose-response curves

Figure 4.6. Risk distributions from head and neck IMRT treatment for different models (risk at any given point corresponds to the absolute cumulative risk if this were present throughout the entire dataset). 3D conformal treatment showed a similar distribution

Table 4.1. Calculation methodology for relative organ sensitivities from incidence data given in BEIR VII for a 5 year old male

List of Abbreviations

3D-CRT	3-dimensional conformal radiotherapy
BEIR	Committee on the biological effects of ionising radiation
CCC	Collapsed cone convolution
CT	Computed tomography
CW	Clockwise
DMPO	Direct machine parameter optimisation
DVH	Dose-volume histogram
EDR	Extended dose range
EPID	Electronic portal imaging device
FDG	fluoro-deoxyglucose
HU	Hounsfield unit
IAEA	International Atomic Energy Agency
IGRT	Image-guided radiotherapy
IMRT	Intensity-modulated radiotherapy
IPEM	Institute of physics & engineering in medicine
kV	Kilovoltage
MLC	Multi-leaf collimator
MRI	Magnetic resonance imaging
MU	Monitor units
MV	Megavoltage
NTCP	Normal tissue complication probability
PET	Positron emission tomography
QA	Quality assurance
RF	Radiofrequency
RPC	Radiological Physics centre
RTP	Radiotherapy planning
SMG	Submandibular gland
TCP	Tumour control probability
TPS	Treatment planning system
VMAT	Volumetric Intensity-modulated arc therapy

List of Publications

Publication #1: A novel imaging technique for fusion of high quality, immobilised magnetic resonance (MR) images of the head and neck with computed tomography (CT) scans for radiotherapy target delineation

BJR, 82 (2009), p497-503

Publication #2: Development of an optimum photon beam model for head and neck IMRT

JACMP, 8(4), 2007 p129-138

Publication #3: Evaluation of the impact of dental artefacts on intensity-modulated radiotherapy planning for the head and neck

Accepted for publication in Radiotherapy and Oncology

Publication #4: Evaluation of larynx sparing techniques with IMRT when treating the head and neck

IJROBP, 72(2), 2008 p617-622

Publication #5: Design and implementation of a head and neck phantom for system audit and verification of intensity-modulated radiation therapy

JACMP, 9(2), 2008 p46-56

Publication #6: Accuracy and precision of an IGRT solution

Med Dosim, 34(2), 2009, p99-106

PhD Abstract

Intensity-Modulated Radiotherapy for the Head and Neck: Improvements in Planning and Delivery

Gareth John Webster
University of Manchester, September 2009

Introduction
Clinical implementation of intensity-modulated radiotherapy (IMRT) for the head and neck is becoming increasingly widespread due to the expected benefits to local tumour control and reduction of associated toxicities. The exploitation of the full benefits of IMRT is currently not well prescribed.

Methodology
A series of investigations into aspects of the radiotherapy treatment process have been carried out that aim to improve the clinical outcomes associated with IMRT for the head and neck, either by quantifying the achievable accuracy and precision or by highlighting improvements to current practice.

Results
Delineation: a novel imaging technique allows fusion of magnetic resonance images with the radiotherapy planning scan with minimal distortion (<1mm to 8.7cm from the scan centre).

Dose calculation: the accuracy of the collapsed cone convolution algorithm is adequate for head and neck IMRT. Accurate modelling of rounded leaf ends, which can reduce localised errors by up to 3%, is recommended.

Planning: simple solutions are recommended for practical problems in the IMRT planning process, while a methodology has been developed that appears to confirm no significant extra risk of secondary malignancies from the use of IMRT for the head and neck (<0.2% extra risk).

Delivery: accuracy of treatment delivery to within ±3% in absolute dose has been demonstrated, including within beam penumbra, using a custom-made phantom. Development of a technique to calculate the delivered dose directly to the patient dataset is ongoing.

Conclusion
Implementation of IMRT for the head and neck is associated with significant benefits to patients. Widespread implementation should proceed cautiously with simple solutions being sought for practical problems encountered.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright statement

Acknowledgements

I would like to acknowledge the contributions of Carl Rowbottom and Ranald Mackay, the supervisors of this thesis, who have offered much-needed guidance, reassurance and assistance throughout the last four years and without whose help this work would not have been completed.

I am also grateful to my other colleagues within North Western Medical Physics who have often provided invaluable assistance with issues that have arisen throughout the course of this work, in particular Steve Smith, for papering over the substantial cracks in my computer-programming abilities.

Finally, I am most grateful to my parents for their unquestioning belief, encouragement and support throughout the past thirty years. Thank you.

The Author

MPhys, Physics, University of Manchester, 2001
MSc, Medical Physics & Clinical Engineering, University of Sheffield, 2002

The Author has since worked at the Christie Hospital NHS Foundation Trust, Manchester, completing the IPEM Parts I & II clinical scientist training in January 2004 and September 2008 respectively. The author has subsequently worked as a HPC-registered clinical scientist in Radiotherapy.

Other research experience includes the development and implementation of several local and national clinical trials and the development of local treatment techniques.

1.1 Head and neck cancer

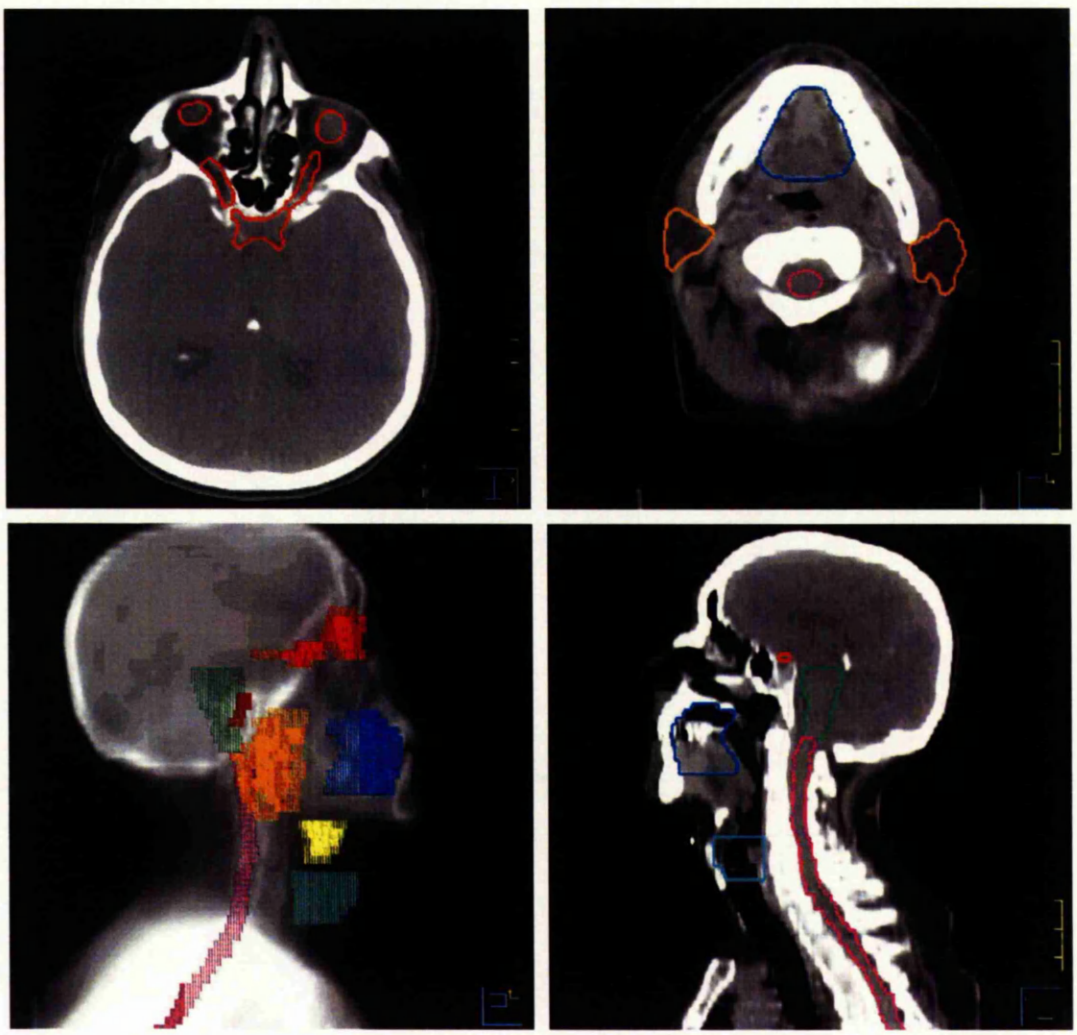


Figure 1.1. Illustration of relevant anatomy in treatment of head and neck showing brainstem (dark green), spinal cord (purple), optic structures (red), parotid glands (orange), submandibular glands (yellow), oral cavity (blue), larynx (sky blue), cochlea (brown)

Figures compiled by Cancer Research UK for 2006-2007 suggest that approximately 7800 patients are diagnosed with head and neck cancer (defined here as cancers of the larynx, oral cavity and nasopharynx) in the UK each year, leading to approximately 2,700 deaths. Most of these treatments will involve radiotherapy as either part or all of their treatment. The

aim of radiotherapy is to achieve local control of disease by the killing of malignant cells. This is done by exploiting the sensitivity of tumour cells to radiation, relative to healthy tissue. In head and neck, curative radiotherapy courses can lead to acute and late reactions that affect the quality of life for patients surviving radiotherapy. The amount of radiation that can safely be given to malignant cells is limited by the radiosensitivity of these surrounding normal tissue structures.

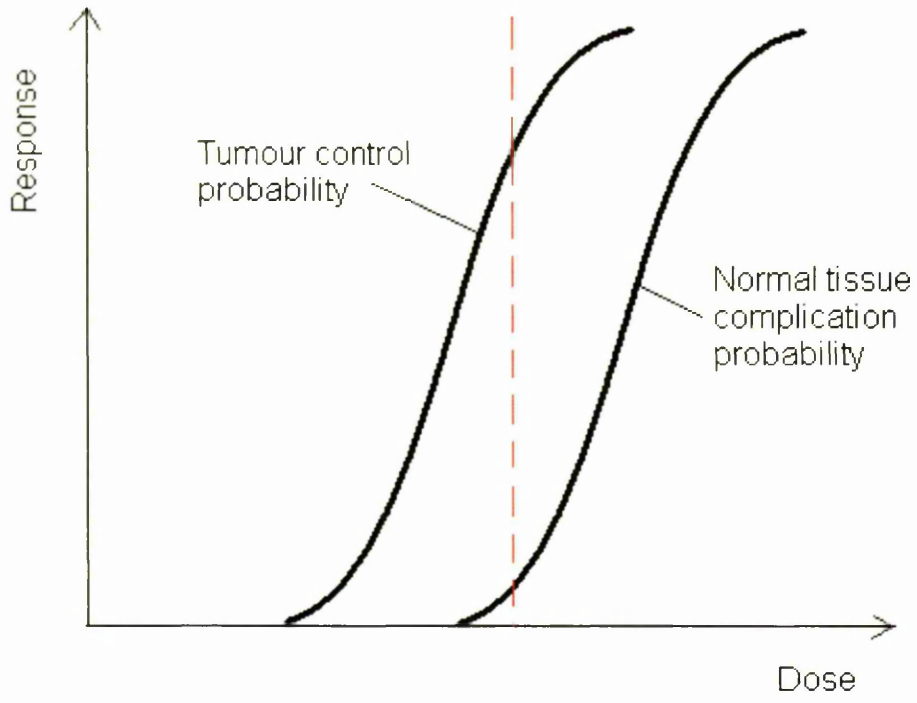


Figure 1.2. Illustration of the therapeutic window: when selecting the optimal treatment dose there is a trade-off vetween the probability of tumour control and the frequency and extent of normal tissue toxicity

Tissues of particular concern in head and neck cancer are illustrated in Fig 1.1: irradiation of the parotid and submandibular glands (SMG) can lead to temporary or permanent loss of salivary function; irradiation of the oral cavity can lead to temporary or permanent oral mucositis; high superficial doses can cause acute skin reactions; irradiation of the larynx can lead to dysphagia and loss of vocal capacity; partial or complete blindness can result from irradiation of the optic chiasm and optic nerves; cataracts can be caused by irradiation of the lens; irradiation of the cochlear and surrounding auditory structures can lead to loss of

hearing; irradiation of a significant volume of the cerebellum can result in ataxia. Depending upon the individual case and the likely severity of the reaction, each of these toxicities may be an accepted compromise if it is necessary to achieve tumour control. Some of these effects can be reduced by irradiating only part of the structure or by reducing the delivered dose and the extent of any reactions can vary considerably between patients due to factors such as inherent radiosensitivity, lifestyle and underlying pathologies. However, the most critical structures in the head and neck are the serial structures of the spinal cord and brainstem. Exceeding the tolerance dose for these structures could result in myelopathy and is therefore the limiting factor in the amount of dose that can be delivered to the tumour.

There is evidence that head and neck tumour control probability (TCP) could be increased by escalating the prescribed dose to the tumour, at a rate of approximately 5% for every 5Gy escalation [1]. Several studies of patterns of failure for the head and neck reinforce this, reporting that most failures occur in the high dose region, which suggests that dose escalation would be beneficial [2-5]. However, Fig 1.2 illustrates the problem encountered when trying to exploit this possibility: the therapeutic window, at which tumour control can be achieved without unacceptably increasing the normal tissue complication probability (NTCP), is narrow and any attempt to prescribe a higher dose to the target would increase the likelihood of unacceptable toxicity. Ideally, dose escalation to the target would be achieved while maintaining the dose to the critical structures at a safe level. An alternative approach may be to reduce toxicity while maintaining current levels of local control.

1.2 Treatment techniques in head and neck radiotherapy

Recently established radiotherapy techniques such as intensity-modulated radiotherapy (IMRT) and image-guided radiotherapy (IGRT) may allow further exploitation of the therapeutic window and could be of significant benefit, both in terms of disease control and reduced toxicity, for many patients. The following will explain the advantages and limitations of the currently established and widely used 3D conformal technique, before evaluating the rationale for both IMRT and IGRT.

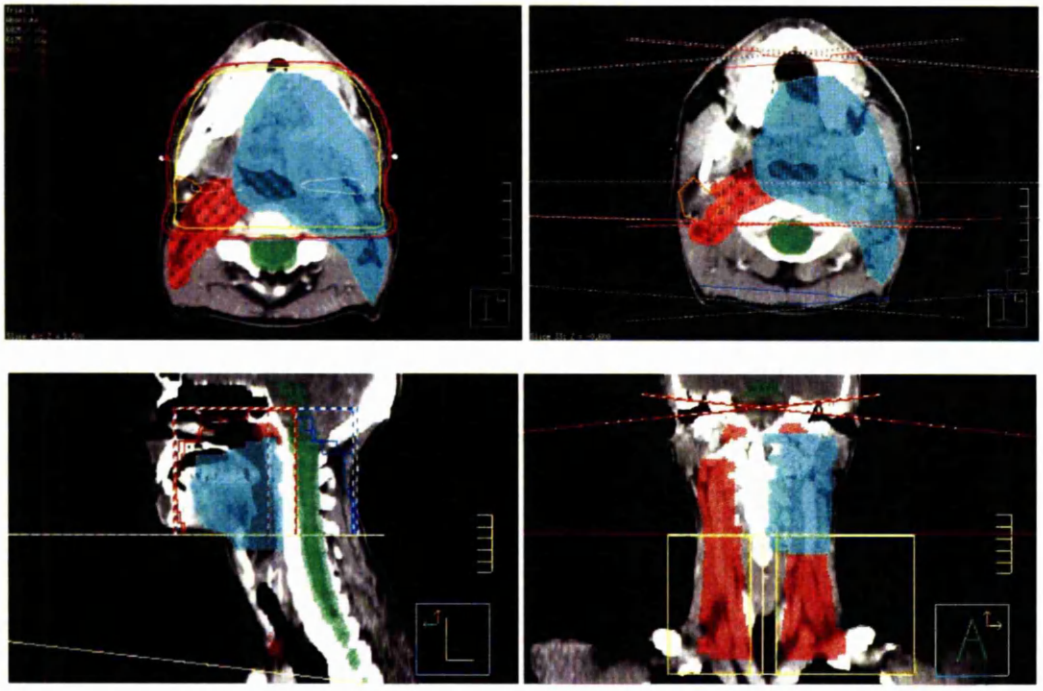


Figure 1.3. Conventional parallel-opposed 3D conformal technique for head and neck radiotherapy. Blue and red colourwash volumes represent PTV1 and PTV2 respectively. Top left: Heterogeneously calculated axial dose distribution (dose to posterior nodes would be increased using electron fields, the impact of which is not shown here). Other images show the first (blue) and second (red) phase lateral fields and the matched anterior neck field (yellow) for axial, coronal and sagittal slices respectively (CW from top right)

A conventional 3D conformal treatment technique for the head and neck is illustrated in Fig 1.3 and is the standard local technique used to treat oropharyngeal cancers. It is a 2-phase treatment consisting of parallel-opposed pairs of photon beams with multi-leaf collimators (MLC) used to shape the beam fluence. The first phase covers the whole of the superior part of the target, while the posterior jaws of the second phase are pulled anteriorly to avoid the spinal cord and brainstem. The dose to the target volume lateral to the spinal cord is then boosted using electron fields, while the inferior part of the target is treated using a matched anterior neck field with spinal blocking along the central axis. This technique has the advantage of simplicity, proven safety with regard to acceptable levels of toxicity, generally good coverage of the primary site of disease, delivers low dose to the oral cavity and is associated with low rates of dysphagia due to the sparing of the larynx from the spinal

blocking. However, the scope for dose escalation is limited due to the large volume of normal tissue being irradiated. It can also fail to cover the entire target volume, particularly in, for example, late stage nasopharyngeal cases, in which the disease can approach very close to the brainstem and optic structures, or for elective nodal volumes in the neck due to the medial blocking on the anterior neck field [3]. It also delivers very high doses to both parotid glands, with an associated high degree of permanent loss of function, with the PARSPORT trial observing ≥Grade 2 xerostomia in 74% of patients for the conformal arm of the trial (LENT-SOMA scale) [6]

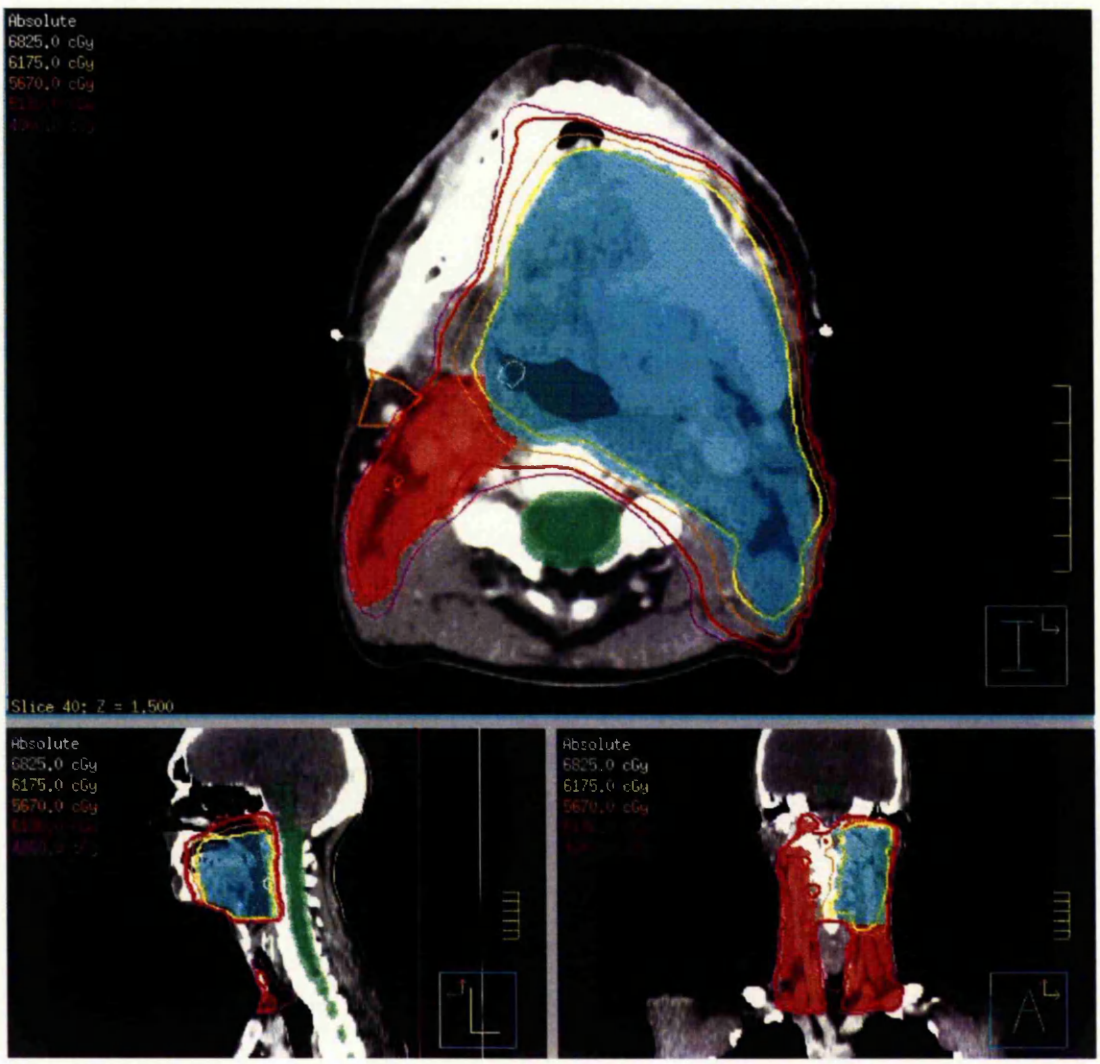


Figure 1.4. Dose distribution obtained with IMRT for the head and neck showing reduced dose to the contralateral parotid

IMRT uses a non-uniform beam fluence to conform the dose around target volumes and away from sensitive structures. This is achieved either by delivering the beam as a series of static beams (step-and-shoot IMRT) or by moving the MLCs across the field at varying speeds while the radiation is being delivered (dynamic IMRT). Both of these techniques are inverse-planned, using an initial set of dose constraints to define a desired solution, such as that shown in Fig 1.4. IMRT can achieve highly conformal dose distributions, delivering high and homogenous doses to the target volume while sparing normal tissues.

Several UK centres have implemented IMRT into routine clinical practice via enrolment in the PARSPORT study, a phase III randomised trial investigating the role of IMRT in reducing ≥Grade 2 xerostomia in head and neck cancer patients [6]. This study has demonstrated significantly reduced rates of xerostomia after treatment with IMRT, relative to conventional conformal techniques, with no observed changes in survival or local control, although acute fatigue was observed to increase, possibly as a result of the increased volume of brain irradiated in the IMRT arm. However, there are several risks inherently associated with the technique that should be considered: the increased conformality means that there is increased sensitivity to errors in patient positioning or delineation, which could lead to loss of local control via peripheral recurrences [4]; the proximity of high dose regions to critical structures means that greater accuracy is required in dose calculation and radiation delivery than has been required for conventional techniques; the parameters of the treatment (i.e. monitor units, beam directions, etc) are often non-intuitive and so errors are less easy to detect, requiring more thorough verification procedures. There are also practical concerns over the implementation of IMRT: increased laryngeal doses due to the loss of conventional spinal cord blocking have led to significant increases in dysphagia [7]; the use of anterior beams increases the dose to the oral cavity, which could cause increased levels of mucositis; increased levels of low dose to the patient, both through changes to the dose distribution and the increased monitor units, could lead to an associated increased risk of secondary malignancies [8].

A recent survey carried out by the Royal College of Radiologists quantified the use of IMRT and IGRT throughout the UK. It found that only 27% of centres use IMRT as part of the routine clinical service, although only 33% of centres did not intend to use IMRT within the next 12 months either in a clinical or research capacity. This, together with the most cited reason for not implementing IMRT being a lack of physics expertise rather than a lack of motivation, suggests that the technique is ultimately likely to be in widespread use throughout the UK. The most popular site for which IMRT is used is the head and neck (35%), with over 50% of centres intending this service to be available within 12 months [9].

The use of volumetric cone beam imaging at the treatment machine is increasingly being incorporated into linear accelerators, raising many possibilities in the fields of image-guided and adaptive radiotherapy. This facility could overcome some of the problems associated with IMRT in terms of the precise positioning of patients during treatment, sensitivity to changes in patient and target shape throughout the full treatment course and the assessment of the impact of delivery errors on the dose distribution delivered to the patient.

1.3 The radiotherapy treatment process

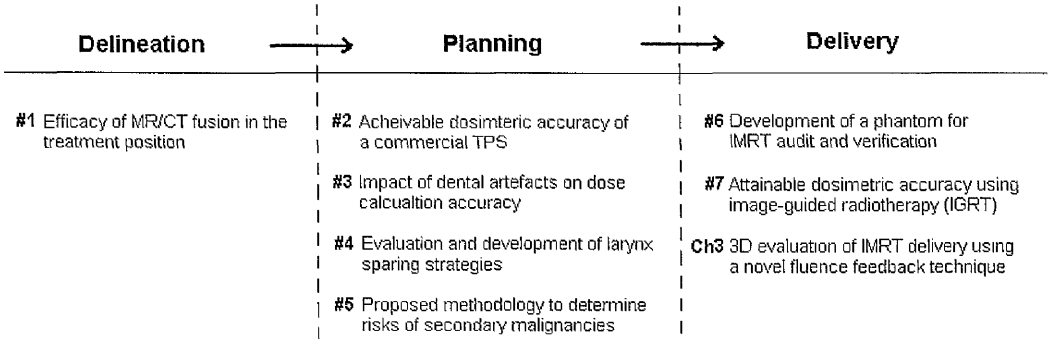


Figure 1.5. Structure of this thesis in relation to the radiotherapy treatment chain.

The radiotherapy treatment chain consists of several links. Those relating directly to the technical application of radiotherapy are the focus of this thesis and are illustrated at the top of Fig 1.5. The full potential of advanced treatment techniques cannot be realised if

significant errors are generated at any stage in this process. The head and neck represents a particularly complex environment for radiotherapy for several reasons: the proximity of critical structures to the target volume; the large concave volumes that are routinely treated, ideally to different dose levels; the presence of a wide range of heterogeneities and the changing surface contour. It is precisely because of the associated increased difficulty in treatment that the head and neck stands to benefit most from the safe implementation of advanced techniques. This work aims to strengthen each major link of the radiotherapy chain in order to establish the limits of possible improvements when using IMRT in head and neck radiotherapy.

1.4 Delineation

Any error introduced into the treatment process at the imaging or delineation stage will be systematic, propagating through the entire treatment and, in the case of radiotherapy, affecting every delivered fraction. Computed tomography (CT) scans are a prerequisite to planning treatments in heterogeneous regions of the body such as the head and neck, as they directly represent the electron density information for accurate dose calculation. However, despite this intrinsic advantage, the modality has limitations: soft tissue contrast (e.g. between tumour and muscle or fat) is very poor, often leading to difficulty in precisely determining the extent, or even presence, of tumour tissue; the cranio-caudal resolution in the sagittal or coronal planes is often limited to 3-5mm due to the limitations of the CT scanner, which can lead to uncertainty in defining tumour extension or, in the particular case of nasopharyngeal cancers, identifying invasion into the skull.

Other modalities have therefore been introduced into the staging and delineation of head and neck disease to complement the information provided by CT. One such modality is positron-emission tomography (PET) using 18F-fluoro-deoxyglucose (FDG) tracer that is taken up into cells by glucose-receptors, which are over-expressed in tumour cells [10; 11]. This functional imaging FDG-PET imaging technique permits visualisation of malignant tissue and is in widespread use for staging of disease, including identification of lymph node

involvement and metastases. Increasingly, studies aiming to improve rates of local control are using IMRT to deliver boost doses to PET-avid volumes, demonstrating that the dose distribution in other areas of the plan need not be affected [12] and that such targeted dose escalation is well tolerated by patients [13], with Soto et al demonstrating that most (although not all) loco-regional failures occur within the pre-treatment PET-avid GTV [5]. However, PET scanning is associated with limitations in terms of both specificity, due to high tracer uptake in any active musculature and normal tissues such as the heart and brain, and sensitivity, since not all malignant tissue will be associated with high glucose uptake. The latter problem carries a risk of geometric miss is PET were to be relied on as a sole guide to target delineation. A recent IAEA report into the use of PET and PET/CT for radiotherapy treatment planning therefore recommends the cautious consideration of FDG-PET information into target volume delineation for the head and neck highlighting that, despite improved correlation with pathological specimens relative to other modalities [14], exclusive reliance on the technique is associated with a significant risk of geographic miss [15]. Other functional imaging modalities that are being explored for head and neck radiotherapy, particularly in distinguishing different tumour types, include diffusion and perfusion techniques using magnetic resonance imaging (MRI) [16].

The focus of Publication #1 is on improvements to target volume delineation using anatomical, rather than functional, information. The most widely used alternative imaging modality for the head and neck is MRI, which provides improved soft tissue contrast and, depending upon the parameters of the imaging sequence used, improved resolution. The main limitation of MRI for this purpose is the difficulty in simultaneously replicating the treatment position while acquiring high quality images suitable for target delineation. Publication #1 describes a novel methodology for overcoming this problem by using superficial radiofrequency (RF) coils to acquire high quality images of the patient within the immobilisation shell used for radiotherapy treatment.

The precision of the Syntegra image fusion software with the Pinnacle Treatment Planning System (TPS) was evaluated for translational and rotational accuracy by quantifying the

shifts between identifiable landmarks in both the CT and MR image sets to assess the automated registration software, which was based on a mutual information algorithm.

1.5 Accuracy and precision of dose calculation and delivery

Systematic errors, shown to be of greater impact than random errors [17; 18], will also arise due to the limited accuracy of the dose calculation algorithm and the integrity of the image on which it is based. The mathematical model of the linear accelerator used for dose calculation must be sufficiently robust that it allows an accurate calculation of the delivered beam and the resulting dose distribution within the patient CT dataset. The current gold standard for dose calculation in radiotherapy is the Monte Carlo simulation, which combines a mathematical representation of the linear accelerator with the fundamental physics of the photon and electron interaction processes to determine the measured interaction probabilities and the resulting dose distribution. The probabilistic nature of this technique means that the accuracy of the simulation increases with the number of generated photons. In practice, millions of photons are required and the resulting computation power has limited the clinical application of the technique. There is currently no commercially-available TPS that incorporates Monte Carlo simulation for photons.

The most advanced commercial products, including the Pinnacle TPS used in this work, use the collapsed cone convolution (CCC) algorithm [19]. In Publication #2 the accuracy of the beam model representation in Pinnacle and it's efficacy for use with IMRT is evaluated by gamma analysis comparison of a series of clinical IMRT beam fluences calculated and measured at depth d_{max} in a solid water slab phantom. The sensitivity of the dose calculation to the individual modelling parameters was similarly assessed by comparison of the fluences calculated using beam models with and without the individual parameters. The intention of this work was to evaluate the efficacy of the CCC algorithm for use in head and neck IMRT and to isolate the additional modelling parameters that have the greatest influence and require accurate representation within the beam model.

The CT datasets used for treatment plan optimisation can be subject to streaking artefacts of varying severity due to the presence of high density dental amalgams. The accuracy of the heterogeneous dose calculation can be compromised by these regions of artificially high and low density [20], although their impact on the clinical IMRT dose distribution and the efficacy of simple correction strategies has not been thoroughly investigated. When using the 3D-CRT technique detailed in section 1.2, these artefacts are not a problem from the perspective of dose calculation accuracy, primarily because the lateral-opposed beams do not pass directly through the affected region, but also because many such treatments are calculated homogeneously and so do not account for the variation in density. With the advent of IMRT, the increased accuracy required means that heterogeneous calculations are highly recommended, while the ease with which satisfactory dose distributions can be achieved would be severely compromised by restricting the use of anteriorly-incident beams.

Publication #3 describes an investigation into the detrimental impact of planning on these compromised datasets by comparing the resulting dose distribution to that actually delivered, as simulated by applying the plan to a gold standard dataset corrected incorporating bulk density corrections for amalgam, teeth and soft tissue. The work aimed to quantify the extent of any dosimetric or radiobiological discrepancies between the planned and delivered dose distributions, to suggest criteria under which corrective action is required, and to evaluate the efficacy of simple correction strategies.

Reliable and accurate delivery of the IMRT treatment plan is crucial in ensuring that the intended dose is delivered to critical regions of the tumour and patient anatomy. Repeatable, precise positioning of the patient during treatment can allow target expansion margins to be reduced in much the same way as described in relation to improving confidence in delineation, allowing the potential of IMRT to be more fully exploited and optimising the therapeutic ratio. The advent of IGRT, in the form of on-board kV-imaging devices capable of acquiring volumetric scans, has been widely incorporated into clinical practice, particularly to verify the patient setup prior to treatment. Improvement in setup accuracy available using

this facility compared with portal imaging has been demonstrated [21], finding that these improvements in delivery precision could translate into reduced margins, leading to optimal application of IMRT techniques.

Publications #5 & #6 investigate the accuracy and precision of delivery of an IMRT plan, with the latter quantifying the potential gains of IGRT with on-board kV cone-beam imaging. Increasing the precision of delivery is vital in exploiting the full benefits of IMRT in terms of reduced PTV margins and the therapeutic window. Publication #6 details the development of a semi-anthropomorphic phantom intended for use in head and neck verification (HANK), while subsequent modifications are described in Publication #6. The design of the phantom was intended to improve on both the range of commercially-available geometric phantoms, which often were not representative of patient anatomy and the RANDO (originally produced by Alderson Research Laboratories, Stamford, CT) and the Radiological Physics Centre (RPC) phantoms, which provide a realistic patient representation but are limited in terms of flexibility of measurement. The phantom design therefore allows flexibility in measurement position either within the target or near normal tissue structures, can be used to measure both absolute point dose values and relative dose distributions and can incorporate a range of heterogeneities. Such phantoms are integral to the safe implementation, as well as ongoing audit, of modern treatment techniques.

The focus of Publication #6 is an attempt to quantify the potential gains of IGRT in terms of the accuracy and precision of plan delivery. IGRT with kV cone-beam volumetric imaging allows acquisition of images of the patient in the treatment position, which can then be used to precisely position the patient prior to treatment. This technique is already in clinical use but will be associated with an error that is currently not quantified. Using repeat measurement of an arrangement of 20 microMosfet detectors within the treated volume of the HANK phantom for 4 clinical IMRT plans, the accuracy and precision of the technical aspects of treatment delivery, under image-guidance to precisely align the phantom in the planned configuration, were quantified.

A limitation of current IMRT verification processes, as outlined in IPEM Report 96: Guidance for the Clinical Implementation of Intensity Modulated Radiation Therapy [22], is that the impact of any errors in the dose calculation or delivery can only be interpreted in terms of their impact on the dose distribution within the phantom used for verification or the variation seen on planar measurements of each beam fluence. Ideally, the impact of any errors would be evaluated in terms of the impact on the actual dose delivered to the patient so that the likely clinical importance of a given error in a given position can be assessed. The aim of Chapter 3 is to highlight the magnitude and location of the impact of residual errors in plan delivery on the patient dataset, which represents an important step in the development of IMRT quality assurance techniques: moving the analysis of errors to their direct impact on the dose delivered to the patient, compared with the current situation of attempting to indirectly project the impact of errors in individual beam fluences or on representative phantoms.

This has previously been attempted for homogeneous datasets, in which a sophisticated dose calculation algorithm is unnecessary, by back-projecting the fluence measured with an electronic portal imaging device (EPID) attached to the linear accelerator to a plane within the patient dataset [23]. The approach described in Chapter 3 attempts to incorporate the more sophisticated CCC dose algorithm contained within the Pinnacle TPS. This can be achieved by defining the in-air fluence required to give a planar dose calculated at d_{max} in a solid water slab phantom that matches the film fluence measured in identical conditions. In Chapter 3, a description is provided of the methodologies explored in attempting to achieve this.

1.6 Practicalities of treatment plan optimisation

From a practical perspective, IMRT planning requires a significant change to previously-established methods. Specifically, the inverse nature of the planning process, wherein the planner provides the TPS with an ideal solution by indicating the shape of the required dose-volume histograms, leaving the TPS to define the details of delivery, requires the planner to

think in a different way to that used for conventional planning. A thorough understanding of this process and the means to manipulate the dose distribution resulting from the specified dose-volume histogram (DVH) constraints allows impressive target conformality, dose escalation and sensitive tissue avoidance to be achieved with only modest compromise, even for the most complex sites such as the nasopharynx. However, the work required to obtain this experience is often substantial and IMRT planning can initially appear a daunting proposition. This lack of confidence has led to moves to reign in the scope of IMRT when faced with early problems rather than fully exploiting the potential of the technique.

An example is the clinical concern raised upon observation of increased incidence of dysphagia after treatment with IMRT. This occasionally severe toxicity, which significantly impacts on quality-of-life was correctly attributed to the removal of spinal blocking from the anterior field traditionally used for such treatments. Although the spinal cord had been successfully spared by other means for the IMRT patients, the increased dose to the larynx had been overlooked. In light of this, several centres retreated from full implementation of IMRT by matching the previously used anterior field to a pre-planned IMRT plan in the anatomically superior structures, resulting in substantial compromise to target coverage at the matchline and in the inferior nodes and a risk of over or underdosage at treatment. Other centres attempted to add a constraint to the larynx into the IMRT optimisation process but found that it created errors elsewhere in the plan, eventually re-adopting the conventional neck field [7].

Publication #4 describes a simple method for incorporating larynx sparing into the optimisation process without compromising other aspects of the plan, aiming to achieve similar levels of larynx sparing to those using the spinal blocking technique.

The widespread implementation of IMRT for the head and neck in the UK was driven largely by the PARSPORT Trial [24], a key endpoint of which was to confirm and quantify the expected reduction in xerostomia by reducing the dose to the contralateral parotid gland. Several studies have demonstrated increased saliva flows compared with conventional

techniques [25; 26] generally finding a 1Gy drop in parotid gland mean dose leading to a 4% reduction in salivary flow [27]. However, xerostomia is predominantly a quality-of-life issue and studies looking at patient-reported xerostomia reveal a more uncertain outcome, indicating that objective quantitative measures of saliva flow did not correlate strongly with patient-reported xerostomia [28]. Jellema et al used questionnaires completed by 157 patients to establish strong correlations with xerostomia (identified as a dry mouth) for a combination of both the mean dose to the parotid gland and that to the submandibular gland (SMG), finding that the probability of a given mean parotid dose resulting in xerostomia increased with the mean SMG dose [29]. The saliva flow was found to reduce exponentially as mean SMG dose increased up to a threshold of 39Gy, beyond which no further deterioration in symptoms was seen. Previous studies have demonstrated a similar effect for the parotid gland with a lower threshold of approximately 24Gy [30]. The study by Jellema et al found that 'sticky saliva' correlated only with the mean dose to the SMG and was independent of the parotid gland dose. It has been suggested [31] that this could be explained by the relative lack of mucins (mucosal lubricants) in parotid saliva compared to that from the SMG. Eisbruch et al have also demonstrated that sparing fo the minor salivary glands within the oral cavity is a significant independent predictor of patient-reported xerostomia [26].

A recent study has demonstrated the efficacy of sparing the SMG by including it in the IMRT optimisation [31], although in most centres the proximity to the target volume means that any sparing is applied cautiously. Successful sparing has also been demonstrated from surgical transfer techniques that move the SMG away from the target region [32]. Sparing of the contralateral parotid gland with IMRT is not found to result in increased loco-regional failure [33] although care must be taken when sparing the ipsilateral parotid gland, ensuring that compromises to target coverage are not made in high-risk lymph node levels, regardless of the apparent metabolic activity observed on PET scans [34].

1.7 Risk of secondary malignancies

Further clinical concerns over the implementation of IMRT have been expressed due to the increased number of MU required for delivery and the increased volume of low dose irradiation to the patient due to the more isotropic beam arrangement, both of which are expected to increase the risk of the treatment inducing a secondary malignancy [8]. The increasing prevalence of IMRT, together with increasing survival after radiotherapy, have meant that this concern is being taken increasingly seriously and attempts have been made to calculate the risk associated with different treatment techniques.

Models for calculating this risk, for which large-scale studies suggest radiotherapy-induced secondary cancer in approximately 6-9% of patients [35; 36], are currently associated with large uncertainties due to the lack of established data available to validate them. Due to the relatively small risks and very late toxicities being observed, such studies require large cohorts of patients and long follow-up times. As a result, these models tend to be very simplistic, not directly taking into account the precise dose distribution relative to sensitive organs, inter-organ sensitivity variations or the dependency of secondary cancer risk on patient age at exposure [37].

A methodology that aims to expand and improve on those currently available is described in Chapter 4. A voxel-based, rather than organ-based, approach is used that allows calculation according to the precise dose distribution and sensitivity of the tissue in question and will allow increased flexibility to include relevant data acquired in the future. Corrections are made for age, gender and relative organ sensitivity (based on established data from BEIRVII) and a comparison made of conventional and IMRT techniques for treatment of the head and neck based on dose calculation in adult and paediatric phantoms.

Publication #1

A novel imaging technique for fusion of high quality, immobilised magnetic resonance (MR) images of the head and neck with computed tomography (CT) scans for radiotherapy target delineation.

Published in The British Journal of Radiology, 82 (2009), p497-503

Authors:

Webster GJ

Killgallon JE

Ho KF

Rowbottom CG

Slevin NJ

Mackay RI

18 pages

Abstract

Introduction

Uncertainty and inconsistency is observed in target volume delineation in the head and neck for radiotherapy treatment planning based only on computed tomography (CT) imaging. Alternative modalities such as magnetic resonance (MR) imaging have previously been incorporated into the delineation process to provide additional anatomic information. This work aims to improve on previous studies by combining good image quality with precise patient immobilisation in order to maintain patient position between scans.

Methods and materials

MR images are acquired using quadrature coils placed over the head and neck while the patient is immobilised in the treatment position using a 5-point thermoplastic shell. The MR image and CT images are automatically fused in the Pinnacle Treatment Planning System (TPS) using Syntegra software. Evaluation of image quality, distortion and the accuracy of the image registration using patient anatomy are described.

Results

Image quality was found to be superior to that acquired using the body coil, while distortion was <1.0mm to a radius of 8.7cm from the scan centre. Image registration accuracy was found to be 2.2mm (±0.9mm) and <3.0 degrees (n = 6).

Conclusions

A novel MR imaging technique that combines good image quality with rigid patient immobilisation has been developed and is now in clinical use. The scan duration of approximately 15 minutes has been well tolerated by all patients.

Introduction

Many institutions, including the Christie Hospital, currently delineate gross tumour volume (GTV) and clinical target volume (CTV) in the head and neck regions using only a computed tomography (CT) scan, possibly with the aid of prior diagnostic CT. The limitations of this approach are inherent to the imaging modality; CT scans result in poor soft tissue contrast due to the similarity of the electron density between healthy soft tissue and malignant disease and are also susceptible to significant streak artefacts due to dental amalgams near the target region. Another inherent problem with using an axial CT dataset for volume delineation is the presence of a partial volume effect, which limits the image resolution in the cranio-caudal axis to the slice width of the original scan. These factors can make the delineation of the target volume uncertain, leading to increased inter-observer variation and reduced confidence in treatment margins.

The potential of modern treatment techniques, such as intensity modulated radiotherapy (IMRT), which allow improved dose conformality and possible dose escalation, cannot be fully realised unless confidence exists in our knowledge of the target volume. Any error in target delineation will appear as systematic throughout treatment and could lead to a significant under-dosage to a region of the tumour or unnecessary dosage to a critical structure; risks that should be accounted for in target volume expansion strategies. A study by Hong et al looked at the delineation process for several centres performing IMRT for the head and neck; even using a pre-defined GTV for the patient in question, thus removing one possible source of uncertainty, the study found significant variation in the delineated target volume, ranging from 35-175cc for the CTV in the ipsilateral neck [1].

To reduce uncertainties in delineation, alternative imaging modalities have been used to complement CT. The Royal College of Radiologists recommend using magnetic resonance (MR) imaging for selected sub-sites of the head and neck such as oral cavity, oropharynx,

nasopharynx and hypopharynx [2]. The benefits of MR imaging for target delineation are the improved soft tissue contrast and cranio-caudal resolution compared to standard CT, particularly when tumour conspicuity is enhanced by the uptake of Gadolinium-based contrast agents or fat suppression techniques [3, 4]. The impact of dental artefacts is also reduced using MR and there is no associated radiation risk. Although it is possible to use helical CT to improve the z-axis resolution, the poor soft tissue contrast must be accepted. MR is currently used as the principal modality for delineation of head and neck tumours at several institutions for sites such as the oral cavity, oropharynx [5] and nasopharynx [6]. However, for applications in head and neck radiotherapy, MR should remain as a complimentary imaging modality, rather than a replacement for CT, due to the need for heterogeneity correction, which is reliant on electron density information not provided by MR.

The impact of MR has been demonstrated in several studies, generally finding that MR-derived GTVs in the head and neck are smaller and subject to significantly less inter-observer variation than those derived with CT [7]. However, this finding is not applicable to all clinical sites in the head and neck. Emami et al found that a sample of 8 nasopharyngeal carcinoma patients showed consistently larger tumour volumes when delineated with MR only compared to CT only (mean volume 76% greater for MR) due to the improved detection of intracranial spread with MR [8]. However, this study made no attempt to immobilise the head for the MR scan. The findings were reinforced by Chung et al, who compared the detection rate of CT and MR images of intracranial infiltration for 51 nasopharynx patients with pterygopalatine fossa involvement. The measured sensitivities of 96.1% and 56.9% for MR and CT respectively indicate that target delineation was significantly improved by the more detailed disease classification provided by MR.

A study by Geets et al, which looked at inter-observer variability as well as a comparison of MR and CT-assisted delineation, suggested that for pharyngo-laryngeal tumours there were no significant benefits to the use of MR above CT. However, this study was limited by the non-

optimal MR image acquisition method that used the integrated body coil as a receiver rather than head or surface coils, the use of only axial MR data in order to reduce acquisition time and a small sample size (n=10 for each tumour site) [9].

Rasch et al conducted a small study comparing GTV delineation by 4 observers for 6 advanced head and neck cancer patients using axial CT, axial MR and coronal/sagittal MR images [7]. The results indicated that CT delineation resulted in significantly greater observer variation and tended towards larger volumes (mean CT volume greater for 4/6 patients). The study covered several sites and served to demonstrate particular cases in which the different modalities could be of benefit; for example, the spread through the base of skull from an ethmoid sinus patient was missed on axial CT and axial MR but identified from coronal MR sections. This removal of the partial volume effect allowed clearer tumour visualisation and led to greater consistency from using sagittal/coronal MR images compared to axial MR alone. This study, as well as others [10], indicated that the two modalities are complimentary with different tumour extensions being better visualised by the different modalities. MR was found to be better for soft-tissue borders and bone marrow invasion, while CT provided better visualisation of tumour-fat boundaries and bone cortex invasion.

The current work has been developed in support of a local trial, PRECISE, that aims to quantify the differences between both delineated tumour volume and inter-observer variability when comparing fused MR/CT with CT alone for pharyngeal cancers. Local ethics approval was obtained for the MR images to be used retrospectively and were not part of the planning process. The study advances from those outlined above in that it aims to combine good MR image quality, which normally requires use of the head coil, with patient immobilisation to allow rigid-body fusion with the CT scan, the latter being incompatible with commercially-available head coils.

Several practical issues were considered in the development of the MR-imaging technique used for the study; image distortion; image quality; setup repeatability; patient comfort and scan duration; scan parameters; MR-compatibility of equipment; image registration software.

Methods and materials

Imaging protocol

The patients for this study are scanned on a Siemen's Avanto 1.5T scanner (Siemens, Erlangen, Germany). The patient is positioned on a Sinmed Posicast IMRT baseboard (Sinmed, Reeuwijk, The Netherlands) that allows 5-point fixation of thermoplastic shells over the shoulders so as to ensure repeatability of patient position between CT scan, MR scan and treatment. The patient positioning is identical to treatment other than the need to wear earplugs for the MR scan and the total scan time of approximately 15 minutes is comparable to a typical IMRT treatment delivery time. Although non-immobilised images are often fused with CT for whole-brain imaging, it was felt that the risk of neck flexation and changing position of the head relative to the shoulders without immobilisation using a 5-point fixation shell would be too great. There is also concern that distortions in the patient contour between scans could impact negatively on image registration accuracy.

Table 1. Optimised scan sequence parameters

	T_1 pre- and post- contrast	T_1 fat suppression
TR (ms)	8.4	9
TE (ms)	4.76	3.35
Flip angle (degrees)	20	10
Field of view (mm)	320	320
Matrix	192 x 256	144 x 256
Voxel size (mm)	1.3 x 1.3 x 1.3	1.7 x 1.7 x 1.3
Number of slices	192	192
Number of averages	2	2

TE, echo time; TR, time to repeat

The MR imaging setup is illustrated in Figure 1. A simple Perspex support, designed to elevate the coil away from the patient so as to improve patient comfort, was designed to be positioned under the fixation board after the patient is immobilised. The coil is positioned so that its centre covers the approximate target area. The equipment used is MR-compatible regarding the absence of ferrous materials and can be setup consistently between patients. Surface coils are used as the immobilisation shell precludes the use of the standard head coil, while the surface coils are expected to provide superior image quality to the quadrature body coil used in previous studies. This assumption has been checked by measuring the signal-to-noise ratio (SNR) using a spherical flood phantom (diameter 25cm) to give a comparison between the body coils and the setup used for this study. The analysis was carried out in the Pinnacle Treatment Planning System (TPS) v7.4f (Philips Medical Systems, Andover, MA) and the relative SNR was determined within the phantom using the equation

$$SNR = \frac{S_{water} - S_{air}}{SD_{water}}$$

where S_{water} is the mean signal value within a central sphere of approximate diameter 13cm within the 23cm diameter flood phantom. SD_{water} is the standard deviation on this value and S_{air} is an average signal value from the image outside the phantom.

The MR imaging protocol consists of a 3D T1 Gradient Echo (GE) sequence acquired pre- and post- administration of a Gadolinium(Gd)-based contrast agent, supplemented by a post-contrast 3D T1 GE sequence with fat-suppression. 3D sequences were chosen due to their inherently lower distortion than 2D sequences. Additionally, the isotropic resolution of the 3D T1 GE pre- and post- contrast sequences permits scanning to be performed in the sagittal plane, reducing acquisition time compared to a transaxial acquisition with the equivalent foot head coverage, but enabling reformatting of the data into the transaxial plane for image fusion with the CT scan. The imaging sequences were optimised using repeated scans on volunteers with regard to image

resolution, SNR and acquisition time and the optimised parameters are listed in Table 1. The results shown in this work were acquired using the fat-suppressed sequence, although no significant differences were observed using the other sequences. The CT scans were acquired on a GE Lightspeed scanner with a 3mm slice thickness throughout the approximate disease volume and 5mm inferiorly. The axial images are 512x512 arrays of 0.9mm resolution pixels.

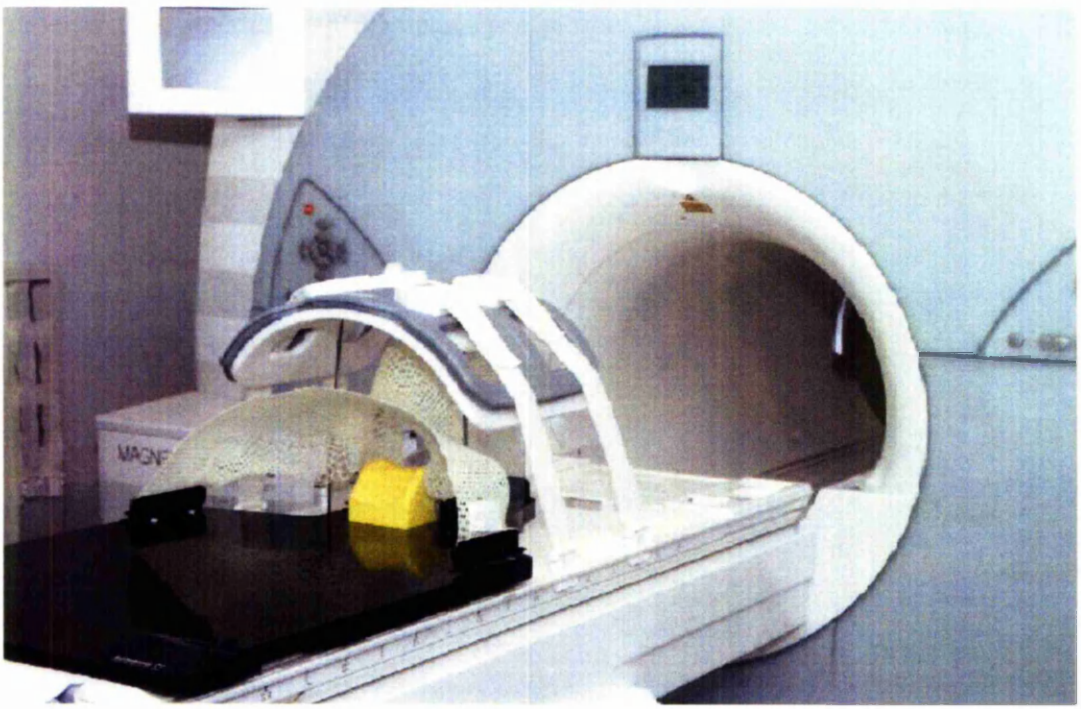
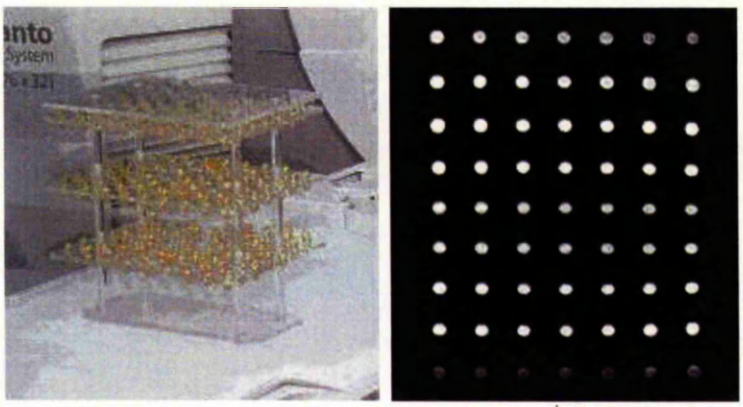


Figure 1. Scan setup in treatment position using immobilisation shell and surface coil

Distortion

Image distortion is a common problem in MR imaging and is due both to non-uniformities in the magnetic field and non-linearities in the applied gradient. Typically, the effect is minimal at the centre of the image but can extend to several centimetres at the field extremities. The extent of the distortion can be influenced by the scan parameters and coil position and so for any images that will be used for volume studies it is advisable to quantify the distortion prior to use. An

existing phantom [11], designed to assess distortion over larger field of views than of interest here [12], was unsuitable for use with our surface coil. It was decided that a simplified phantom could be used to quantify the distribution of distortion.



a) b)

Figure 2. a) Linearity test object used for quantifying and mapping image distortion, b) simplified in-house phantom and c) resulting axial slice of MR image

The phantom, illustrated in Figure 2a, consists of a 9 x 7 x 3 array of cod-liver oil capsules. These are positioned in 3 layers within a Perspex support approximating the dimensions of a head (25cm x 20cm x 15cm). The capsule separation is 2.5cm in each direction along each plane and 7.5cm between planes, the capsules themselves have a central diameter of ~8mm and a length of ~2cm. To achieve 2.5cm resolution in all axes, over a volume of 15x20x25cm (lateral, anterior-posterior and superior-inferior axes respectively), scans of the phantom in 6 configurations were required and are illustrated in Figure 3. The scans were exported to the Pinnacle TPS and the central coordinates of each capsule were manually recorded. It was attempted to do this by automatically outlining each capsule using the thresholding facility on Pinnacle and using the volume to determine the centre of mass, but due to the noise on some of the images this was not practical. A selection of suitable low-noise capsule images revealed minimal differences between the manual and automatic results. The manual method was then repeated on a small sample of markers and resulted in negligible variation from the initial values. The manually-determined coordinates of each marker can then be compared to a regular grid, which will allow the distortion to be determined. In-house software written in IDL (ITT Visual

Information Solutions, Boulder, CO) was used to determine the distance from the scan centre to which a particular magnitude of distortion is not exceeded.

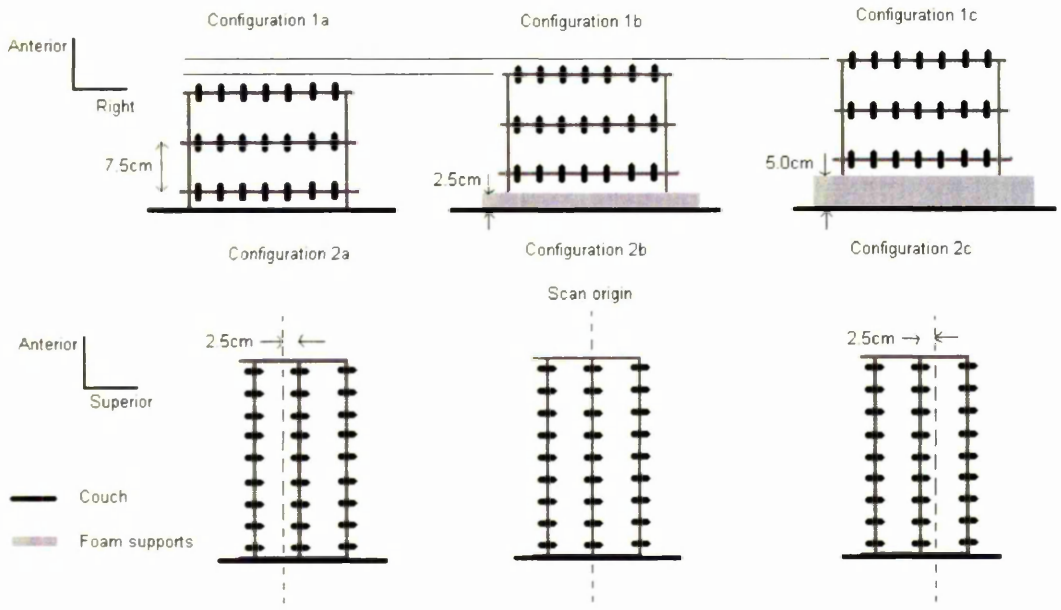


Figure 3. Distortion phantom configurations necessary to acquire 3D vector distortion data for the MR scan sequence. Configurations 1a-c allow quantification of lateral and superior-inferior distortion in nine coronal planes but do not provide information on anterior-posterior distortion as the precise position of the centre of the markers in this axis is uncertain. Configurations 2a-c are therefore used to quantify this distortion in nine axial planes.

Image registration

A final technical concern for this process is the accuracy of the image registration software image registration, which has not currently been rigorously tested. A phantom was not thought suitable for this test as it would not have the complex anatomical variations present in a patient and would therefore not be a sufficiently robust test. The scans from the first 6 patients recruited to the PRECISE study were therefore automatically fused within the Syntegra software on the Pinnacle TPS. Previous work has shown that the mutual information algorithm is more robust for this purpose [13]. It was decided to use anatomical volumes rather than points to assess the

image accuracy as, with this method, discrepancies in windowing or edge detail between the two modalities would lead only to volumes of different size rather than points in the incorrect position. Since patient scans are being used this is not strictly a test of the image registration software but of our imaging technique, which includes the software and the patient immobilisation.

The method used to determine the registration accuracy is illustrated in Figure 4. The spinal canal was chosen as a volume for comparison as it is clearly visible on both the CT and MR datasets, will not move within the patient either during or between scans and, due to it's central position within the patient, will be minimally affected by any distortion. The canal was manually contoured on individual slices every centimetre on both CT and MR modalities and the Pinnacle TPS used to determine the centroid coordinates on each slice. The lateral and anterior-posterior positions were then plotted against the z-axis position and equations established of linear best fits for each modality (where significant curvature was present, a linear region of >4cm was isolated). The resulting gradients were then used to establish the coronal and sagittal rotational discrepancies (Fig.4a & b). The MR fit was then rotated about the central point of the plot to match the gradient of the CT plot and the new equations used to determine the required lateral and anterior-posterior shifts.

The roughly cylindrical shape of the spinal canal means that it is unsuitable for determining z-axis translational or axial rotational inaccuracies. The average superior-inferior translational error was determined using the relative positions of the centroids of the eyes after manual volumetric delineation in each dataset (Fig.4c), while the axial rotational error was determined by measuring the angle created from the anterior axis by a line from the centre of the spinal canal and the centroid of the eye volume (Fig.4d). This process was carried out using both Syntegra software in v7.4f of the Pinnacle TPS and v8.0h, the latter of which allowed a region of the dataset to be specified for fusion. For all patients, this region ran from approximately the optic chiasm to the hyoid bone, containing the entire patient contour within these slices and omitting the shoulder region from the registration volume.

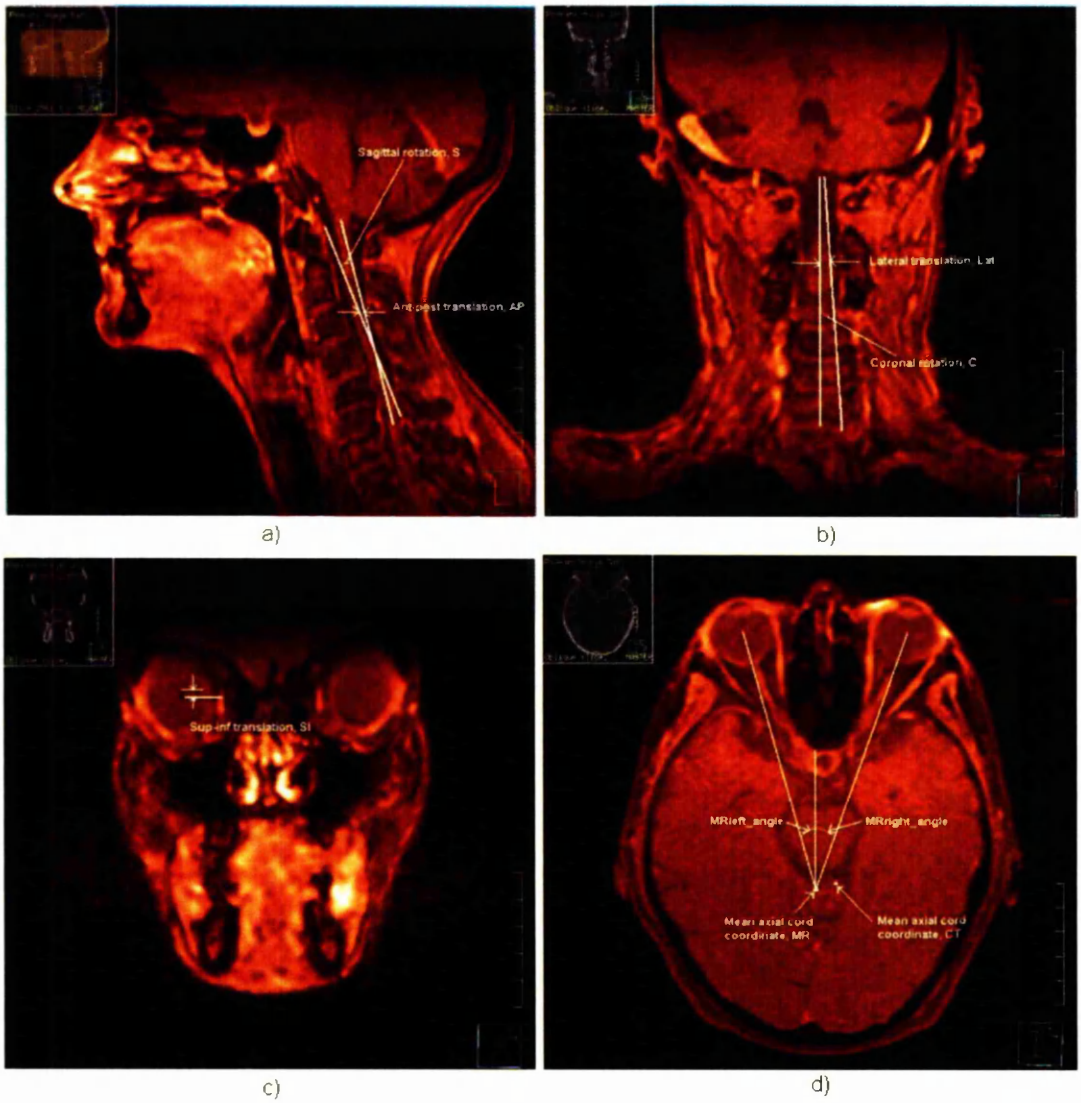


Figure 4. Image registration procedure: measurement of a) sagittal (S) and b) coronal (C) rotation from linear plots of the spinal canal central coordinates, also used to determine lateral (Lat) and anterior-posterior (AP) translational discrepancies. c) Superior-inferior translational error (SI) is calculated as the shift from the centroid of the eye outlines from the MR and CT datasets, averaged from both eyes. Axial rotation (A), using the mean coordinate of the spinal canal position as a pivot around which to determine the relative rotation of the centre of mass of the eyes in each dataset, is calculated from the equation

$$A = \frac{(CT_{left\_angle} - MR_{left\_angle}) + (CT_{right\_angle} - MR_{right\_angle})}{2}$$

Table 2. Signal-to-noise (SNR) results

Coil	Volume (cm^3)	S_{mean}	S_{air}	$s.d._{water}$	SNR
Surface	1314.4	33.82	0	1.02	33.2
Body	1344.3	52.24	1	1.89	27.1

S_{air}, average signal value from the image outside the phantom; $s.d._{water}$, standard deviation of S_{water}; SNR, signal-to-moise ratio; S_{water}, mean signal value within a central sphere of approximate diameter 13cm within the 23cm diameter flood phantom

The image distortion assessment indicates that magnetic field inhomogeneity is within acceptable levels. The distortion map shown in Figure 5 demonstrates good image integrity; <1.0mm distortion within a radius of 8.7cm from the scan centre (although it should be remembered that the measured region extends only 7.5cm either side of the scan centre in the left-right axis). Similarly, the scan distortion is within 2.0mm to a radius of 12.8cm. This compares favourably with measured distortions by Wang et al for a range of commercial scanners; after vendors corrections were applied 3D vector distortions of >1mm were found within 5.0cm and distortions >2mm within 10cm of the scan centre [14]. The repeatability of the scan distortion and sensitivity to coil position has been confirmed by repeated setup and acquisition of the same imaging sequence with the coil deliberately shifted 3cm superiorly and laterally. No significant effect on the distortion was seen, suggesting that imaging integrity is not limited by the precise positioning of the coils and that longer volumes could be scanned. The signal-to-noise ratio (SNR) is increased (Table 2) for our imaging setup in comparison to the sequence adapted for the body coil.

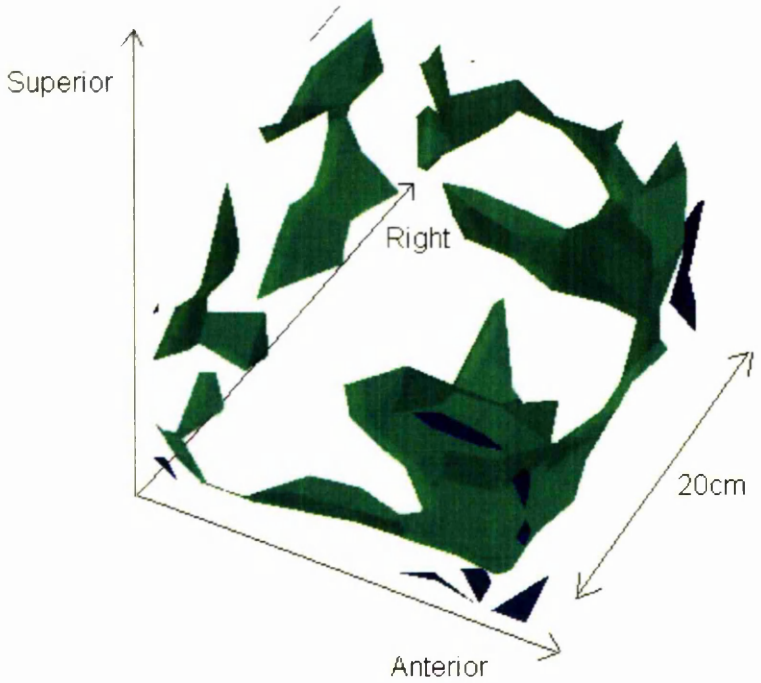


Figure 5. Distribution of distortion. The green region represents the distortion isocontour at 1mm, the blue represents that at 2mm.

The results in Tables 3 and 4 show the accuracy of image registration on 6 patients using 2 versions of the Pinnacle software. The facility in v8.0h to isolate a region of either dataset for fusion allows improved accuracy compared to the registration using the full CT and MR datasets used in v7.4f. This improvement is seen mainly as a result of a reduction in the extent of outliers (the later version has no individual translational errors >3mm and no rotational errors >3degrees). This is probably due to the ability of the image fusion software to exclude the shoulder contours from the image registration, which are likely to be subject to greater positional uncertainty than superior regions of the patient. Moore et al found a maximum discrepancy of 0.18cm in the registration of images of a semi-anthropomorphic phantom between the two modalities (compared to a maximum of 0.33cm in this study), which suggests that the patient-induced errors in the current work are in the region of 0.15cm [15].

Table 3. Image registration results for v7.4f, in which the full data set was incorporated into the fusion

	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6	Mean	SD	RMS
Lateral (cm)	-0.34	0.11	0.14	-0.06	-0.12	-0.04	-0.05	0.17	0.17
Ant-Post (cm)	0.02	-0.4	0.13	0.1	-0.11	-0.13	-0.07	0.2	0.19
Sup-Inf (cm)	-0.05	-0.06	-0.28	-0.13	0.28	0.13	-0.02	0.19	0.18
Vector shift (cm)	0.34	0.42	0.33	0.17	0.32	0.18	0.3	0.1	0.31
Coronal (deg)	3.39	0.35	0.17	-1.48	1	1.36	0.8	1.6	1.67
Sagittal (deg)	-0.29	2.27	-0.47	-1.15	-0.04	-1.06	-0.12	1.25	1.15
Axial (deg)	2.42	-1.12	-0.88	0.99	3.26	-0.68	0.66	1.86	1.82

RMS, root-mean-square; SD, standard deviation; Ant-Post, Anteroposterior; Sup-Inf, Superior-inferior

Table 4. Image registration results for v8.0h, in which the CT data set used for fusion was limited to a specific volume

	Patient 1	Patient 2	Patient 3	Patient 4	Patient 5	Patient 6	Mean	SD	RMS
Lateral (cm)	-0.23	-0.04	0.08	-0.08	-0.04	-0.21	-0.09	0.12	0.14
Ant-Post (cm)	0.08	-0.08	0.11	0.15	0.05	-0.01	0.05	0.08	0.09
Sup-Inf (cm)	0.2	-0.06	-0.14	-0.28	0.13	-0.11	-0.04	0.18	0.17
Vector shift (cm)	0.32	0.11	0.2	0.33	0.14	0.23	0.22	0.09	0.24
Coronal (deg)	-2.49	0.56	-1.25	-0.1	-1.1	-0.19	-0.76	1.08	1.25
Sagittal (deg)	0.5	1.42	0.76	-1.81	-0.42	0.84	0.22	1.16	1.08
Axial (deg)	-0.16	-0.12	-0.46	0.91	1.21	2.38	0.63	1.08	1.17

RMS, root-mean-square; SD, standard deviation; Ant-Post, Anteroposterior; Sup-Inf, Superior-inferior

Discussion

The imaging technique has been implemented and is now in clinical use. The PRECISE study, which will quantify any advantage that MR images offer to delineation of GTV for pharyngeal cancers, is currently in progress. The scan setup is simple and stable, with patients reporting a similar level of comfort to treatment during a total scan time of approximately 15 minutes.

The MR images acquired using the surface coil are not of optimal quality but offer superior signal-to-noise compared to those obtained with a body coil and are suitable for target delineation for the purposes of radiotherapy treatment planning. The setup procedure is practical and repeatable between scans. System-induced distortion is found to be acceptable (<1mm)

within the likely target region and does not appear to significantly impact on the ability of the image fusion software to accurately co-register the images (see Tables 3&4).

Another benefit of the MR images is that they allow much greater confidence in the delineation of critical structures, particularly the spinal cord and brainstem, due to the significantly improved soft tissue contrast.

Subsequent work will attempt to compare the MR technique detailed in this work with metabolic imaging techniques such as CT-PET, although there are still many unresolved questions regarding the latter modality [16]. Attempts to further improve our image quality using the current technique are limited by the necessary increases in scan duration required to improve the signal-to-noise ratio with the current generation of scanners. The image display on the TPS, which is required for fusion with other modalities, is also of limited quality as the system was not specifically designed to display high quality images.

Conclusion

A novel MR imaging technique has been developed and clinically-implemented that uses surface coils to optimise the reliability of fusion with CT images in terms of image quality and patient immobilisation. Practical and technical aspects have been investigated and found to be acceptable for clinical use.

References

1. Hong TS, Tome WA, Chappell RJ and Harari PM. Variations in target delineation for Head & Neck IMRT: An international multi-institutional study. Int J Radiat Oncol Biol Phys, 2004;**60**(1):S157-S158
2. RCR, T.R.C.o.R., Recommendations for Cross Sectional Imaging in Cancer Management. 2006.
3. Tien R and Robbins K. Correlation of clinical, surgical, Pathologic and MR fat suppression results for head and neck cancer. Head Neck, 1992. **14**(4): p. 278-284.
4. Phillips CD, Gay SB, Newton RL and Levine PA. Gadolinium-enhanced MRI of tumours of the head and neck. Head Neck, 1990;**12**(4):308-315
5. Weber AL, Romo L and Hashmi S. Malignant tumours of the oral cavity and oropharynx: clinical, pathologic and radiologic evaluation. Neuroimaging Clin N Am, 2003;**13**(3):443-464
6. Chung NN, Ting LL, Hsu WC, Lui LT and Wang PM. Impact of magnetic resonance imaging versus CT on nasopharyngeal carcinoma: primary tumour target delineation for radiotherapy. Head Neck, 2004;3;241-246
7. Rasch C, Keus R, Pameijer FA, Koops W, De Ru V, Muller S et al. The potential impact of CT-MRI matching on tumour volume delineation in advanced head and neck cancer. Int J Radiat Oncol Biol Phys, 1997;**39**(4):841-848
8. Emami B, Sethi A and Petruzzelli GJ. Influence of MRI on target volume delineation and IMRT planning in nasopharyngeal carcinoma. Int J Radiat Oncol Biol Phys, 2003;**57**(2):481-488
9. Geets X, Daisne JF, Arcangeli S, Coche E, De Poel M, Duprez T et al. Inter-observer variability in the delineation of pharyngo-laryngeal tumor, parotid glands and cervical spinal cord: comparison between CT-scan and MRI. Radiother Oncol, 2005;**77**:25-31
10. Weltens C, Menten J, Feron M, Bellon E, Demaerel P, Maes F et al. Interobserver variations in gross tumour volume delineation of brain tumours on computed

tomography and impact of magnetic resonance imaging. Radiother Oncol, 2001;**60**:49-59

11. Wang D and Doddrell DM. A proposed scheme for comprehensive characterization of the measured geometric distortion in magnetic resonance imaging using a three-dimensional phantom. Med Phys, 2004;**31**(8):2212-2218

12. Doran S, Charles-Edwards E, Reinsberg SA and Leach MO. A complete distortion correction for MR images: I. Gradient warp correction. Phys Med Biol, 2005;**50**:1343-1361

13. Loi G, Dominietto M, Manfredda I, Mones E, Carriero A, Inglese E et al. Acceptance Test of a Commercially Available Software for Automatic Image Registration of Computed Tomography (CT), Magnetic Resonance Imaging (MRI) And 99mTc-ethoxyisobutylisonitrile (MIBI) Single-Photon Emission Computed Tomography (SPECT) Brain Images. J Digit Imaging, 2007. **Online**.

14. Wang D, Strugnell W, Cowin G, Dodrell DM and Slaughter R. Geometric distortion in clinical MRI systems Part I: evaluation using a 3D phantom. Magn Reson Imaging, 2004;**22**:1211-1221

15. Moore CS, Liney GP and Beavis AW. Quality assurance of registration of CT and MRI data sets for treatment plnning of radiotherapy for head and neck cancers. J Appl Clin Med Phys, 2004;**5**(1):25-35

16. MacManus MP and Hicks RJ. Where Do We Draw the Line? Contouring Tumors on Positron Emission Tomography/Computed Tomography. Int J Radiat Oncol Biol Phys, 2008;**71**(1):2-4

Publication #2

Development of an optimum photon beam model for head and neck IMRT

Published in the Journal of Applied Clinical Medical Physics, 8(4), 2007

Authors:

Webster GJ

Rowbottom CG

Mackay RI

12 pages

This page is intentionally blank

JOURNAL OF APPLIED CLINICAL MEDICAL PHYSICS, VOLUME 8, NUMBER 4, FALL 2007

Development of an optimum photon beam model for head-and-neck intensity-modulated radiotherapy

Gareth J Webster, Carl G Rowbottom, and Ranald I Mackay
North Western Medical Physics, Christie Hospital NHS Trust, Manchester, United Kingdom.
Gareth.Webster@physics.cr.man.ac.uk

Received 25 June 2007; accepted 23 August 2007

Intensity-modulated radiotherapy (IMRT) for complex sites such as tumors of the head and neck requires a level of accuracy in dose calculation beyond that currently used for conformal treatment planning. Recent advances in treatment planning systems have aimed to improve the dose calculation accuracy by improving the modeling of machine characteristics such as interleaf leakage, tongue and groove, and rounded multileaf collimator (MLC) leaf ends. What is uncertain is the extent to which these model parameters improve the agreement between dose calculation and measurements for IMRT treatments.

We used Pinnacle version 7.4f (Philips Medical Systems, Andover, MA) to carry out optimization of additional photon-beam model parameters for both an Elekta Precise (Elekta, Stockholm, Sweden) and a Varian (Varian Medical Systems, Palo Alto, CA) linear accelerator (LINAC). One additional parameter was added to the beam models in turn, and associated models were commissioned to investigate the dosimetric impact of each model parameter on 5 clinical head-and-neck IMRT plans. The magnitude and location of differences between the models was determined from gamma analysis of the calculated planar dose maps. A final model that incorporated all of the changes was then commissioned. For the Elekta Precise, the impact of all the changes was determined using a gamma analysis as compared with measured films.

Cumulative differences of up to more than 3%/3 mm were observed when dose distributions with and without all of the model changes were compared. Individually, for both LINACs, the addition of modeling for the rounded MLC leaf ends caused the most dramatic change to the calculation of the dose distribution, generating a difference of 3%/3 mm in up to 5% of pixels for the 5 patient plans sampled. The effect of tongue-and-groove modeling was more significant for the Varian LINAC (at 1%/1 mm, mean of 25% of pixels as compared with 5% of pixels with the Elekta Precise LINAC). The combined changes to the Elekta model were found to improve agreement with measurement.

Current commercially available treatment planning systems offer accuracy sufficient for clinical implementation of head-and-neck IMRT. For this treatment site, the ability to accurately model the rounded MLC leaf ends has the greatest affect on the similarity of the calculated dose distribution to measurements. In addition, for the Varian LINAC, modeling of the tongue-and-groove effect was also advantageous.

PACS numbers: 87.53.-j, 87.53.Bn, 87.53.Tf

Key words: IMRT, TPS, commissioning, verification, rounded leaf ends

[a] Corresponding author: Gareth Webster, North Western Medical Physics, Christie Hospital NHS Trust, Wilmslow Road, Manchester, U.K.; phone: +44-0161-446-3536; fax: +44-0161-446-3532; email: Gareth.Webster@physics.cr.man.ac.uk

I. INTRODUCTION

The delivery of intensity-modulated radiotherapy (IMRT) may require greater accuracy in the beam model used for calculation of the dose distribution than has traditionally been demanded for conformal radiotherapy techniques. The increased complexity of IMRT plans can result in a significant proportion of the dose in high-dose regions being a summation of dose contributions from penumbra and out-of-field regions from several individual beam segments. Dose inaccuracies that would be accepted in low-dose regions can become very important when they contribute to high-dose regions—for example, close to critical structures. The case of penumbrae is particularly relevant, because a slight error in penumbral position could cause a marked change in the dose at a particular point.

A study by Schwarz et al.[1] investigated the reliability of a standard treatment planning system (TPS) in calculating an IMRT dose distribution. Those investigators first verified that two commercial TPSs fulfilled the dose calculation accuracy criteria for conventional treatment techniques, as reported by the American Association of Physicists in Medicine and others.[2] They then used the two beam models to calculate the dose distribution to a series of prostate and head-and-neck IMRT plans for comparison with measured segment doses, demonstrating good dose agreement (<3%) for points in the center of 95% of segments, but errors of up to 40% (more than 5% of isocentric dose) when measured beneath the jaws. The discrepancies between the two systems were compounded by the calculated values being either side of the measured values. Their work confirms that the accuracy criteria required in beam models for conventional radiotherapy is insufficient for the clinical implementation of IMRT, because a critical dose to, for example, the spinal cord could be significantly underestimated if it is partly the summation of out-of-field or penumbral doses from several segments.

Modern commercial TPSs form mathematical representations of the radiation beam based on a number of variable parameters that reflect particular aspects of the linear accelerator (LINAC). The Philips Pinnacle TPS (Philips Medical Systems, Andover, MA) has recently been upgraded, incorporating a number of new features into the modeling software[3] to improve on the limitations of the previous release.[4] These include an improved auto-modeling facility that may remove the previous subjectivity from the modeling process, and refinements to the model itself, such as tongue-and-groove width, interleaf leakage, independence of x and y jaw transmissions, modeling of the flattening filter as a radially-symmetric arbitrary profile rather than as a cone, and modeling of rounded MLC leaf ends and leaf offsets.

Test patterns have been developed to allow accurate determination of several of these modeling parameters by iterative adjustment to agreement with measured profiles.[3] Values for the groove width and MLC transmission can be determined by matching profiles from measurement and from the TPS. Similar tests were carried out to determine the curvature of rounded MLC leaf ends and interleaf leakage. The resulting beam model was applied to a clinical head-and-neck IMRT dose distribution, using film measurement to demonstrate good agreement (mostly within 4%/4 mm), with the error partly attributable to over-response of the film in low-dose regions. No attempt was made to quantify the clinical or dosimetric benefits of accurately determining these parameters and so did not confirm their clinical impact on the dose distribution.

Williams and Metcalfe[5] describe, for Pinnacle v7.4f, the effect on penumbra width and profile modeling of changes in specific modeling parameters such as leaf-end radius of curvature and intraleaf transmission. They applied an optimized model to a clinical IMRT field to demonstrate the effects of tongue-and-groove width and rounded leaf-end modeling on a sample IMRT plan. They used gamma analyses (2 mm/3%) to quantify these effects for the two software versions as compared with a dose-to-water image obtained with an electronic portal imaging device and found improved accuracy to 2.2% from 7.0% pixel failure with the new version, although how much of the improvement was attributable to the improved modeling of the

rounded MLC leaf ends and how much to the addition of the groove width to the model was unclear. Absolute dose calculation also demonstrated improved accuracy as compared with measurement.

Although much work has been done to optimize current beam modeling facilities and to move toward accurate measurement of the required parameters, the necessary work—in terms of its effect on clinical plans—has yet to be quantified. The aim of the present study was to evaluate the dosimetric impact of the improvements to the photon beam model description within Pinnacle and to identify the parameters to which the plans are most sensitive (warranting the most attention and being less amenable to compromise when modeling). We aimed to assess the individual impact of each of the previously described parameters on a series of clinical plans so as to determine which ones have a significant clinical impact. The study was carried out for both an Elekta Precise and a Varian 600C/D LINAC. Head-and-neck IMRT plans were chosen, because they represent more complex cases, in which the impact of the experimental changes is likely to be greater. The intention was that this work provide valuable information on which model parameters are required to be accurately modeled for IMRT treatments and the required sophistication of the beam modeling software.

II. MATERIALS AND METHODS

A thorough explanation of the beam modeling process and the sequence followed by the Pinnacle TPS auto-modeling software has been given by Starkschall et al.[6] Although the modeling software has evolved since that publication, the methodology and most of the parameters used have remained unchanged. Only the modifications are described here.

The flattening filter attenuates and hardens the beam to a greater extent in the beam's center. The attenuation has therefore previously been modeled as a subtraction of an inverted cone from the beam fluence, where the depth and radius of the cone can be specified. The attenuation can now be modeled as an arbitrary profile with radial symmetry. Fig. 1 illustrates both approaches.

The presence of an MLC can be designated by the user, whereupon a range of associated parameters, such as the tongue-and-groove width and interleaf leakage, are now available for definition. In regions below a full MLC leaf, a user-defined MLC transmission factor is used to appropriately attenuate the initial beam intensity to that point. For regions at the field edges affected by the tongue-and-groove effect, the intensity is reduced by half of that to the full leaf, and for regions between two adjacent leaves, the user-defined leakage value applies a correction factor to the initial intensity. For regions beneath the jaws, the intensity is attenuated according to the appropriate jaw transmission factor, which can now be assigned separately for the x and y jaws, an improvement on previous software releases, when an average value was required.

For regions underneath the leaf tip, the radius of curvature of the rounded MLC leaf tip, leaf offset position, and the effective attenuation coefficient are used to calculate the relative transmission through the leaf tip compared with that through the full leaf, as described in detail by Cadman et al.[3] In the present work, the radius of curvature of the rounded MLC leaf ends and the leaf offset table were initially taken from values recommended by the Pinnacle planning system. The leaf offset table was then adjusted to better fit the measured data.

The individual impact of the various additions to the beam modeling facility were evaluated by commissioning 7 beam models, each with additions to a baseline model derived by the auto-modeling software:

- Model i: an original baseline model, including none of the enhancements to the beam model
- Model ii: incorporates only the tongue-and-groove effect, using the recommended leaf width of 0.1 cm

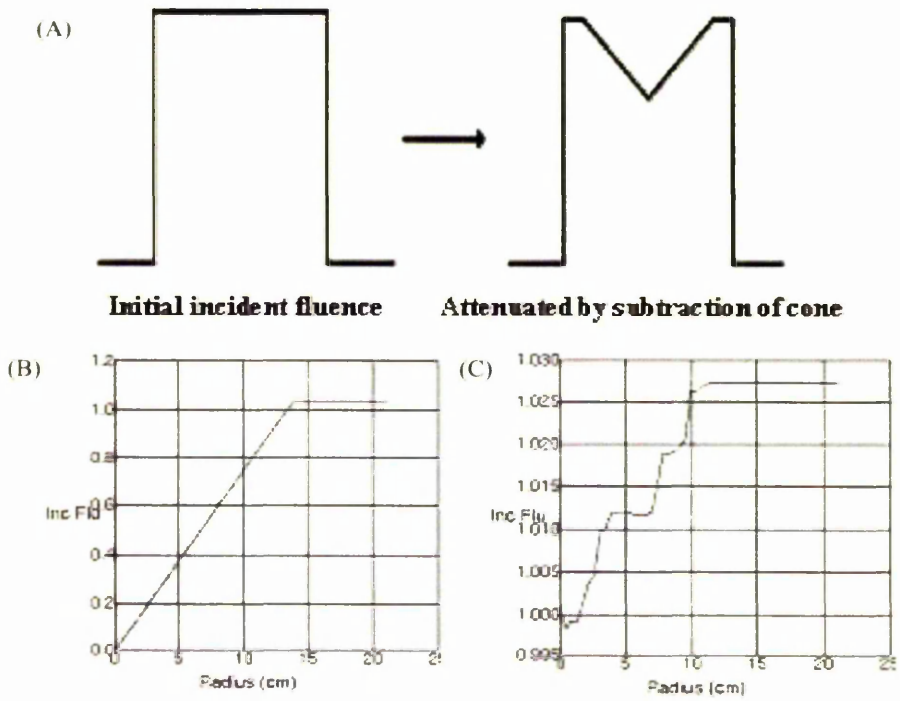


FIG. 1. (A) A radially symmetric volume is subtracted from the initial fluence to account for the flattening filter. It can be (B) a uniform cone, or (C) the Pinnacle v7.4f treatment planning system provides the option of an arbitrary profile.

- Model iii: incorporates 2% interleaf leakage
- Model iv: allows for the independence of x and y jaw transmissions, with optimized values (x transmission of 0.08, y transmission of 0.02) as compared with a value used in the baseline model (x and y transmission of 0.02)
- Model v: improves modeling of the flattening filter, using a radially-symmetric arbitrary profile
- Model vi: accurately models the rounded MLC leaf ends and optimized leaf offsets
- Model vii: a combined model that incorporates all the modifications

The beam model investigated on the two LINACs involved a 6-MV beam. The present work should highlight any differences in the effects of the modeling changes between the two LINACs—changes that would likely be attributable to the changes in the IMRT segment delivery process illustrated in Fig. 2. On the Elekta machine, the jaws track the leaf motion for each segment, but on the Varian, the jaws remain fixed throughout beam delivery. As a result, for Varian delivery, almost all segments are defined by the rounded leaf ends and the leaf edges. The radius of curvature of the rounded leaf ends is set to 15 cm for the Elekta and to 12 cm for the Varian, in line with manufacturer-recommended values. Because the jaws always remain outside the target volume for the Varian, model iv was omitted for that machine.

Each model (i – vii) was commissioned, and the resulting dose distribution calculated on a 2-mm dose grid for a series of 5 head-and-neck IMRT patient plans on a semi-anthropomorphic Perspex head-and-neck phantom that is routinely used in the verification of clinical head-and-

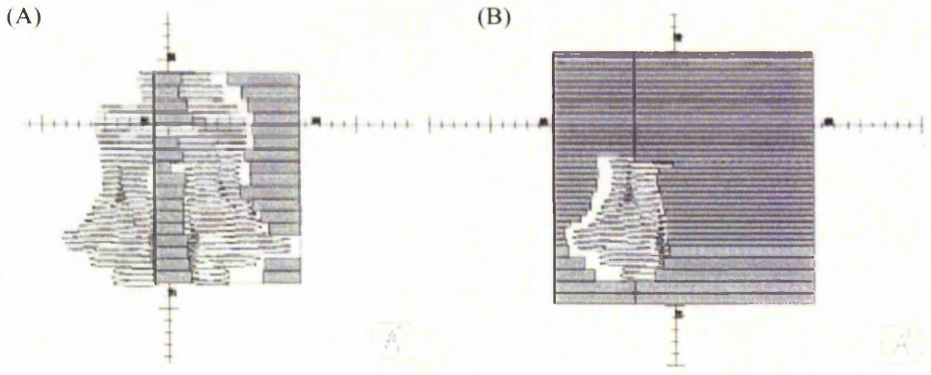


FIG. 2. Typical segments defined for delivery on (A) the Elekta Precise (Elekta, Stockholm, Sweden) linear accelerator, in which the jaws track the leaf positions, and (B) the Varian 600C/D (Varian Medical Systems, Palo Alto, CA) linear accelerator, in which the jaws remain fixed as the leaf-defined aperture moves between segments.

neck IMRT treatments. To be able to examine the effect of these changes in regions of high dose, low dose, high dose gradient, and at a distance from the central axis, coronal dose planes were acquired through the phantom at the level of the spinal cord on Pinnacle. In-house gamma analysis software written in IDL[7] was then used to calculate the location and extent of any changes to the dose distribution resulting from the individual (ii – vi) and combined (vii) model changes. Each plan used 5 IMRT fields at 6 MV with a total of 65 – 82 segments and was planned according to the departmental protocol for the Parsport trial, which delivers a dose of 65 Gy to the tumor clinical target volume and involved nodes and 54 Gy to the elective nodes, and which attempts to spare the contralateral parotid.

To confirm that the changes to the beam model resulted in improved dose calculation accuracy, the baseline (i) and modified (vii) models for the Elekta were then used to calculate individual beam fluence maps for comparison with measured Kodak extended dose range (EDR2) film for the same cohort of patients for every field. This analysis was conducted using routine clinical practice, wherein each fluence map is analyzed with the 100% isodose normalized to fall within the high-dose region. Similar work has been done for a single IMRT field,[5] but has not been established over a number of clinical fields. Because access to film processing equipment was discontinued during the study, a comparison of the Varian model with film was not carried out.

III. RESULTS

Fig. 3 illustrates the effect on the coronal dose distribution for the individual and collective modifications to the Elekta Precise beam model for a typical clinical head-and-neck IMRT plan copied to a verification phantom. Table 1 (Elekta Precise) and Table 2 (Varian 600C/D) present the gamma analysis results for the sample of 5 patients. The results in these tables are given as percentage pixel failures either within high-dose regions (>80% isodose) or for the overall dose distribution (>20% isodose). The tolerances for the various parameters were chosen to highlight the extent and, in Fig. 3, the location of significant variation. (A common high or low tolerance would have caused some gamma maps to be either empty or saturated as a result.)

The results suggest that the effect of interleaf leakage on planning would be negligible, except directly beneath the leaf gap. The independence of jaw transmission values in the Elekta model has a small impact at the 1%/1 mm level. Fig. 3(A) illustrates the addition to the Elekta

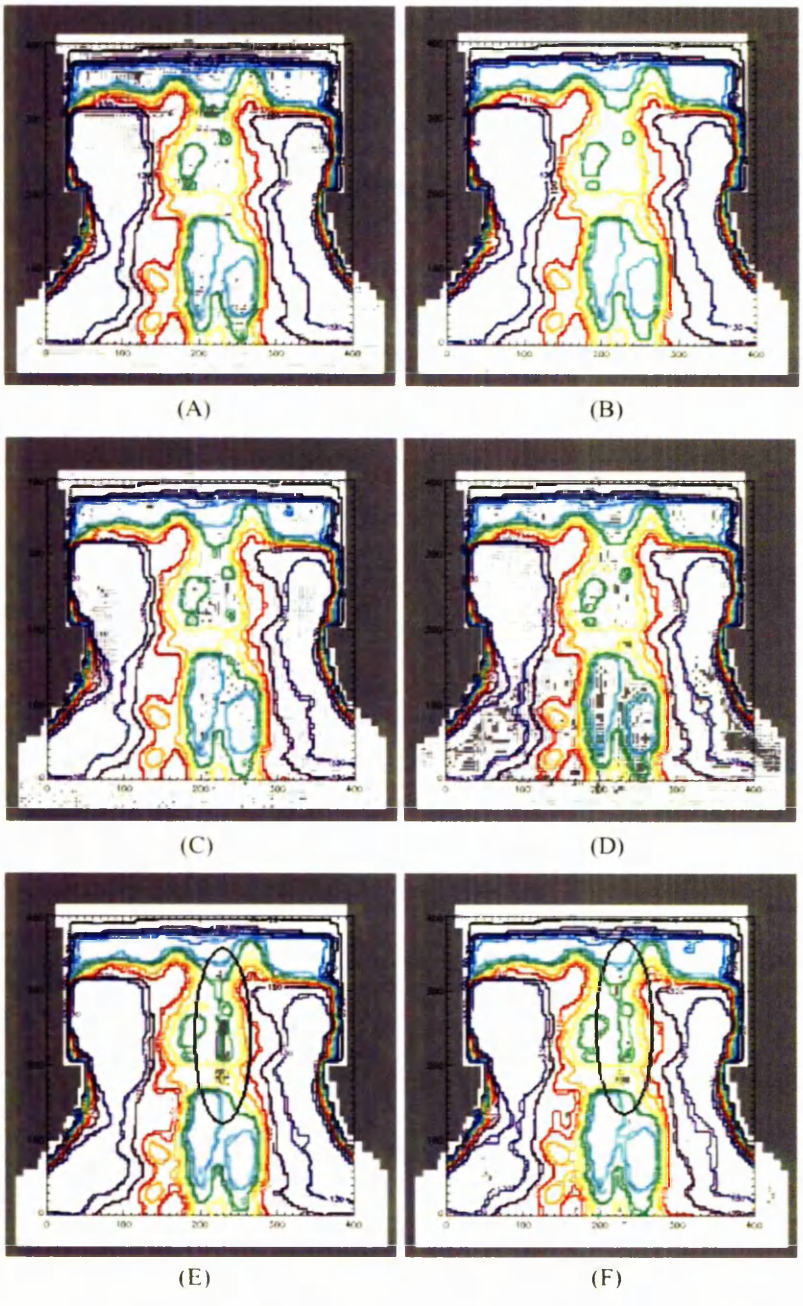


FIG. 3. Typical gamma evaluation maps for coronal slices at the level of the spinal cord as compared with the baseline model: (A) tongue-and-groove width at 1%/1 mm; (B) interleaf leakage at 1%/1 mm; (C) independent jaws at 1%/1 mm; (D) arbitrary profile for flattening filter at 1%/1 mm; (E) full multileaf collimator model at 3%/3 mm, with circled region of pixel failure; and (F) all changes at 3%/3 mm. Regions of pixel failure are shown in grey; white represents pixels within the defined tolerance. The 80% isodose is shown in sky blue and was selected to include doses above the approximate level of the spinal cord tolerance. The 20% isodose, which encompasses almost all of the dose distribution, is shown in black.

model of the tongue-and-groove effect, which is found to have only a small effect on the dose distribution, predominantly in the superior region of the plan (mean of 4.9% of pixels altered at 1%/1 mm tolerance). For the Varian model, this effect is more widespread (24.6% of pixels altered at 1%/1 mm tolerance). The modeling of the flattening filter as a radially symmetric profile rather than a simple cone is found to have a consistent low-level effect (<2%/2 mm) throughout the dose distribution for both LINACs. Fig. 3(E) has been calculated for the Elekta with a 3%/3 mm tolerance and demonstrates the significant impact of the accurate modeling of the rounded MLC leaf ends in high dose gradient regions such as the spinal cord (mean of 1.5% of pixels altered at 3%/3 mm tolerance). The close agreement of this image with the gamma image resulting from the addition of all modifications [Fig. 3(F)] confirms that accurate modeling of the rounded leaf ends is the dominant effect. This finding is also the case for the Varian model, although to a slightly lesser extent.

TABLE 1. Percentage of pixels failing to meet gamma analysis limits, as compared with the baseline model (i) for the Elekta Precise (Elekta, Stockholm, Sweden) linear accelerator[a]

		Isodose			
		80%		20%	
Model		Mean	SD	Mean	SD
At 1%/1 mm					
(ii)	Tongue-and-groove effect	4.9	1.2	5.9	1.8
(iii)	Interleaf leakage	0.0	0.0	0.1	0.1
(iv)	*x* and *y* jaw transmission	3.8	2.1	4.0	1.9
(v)	Arbitrary flattening filter profile	12.6	3.7	11.8	3.7
At 3%/3 mm					
(vi)	Full MLC model	1.5	1.0	1.6	0.9
(vii)	All changes	1.6	1.0	1.4	0.8

[a] For models ii – iv, changes were observed only at the 1%/1 mm level, suggesting that the clinical impact of accurately modeling these parameters is limited, even for complex treatments such as intensity-modulated radiotherapy for the head and neck. The larger changes to the dose distribution (apparent in model vi at 3%/3 mm) are attributable to the precise modeling of the shape and position of the rounded multileaf collimator (MLC) leaf ends.

SD = standard deviation; MLC = multileaf collimator.

TABLE 2. Percentage of pixels failing to meet gamma analysis limits, as compared with the baseline model (i) for the Varian 600C/D (Varian Medical Systems, Palo Alto, CA) linear accelerator[a]

		Isodose			
		80%		20%	
Model		Mean	SD	Mean	SD
At 1%/1 mm					
(ii)	Tongue-and-groove effect	25.5	2.0	22.8	2.7
(iii)	Interleaf leakage	0.0	0.0	0.0	0.0
(iv)	*x* and *y* jaw transmission	N/A	N/A	N/A	N/A
(v)	Arbitrary flattening filter profile	13.8	3.5	12.0	1.9
At 3%/3 mm					
(vi)	Full MLC model	0.3	0.1	0.3	0.2
(vii)	All changes	0.8	0.4	1.1	0.8

[a] The independence of the jaws (iv) was not investigated for the Varian linear accelerator because the jaws always remain outside of the overall fluence. The tongue-and-groove effect (ii) was found to be more significant for the Varian than for the Elekta Precise (Elekta, Stockholm, Sweden). The effect of modeling the rounded leaf ends was diminished slightly, but remained significant.

SD = standard deviation; MLC = multileaf collimator.

Fig. 4 demonstrates that, for all dose levels, the combined effect of changes to the Elekta modeling parameters is to improve agreement with measured dose distributions. The improvements are significant ($P < 0.05$) at the medium dose level for errors greater than 2%/2 mm, which is expected, because this level incorporates a higher proportion of high-gradient areas in

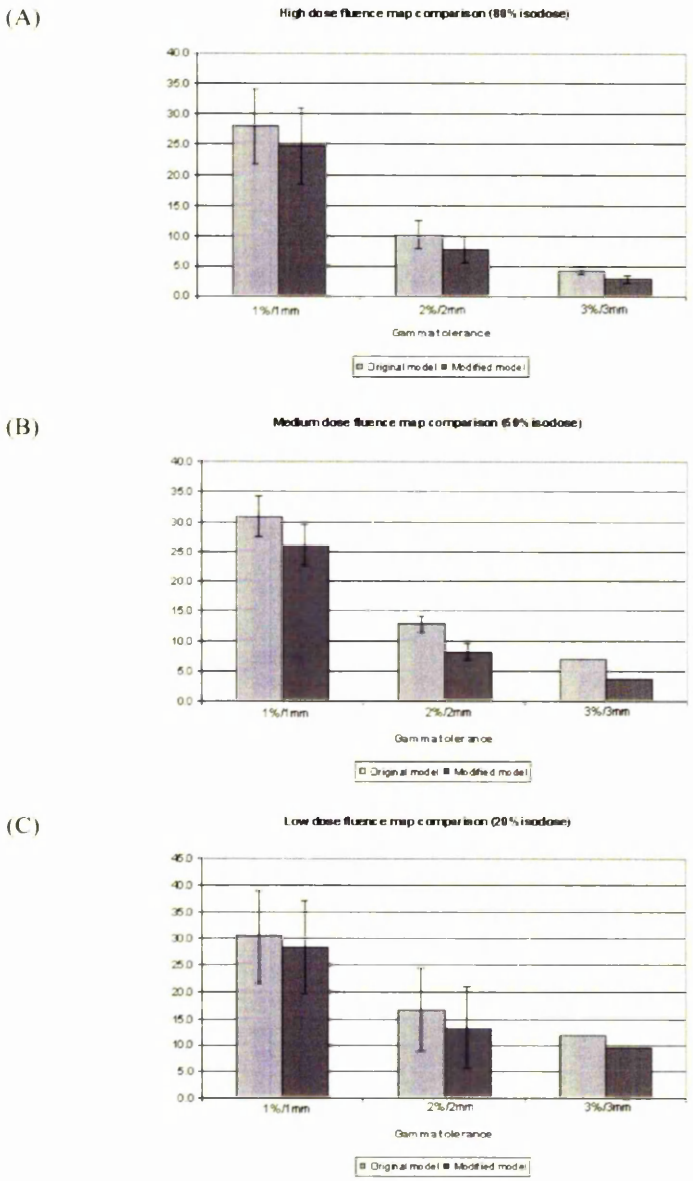


FIG. 4. Comparison of the baseline (i) and fully modified (vii) models relative to measured fluence maps for 5 clinical head-and-neck intensity-modulated radiotherapy plans (25 fields) for the Elekta Precise (Elekta, Stockholm, Sweden) linear accelerator, indicating the improved accuracy of the modified model.

which the effects of errors in the penumbra are more prominent. Significant improvement of the larger errors (3%/3 mm) is also seen in the high-dose region. Given the limited sample size, the improvements were not found to be as significant within the high- and low-dose regions, probably because these regions are predominantly low-gradient areas of the dose distribution.

IV. DISCUSSION

The improvements to the beam modeling facilities in Pinnacle v7.4f have a statistically significant impact on the verification accuracy of IMRT plans for the head and neck. The addition of interleaf leakage values was found not to affect the resulting dose distribution, and values can be easily taken from a manufacturer's machine information. Accurate modeling of the tongue-and-groove effect was found to be more significant for the Varian 600C/D than for the Elekta Precise. That finding was expected, because for the Elekta delivery, the superior region contains a higher proportion of field edges defined by the combination of jaw and leaf edges as opposed to leaf edges alone. Other parts of the dose distribution result from the number of small, low-weighted segments. Because of the static jaws in the Varian LINAC's delivery, almost all segments are defined by the leaf edge only, and so the introduction of the groove width has greater impact. The accurate modeling of the rounded MLC leaf ends is found to be of great importance to the dosimetric accuracy of the IMRT plan calculation. It is recommended that accurate modeling of rounded leaves be the focus of adjustments to the beam model when auto-modeling has been carried out to set other parameters. The accuracy of the resulting beam model significantly and consistently improves agreement with measured fluence films.

Although this study focused on the Pinnacle TPS, its results highlight the need for any TPS to correctly model the penumbra of the MLC and jaws to enable good dosimetric agreement for head-and-neck IMRT treatments. The dose to critical structures such as the spinal cord is often defined by the superposition of multiple penumbrae in such treatments. The addition of the other parameters described is of limited clinical importance.

The ability of Pinnacle to accurately model the dose would be less accurate at higher energies (for example, 18 MV) because of relatively small field sizes and the problem of modeling electron transport at air–tissue interfaces in the head-and-neck region. Such modeling could be more accurate with Monte Carlo simulations, for example—although it is questionable whether treatment of head-and-neck sites would be improved by the use of such high photon energies.

V. CONCLUSION

The accurate modeling of field penumbrae, particularly penumbrae defined by the rounded MLC ends, is of great importance to the commissioning of a TPS for the clinical implementation of head-and-neck IMRT and other sites for which the dose to critical structures is derived from the addition of many beam penumbrae. Current commercially available TPSs can offer sufficient accuracy for this treatment technique (within 3% – 4%),[8] but the user must be aware of the system's limitations, both in terms of the dosimetric capabilities and the effects of the compromises made during the modeling process.

REFERENCES

1. Schwarz M, Bos LJ, Mijnheer BJ, Lebesque JV, Damen EM. Importance of accurate dose calculations outside segment edges in intensity modulated radiotherapy treatment planning. Radiother Oncol. 2003;69(3):305–314.

2. Fraass B, Doppke K, Hunt M, et al. American Association of Physicists in Medicine Radiation Therapy Committee Task Group 53: quality assurance for clinical radiotherapy treatment planning. Med Phys. 1998;25(10):1773–1829.
3. Cadman P, McNutt T, Bzdusek K. Validation of physics improvements for IMRT with a commercial treatment planning system. J Appl Clin Med Phys. 2005;6(2):74–86.
4. Cadman P, Bassalow R, Sidhu NP, Ibbott G, Nelson A. Dosimetric considerations for validation of a sequential IMRT process with a commercial treatment planning system. Phys Med Biol. 2002;47(16):3001–3010.
5. Williams MJ, Metcalfe P. Verification of a rounded leaf-end MLC model used in a radiotherapy treatment planning system. Phys Med Biol. 2006;51(4):N65–N78.
6. Starkschall G, Steadham RE Jr, Popple RA, Ahmad S, Rosen II. Beam-commissioning methodology for a three-dimensional convolution/superposition photon dose algorithm. J Appl Clin Med Phys. 2000;1(1):8–27.
7. Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. Med Phys. 1998;25(5):656–661.
8. Ezzell GA, Galvin JM, Low D, et al. Guidance document on delivery, treatment planning and clinical implementation of IMRT: report of the IMRT subcommittee of the AAPM Radiation Therapy Committee. Med Phys. 2003;30(8):2089–2115.

Publication #3

Evaluation of the impact of dental artefacts on intensity-modulated radiotherapy planning for the head and neck

Accepted for publication by Radiotherapy & Oncology (28th August 2009)

Authors:

Webster GJ

Rowbottom CG

Mackay RI

20 pages

Abstract

Background and purpose

High density materials create severe artefacts in the computed tomography (CT) scans used for radiotherapy dose calculations. Increased use of intensity-modulated radiotherapy (IMRT) to treat oropharyngeal cancers raises concerns over the accuracy of the resulting dose calculation. This work quantifies their impact and evaluates a simple corrective technique.

Materials and Methods

15 oropharyngeal patients with severe artefacts were retrospectively planned with IMRT using two different CT/density look-up tables. Each plan was recalculated using a corrected CT dataset to evaluate the dose distribution delivered to the patient. Plan quality in the absence of dental artefacts was similarly assessed. A range of dosimetric and radiobiological parameters were compared pre and post-correction.

Results

Plans using a standard CT/density look-up table (density $\leq 1.8 g/cm^3$) revealed inconsistent inter-patient errors, mostly within clinical acceptance, although potentially significantly reducing target coverage for individual patients. Using an extended CT/density look-up table (density $\leq 10.0 g/cm^3$) greatly reduced the errors for 13/15 patients. In 2/15 patients with residual errors the CTV extended into the severely-affected region and could be corrected by applying a simple manual correction.

Conclusions

Use of an extended CT/density look-up table together with a simple manual bulk density correction reduces the impact of dental artefacts on head and neck IMRT planning to acceptable levels.

Introduction

High density materials such as dental amalgams can lead to extensive streaking artefacts on computed tomography (CT) scans. These artefacts obscure the underlying anatomy, leading to uncertainty in the delineation of target volumes and compromising the integrity of the density representation that is crucial for accurate dose calculation. Concern exists over the impact this may have on the plan optimisation and dose calculation accuracy of head and neck intensity-modulated radiotherapy (IMRT) plans and whether artefact-correction techniques are required.

Historically, concern over the effect of dental amalgams has been limited since bilateral head and neck disease has conventionally been treated with opposed parallel pairs of beams, using homogeneous calculations, which do not pass through the regions severely affected by artefacts. This is changing with the widespread implementation of IMRT for head and neck, which often consists of radiation delivery from anterior beams passing directly through the affected region, and often the amalgams themselves, prior to reaching the target volume or sensitive structures.

A previous study into the impact of dental artefacts on IMRT planning by Kim et al (2006) evaluated the impact of a range of automated corrections: an IMRT plan was created based on the CT dataset after a gold-standard correction consisting of replacement of corrupted regions of the dataset by either bone or water. This plan was then transferred to the same dataset after having undergone a range of inferior corrections (including no correction) and recalculated. The study predicted a detrimental impact on radiobiological outcome due to observed cold spots in the target volume and hot spots in sensitive structures such as the parotid glands [1].

However, any plan recalculated using a dataset on which the optimisation was not based would be expected to result in a degradation of plan quality. The potential clinical problem is that optimisation based on uncorrected CT datasets could hide detrimental hot or cold spots, which

would become apparent when the plan is delivered to the patient. To quantify this effect, the IMRT plans must be optimised on the original uncorrected dataset and then the delivered dose distribution recalculated on a corrected dataset. This was done in an earlier study by the same authors, although no radiobiological evaluation was carried out [2].

This work aims to address these clinical problems by quantifying the impact on both clinically-relevant and radiobiological parameters, aiming to identify when artefact corrections become necessary, and the efficacy of a simple correction technique. It will also examine the impact of modelling the amalgams themselves and the impact of different CT-to-density look-up tables.

Materials and methods

A Patient selection and delineation

15 consecutive oropharyngeal cancer patients planned using IMRT and subject to moderate or severe artefacts on multiple slices of the CT dataset were selected for this study. The most severely affected slice for each case is illustrated using clinical window settings (W = -624, L = 224) in Figure 1.

Amalgams were outlined by saturating the display (window = -1023, level = 3005) and delineating the visible high density structures. Teeth were outlined at a high level (W = 70, L = 947) to avoid the artificially large volumes that would be delineated using clinical settings and to give a more accurate anatomical representation of dentition. Clinical settings were then used to outline regions of soft tissue affected by the artefacts.

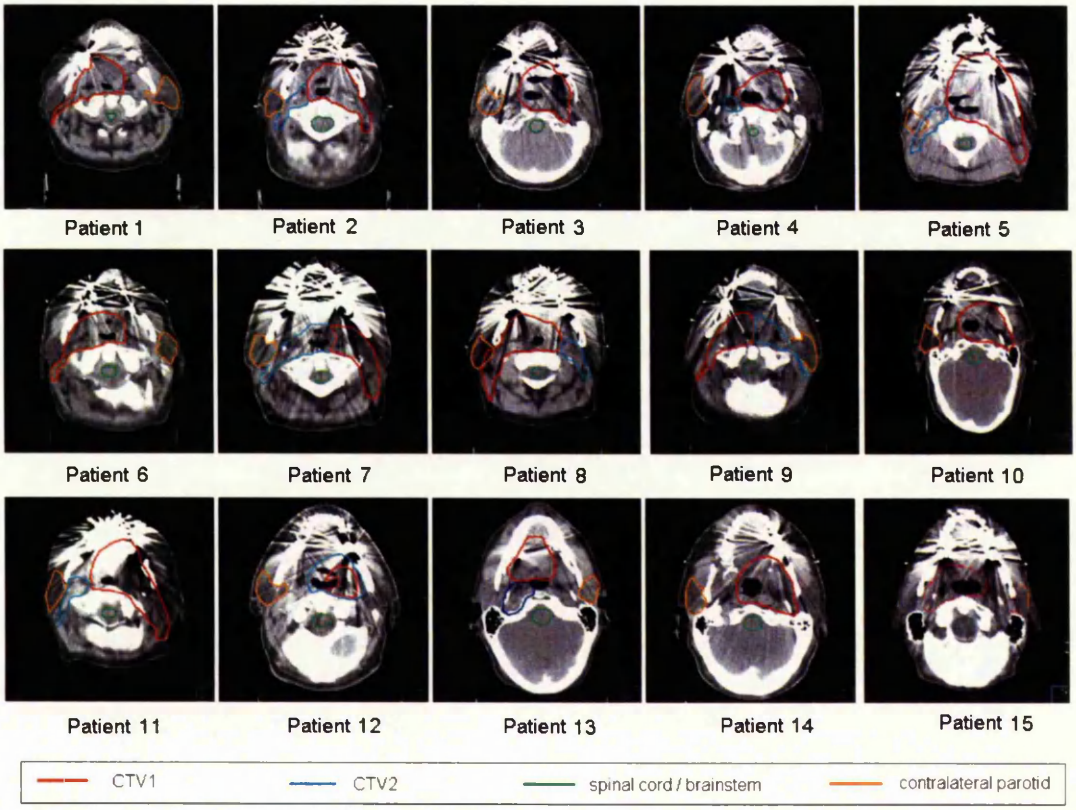


Figure 1. Target volumes and dental artefacts for all patients involved in this study (most severely affected slice is shown)

B Plan optimisation and calculation

All plans were optimised using direct machine parameter optimisation (DMPO) software, which incorporates the deliverable fluence into the optimisation and removes plan degradation due to subsequent segment conversion, in the Pinnacle Treatment Planning System (TPS) v8.0m (Philips, Madison, WI) on a 0.2cm resolution dose grid. Seven 6MV coplanar beams, incorporating a direct anterior beam, were used in all cases to generate 55-60 step-and-shoot segments based on a simultaneous integrated boost of 65Gy to PTV1 and 54Gy to PTV2 in 30 fractions. Each optimisation was carried out using both the local standard CT/density look-up table commissioned using the Gammex RMI tissue characterisation phantom, which saturates at

1.8g/cm^3 (Method A), and an extended table that allowed higher densities to be distinguished (Method B). These CT/density tables are shown in Figure 2. It should be noted that unless a density override is specifically applied when using Method B, the image will saturate at CT number 3072 (density ~3.12g/cm^3) due to the limited dynamic range of the CT-scanned data in the Pinnacle TPS. This will mean that amalgam density is likely to be under-estimated.

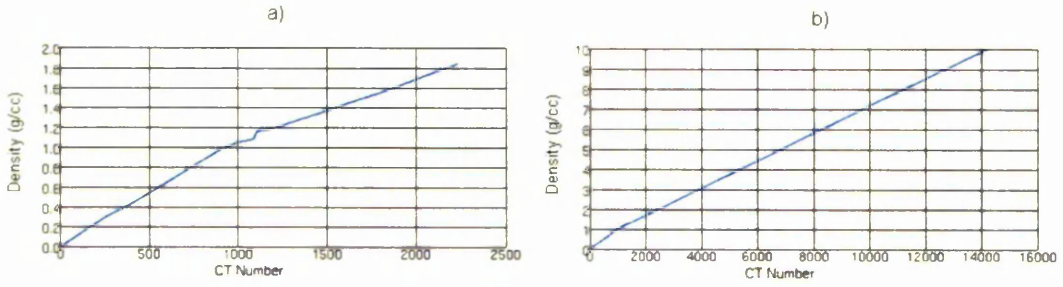


Figure 2. a) Standard clinical and b) extended CT/density look-up tables. The extended table will allow high densities to be incorporated into the calculation of effective path length. [N.b. Pinnacle CT numbers operate in the range 0-4096, effectively adding 1024 to standard CT numbers]

Subsequent recalculation of the plan was then performed on a corrected CT dataset. This corrected CT dataset was generated using manually-outlined bulk density corrections for tissue (1.0g/cm^3) and teeth (1.8g/cm^3), both representing mean values from 4 extraneous patient CT scans that were not subject to artefacts, and amalgam (10g/cm^3), an average value used for a similar study into dose calculation accuracy in the presence of high density dental amalgams [3]. Dose calculations on this corrected dataset were performed using the extended CT/density look-up table.

In the event of a conclusion demonstrating an acceptably low impact of artefacts in terms of masking unwanted hot or cold spots in the dose distribution, indicating no need to apply corrections prior to planning, it would be pertinent to ask what degradation of plan quality results when planning on a dataset with severe artefacts compared to planning on one that has been corrected. Therefore, each plan recalculated using Method B was compared to a third plan

optimised with all density overrides applied and using the extended CT/density table (Method C), This represents a comparison of the achievable plan when all artefacts are removed and the high density amalgam is accounted for in the calculation of effective path length (Method C), against that achieved in the presence of streak artefacts and in which very high densities are ignored (Method B).

C Plan evaluation

Dose changes to spinal cord and brainstem were evaluated using calculated dose-volume histograms (DVH) to determine the maximum point and 1cc doses. Mean dose to the contralateral parotid gland was also recorded. Dosimetric evaluation of target volumes was carried out using several parameters: the dose to 95% of the volume of both CTV1 and CTV2; the dose to the coldest 1cc volume for both CTV1 and CTV2; the dose to the hottest 1cc volume and maximum point dose for CTV1.

Radiobiological modelling was used to obtain comparisons, with tumour control probability (TCP) determined for both CTV1 and CTV2 and normal tissue complication probability (NTCP) calculated for the contralateral parotid gland. TCP calculations for both CTV1 and CTV2 are based on the following formula (adapted from [4] to exclude inter-patient heterogeneity of response):

$$TCP = \prod_i e^{\left[-\rho_{cl} v_i e^{\left[-\alpha D_i \left(1 + \frac{d_i}{\alpha / \beta} \right) \right]} \right]}$$

where the differential dose, D_i, and volume, v_i, are taken directly from differential dose-volume histograms in Pinnacle (bin size 10cGy); $d_i = D_i/n$, where n = 30 fractions; $\alpha = 0.35$; $\alpha/\beta = 10Gy^{-1}$; clonogen density, ρ_{cl}, is optimised for each volume to give TCP values near 50%, maximising the impact of any dose deviations and ensuring that any corrections to these parameters to approach a more realistic clinical scenario would act to reduce the observed impact of dental artefacts.

NTCP for the contralateral parotid gland was evaluated using the Lyman-Kutcher-Burman formalism [5]. This formalism defines a power law to calculate complication probability based on uniform irradiation of a specified volume of the organ.

$$NTCP = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-t^2/2}$$

where

$$t = \frac{D_{ref} - TD_{50}(v)}{m * TD_{50}(v)}$$

$TD_{50}(v)$ is the dose at which a 50% complication rate is seen for uniform irradiation of the partial volume v and is derived from

$$TD_{50}(1) = TD_{50}(v) * v^{n}$$

based on the 50% tolerance of uniform irradiation of the whole organ ($TD_{50}(1)$ = 28.4Gy). m is the steepness of the dose response curve (m = 0.18) and n represents the volume effect (taken as 1, to represent the mean dose). These values were established by Eisbruch et al [6]. Individual DVHs for each plan were used to establish effective partial volumes irradiated to the maximum dose to the organ (D_{ref}) with equivalence to the actual differential DVH (D_i,v_i) according to the formula

$$V_{eff} = \sum \left(\frac{D_i}{D_{ref}} \right)^{\frac{1}{n}} v_i$$

The statistical significance of any changes in dosimetric and radiobiological parameters was determined using a 2-tailed non-parameteric Wilcoxon matched pairs signed rank test.

D Plan validation

The use of bulk density overrides to correct the effects of high density artefacts is not a perfect solution. To quantify the inaccuracies associated with this methodology we would ideally acquire the same image datasets with and without artefacts and compare the resulting optimisations over a patient population.

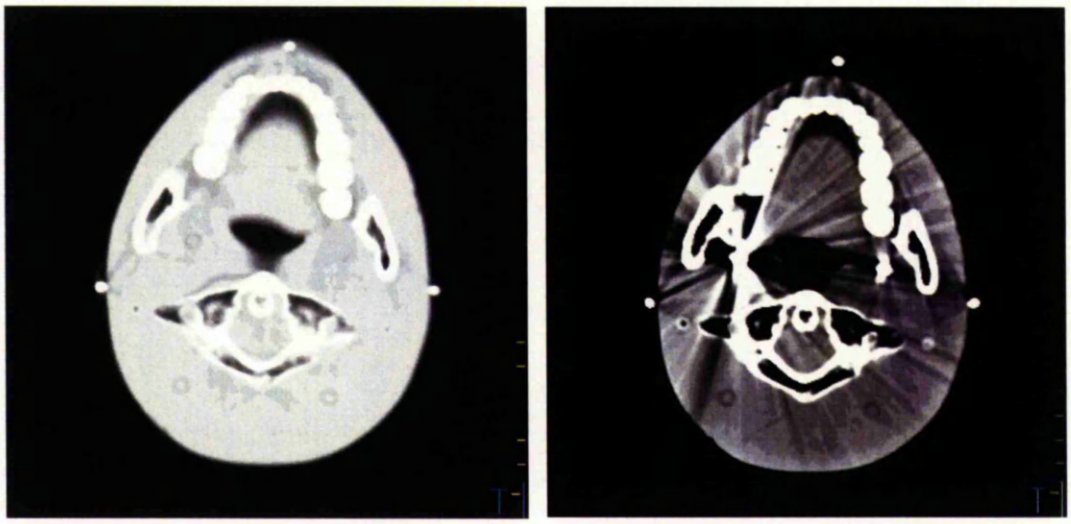


Figure 3. CT scans of anthropomorphic phantom without (left) and with (right) high density implants, used for validation of bulk density correction technique

Obviously this is not possible and so a phantom study has been performed. 2 CT images were acquired of the head and neck region of the RANDO phantom (originally produced by Alderson

Research Laboratories, Stamford, CT): a scan of the phantom itself (Scan1); an identical scan that included 2 high density (~10g/cm^3) implants (Scan2). The resulting images are shown in Figure 3.

Clinical volumes from a representative patient were transferred to the phantom and manually adjusted to fit the anatomy. An IMRT plan was created on both Scan1 (corresponding to Method B) and Scan2 (corresponding to Method C, i.e. using the extended CT/density look-up table and the bulk density corrections under investigation). The two plans were compared to quantify the discrepancies inherent in our correction technique.

Results

Method A - Standard CT/density table

When using the standard CT/density table, dosimetric changes to clinically-relevant parameters for sensitive structures observed on recalculation with the corrected dataset are predominantly detrimental and in some cases statistically significant, but of low magnitude (all changes to spinal cord and brainstem hotspots <0.7Gy, all changes to parotid mean dose <0.3Gy) and unlikely to influence treatment decisions. However, the presence of artefacts results in substantial changes to target volume parameters, indicated in Figure 4a. Statistically significant increases in CTV1 hotspots and CTV2 coverage were observed, leading to increased average TCP, although individual patients showed decreases in TCP (up to 2%). No correlation was seen between the observed error and the artefact severity or position relative to the CTV.

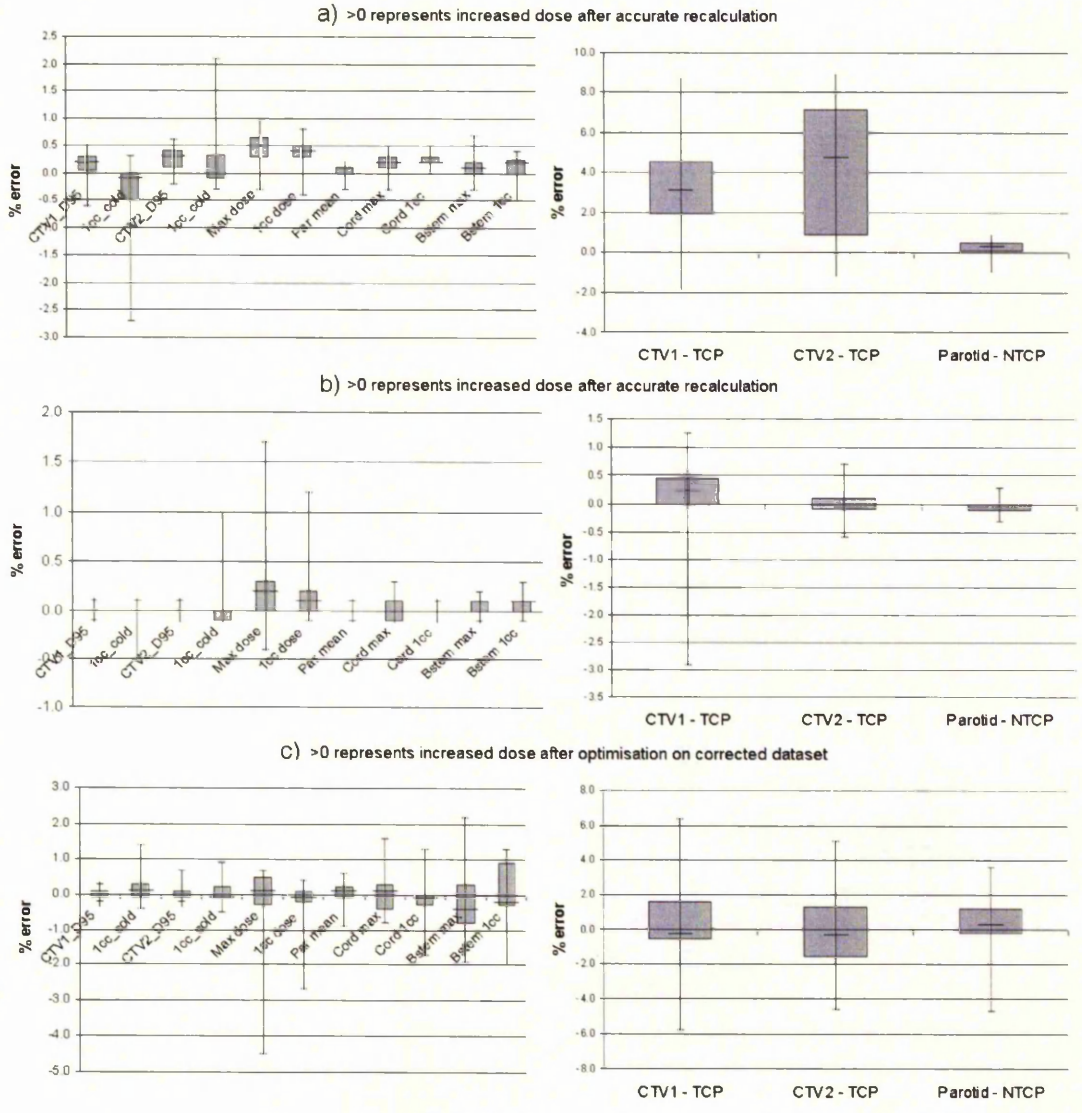


Figure 4. Dosimetric and radiobiological changes due to artefact-correction after optimisation with a) the standard clinical CT/density table b) the extended CT/density table for 13/15 patients with no anterior-encroaching CTV and c) when comparing the latter to initial optimisation on a fully-corrected image dataset for the 13 patients for which Method B is recommended for clinical use (boxes represent upper and lower quartiles, bars show maximum median and minimum values)

Method B - Extended CT/density table

Initial optimisation using the extended CT/density look-up table led to a reduction of the errors seen using Method A for all parameters, with statistically significant discrepancies observed only in the CTV1 hotspots. Across all dosimetric and radiobiological parameters, clinically significant dose discrepancies were observed for only 2 patients: referring to Figure 1, Patients 5&11, showing increased TCPs for CTV1 of 1.0% and 13.7% respectively, with clinically unacceptable hotspots within CTV1 becoming apparent only after recalculation with the corrected CT dataset. No significant impact on the dose to sensitive structures was seen (<0.3Gy).

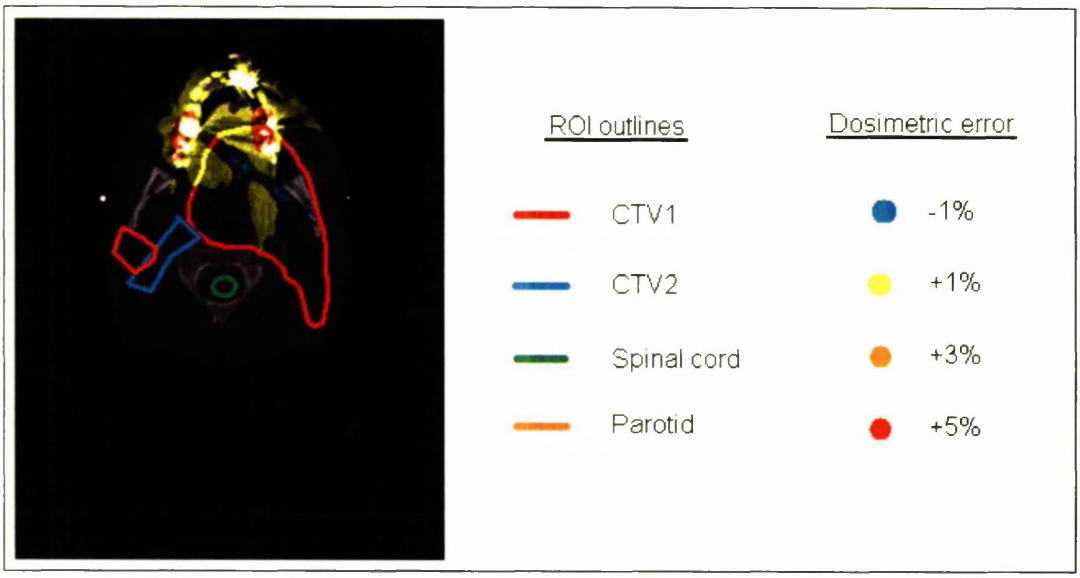


Figure 5. Extreme example (Patient 5) of the impact of dental artefacts on the IMRT dose distribution when using the extended CT/density look-up table (Method B). Positive errors represent hotspots that are not apparent on the initial plan, while negative errors represent hidden coldspots.

Figure 1 suggests that the error seen for patients 5&11 could be due to encroachment of CTV into the region severely affected by image artefacts, the extent of which is not seen in the other patient scans. Quantitative analysis of the remaining 13 patients show no statistically significant changes to any dosimetric or radiobiological parameters (Figure 4b). The discrepancies in the

dose distribution for a severely-affected patient are shown in Figure 5, indicating that the most affected regions are in close proximity to high density artefacts and the amalgams.

Method C - Comparison with artefact-free scan

Comparison of the plan optimised using the extended CT/density table and recalculated using bulk density corrections (Method B) with a plan initially optimised using bulk density overrides (Method C), shows good agreement with no statistically significant differences. The results, summarised in Figure 4c, show that all discrepancies within the target volume were <1Gy and those in organs-at-risk, whilst occasionally approaching 2Gy for the brainstem, were generally <1Gy.

Validation

Comparison of plans optimised based on the RANDO dataset with and without artefacts showed dosimetric discrepancies within 0.5Gy for all target volume parameters, with no observed trend towards a reduction in plan quality when using bulk density corrections to remove the effect of artefacts. Spinal cord dose was reduced, while an observed increase of 2.9Gy in brainstem 1cc dose did not approach clinical tolerance.

Discussion

Streak artefacts due to high density metal amalgams in CT scans can significantly impact on the accuracy of the resulting dose calculation. However, this is highly dependent on the CT/density look-up table used: for most patients, this effect is reduced to acceptable levels (<0.5%) by using an extended dataset that accounts for the high densities, with no manual density overrides required.

This study found the impact of streak artefacts to be minimal in 13/15 patients optimised using an extended CT/density look-up table and no bulk density corrections. Each of these patients demonstrated dosimetric consistency well within the local dose verification tolerance (3% agreement of absolute dose within the target volume, 4% within the spinal cord). The 2 patients in whom larger discrepancies were observed both involved encroachment of CTV into the oral cavity region where the dental artefacts were most severe. It is therefore suggested that patients requiring extra correction be identified in advance on the basis of CTV position relative to the severely-affected region. In these patients it is sufficient to apply a simple manual artefact correction using density overrides, as outlined in Method B.

Further investigation of the two problematic plans confirmed the necessity of outlining the amalgams, with significant discrepancies seen when optimisation was carried out using density overrides for only teeth and tissue. However, replanning using the larger teeth volume outlined when using clinical window settings resulted in negligible differences, indicating that the solution is robust to inter-observer variability in outlining of teeth and amalgams.

When using an extended CT/density look-up table it is important to be aware of inaccuracies likely to be encountered in the resulting dose distribution. The dynamic range of most CT scanners is not designed to operate accurately at such high densities and the amalgams may appear artificially large. This effect was quantified as part of a broader study by Jakel et al, finding an approximate doubling of size for a small metal implant within a cylindrical phantom (intended to represent a dental filling within a head), an effect exaggerated by the presence of multiple high density implants that are often encountered clinically [7]. The accuracy of the dose algorithm at these densities is also limited. Wieslander and Knoos quantified the calculation error of the collapsed cone convolution algorithm in comparison with Monte Carlo simulations after attenuation through high density hip prostheses, finding accuracy of 2-5% downstream of a 4cm thickness of high density material [8]. These errors are likely due to beam hardening and pair

production effects not being accurately modelled by the collapsed cone algorithm, which would be expected to be greatly reduced as the thickness of high density material decreases.

When comparing our proposed correction technique (Method B) with plans initially optimised using corrected scans (Method C), we found that although the mean changes to each parameter were low, improvements to the plan were generally indicated in the latter case. This is to be expected since in the former case we are recalculating based on a different dataset to that on which the optimisation was carried out. However, in no case were these differences felt to be clinically significant. Validation of the bulk density correction, using the RANDO phantom, found clinically insignificant changes, although the increased dose to the brainstem, which was also seen in 1/15 patients, is noteworthy. This is likely due to the ease with which the brainstem tolerance was met in each case, allowing increases in brainstem dose to be made at very little expense to the cost function used in the optimisation.

The NTCP model is limited by a lack of spatial information regarding the dose distribution, assuming a uniform response throughout the organ. This means that the likely reality of the parotid gland response based on a distribution of functional sub-units containing a serial component is ignored and could potentially alter the dose-response curve. However, the combination of consistent observed mean physical dose to the parotid in response to artefacts and the reliance on differences in NTCP values for similar plans rather than the absolute values themselves, means that the conclusions of this work are unlikely to be significantly affected by more sophisticated modelling.

Throughout this study it has been assumed that the patient anatomy is rigid and unchanging. In reality the daily setup of the patient will vary, resulting in a movement of the patient relative to the reference frame of the beam delivery from the linear accelerator and a resultant blurring of the localised impact of the artefacts. Other changes such as the position of the mandible may cause a similar blurring of the dose distribution, further minimising the errors associated with

dental artefacts. Prevost et al measured soft tissue movement of 1-2mm during delivery, which would act to reduce small hot and cold spots in the dose distribution [9]. The negligible differences seen in this study in the volumes of the coldest 1cc of the target volumes suggests that the observed changes in TCP are the result of changes to volumes less than this threshold, further reinforcing the assumption that blurring due to setup variation and intra-fraction motion would reduce their impact.

The impact of patient movement, alongside the likely overestimation of amalgam size due to both saturation of the Hounsfield unit scale and the partial volume effect due to the 5mm slice thickness of the CT scan, as well as the worst-case TCP parameters adopted, imply that the results of this study are likely to correspond to a worst-case-scenario for the patients analysed. The effect has been quantified in reference to Table.A1 in [10], which estimates the dosimetric uncertainty associated with the entire treatment process. We have used the maximum quartile deviations over all target volume parameters (except maximum point dose, for reasons outlined above) to quantify errors due to loss of both dose calculation accuracy (0.25% from Figure 4b) and plan quality (0.3% from Figure 4b) when using Method B to define our intended dose distribution, as well as the uncertainty inherent in our bulk density correction technique (based on the described phantom study yielding a maximum dose deviation of 0.4Gy ~ 0.6%). These give a combined uncertainty of 0.72%, which, when added to the 5.9% quoted by ICRP contributes an additional 0.02%. This indicates that the impact of dental artefacts when using an extended CT/density look-up table is not significant compared to the uncertainty associated with the overall treatment process.

Although this study demonstrates findings contrary to those of Kim et al, they are not necessarily contradictory as the earlier study appears to have made no attempt to model the amalgams themselves, suggesting that a CT/density look-up table with only a limited dynamic range was used.

In response to the results of this study, local practice for patients subject to dental artefacts being treated with IMRT now incorporates an extended CT/density look-up table into the optimisation. This method has the advantage of not requiring any additional outlining or density overrides on the CT dataset. For patients in whom the target region encroaches into a region severely affected by artefacts (in most cases the oral cavity), the amalgams are outlined on saturated window settings, while teeth and tissue are outlined using clinical settings and appropriate bulk density overrides applied.

Conclusions

The impact of dental amalgams on IMRT planning for the head and neck can, for the majority of patients, be minimised by using an extended CT/density table. Patients in whom this would be insufficient can be identified after target volume delineation and a simple manual correction applied.

References

1. Kim, Y. and W. Tome, *On the radiobiological impact of metal artefacts in head and neck IMRT in terms of tumour control probability (TCP) and normal tissue complication probability (NTCP).* Med Bio Eng Comput, 2007. **45**: p. 1045-1051.
2. Kim, Y., W. Tome, M. Bal, T. McNutt, and L. Spies, *The imapct of metal artefacts on head and neck IMRT dose distributions.* Radiother Oncol, 2006. **79**: p. 198-202.
3. Spirydovich, S., L. Papiez, M. Langer, G. Sandison, and V. Thai, *High density dental materials and radiotherapy planning; Comparison of the dose predictions using superposition algorithm and fluence map Monte Carlo method with radiochromic film measurements.* Radiother Oncol, 2006. **81**: p. 309-314.
4. Nahum, A. and G. Kutcher, *Biological evaluation of treatment plans*, in *Handbook of radiotherapy physics*, P Mayles, A Nahum, and JC Rosenwald, Editors. 2007, Taylor & Francis. p. p738.
5. Kutcher, G. and C. Burman, *Calculation of complication probability factors for non-uniform normal tissue irradiation: the effective volume method.* Int J Radiat Oncol Biol Phys, 1989. **16**(6): p. 1623-1630.
6. Eisbruch, A., R. Ten Haken, H. Kim, L. Marsh, and S. JA, *Dose, volume and function relationships in parotid salivary glands following conformal and intensity-modulated irradiation of head and neck cancer.* Int J Radiat Oncol Biol Phys, 1999. **45**(3): p. 577-587.
7. Jakel, O. and P. Reiss, *The influence of metal artefacts on the range of ion beams.* Phys Med Biol, 2007. **52**: p. 635-644.
8. Wieslander, E. and T. Knoos, *Dose perturbation in the presence of metallic implants: treatment planning system versus Monte Carlo simulations.* Phys Med Biol, 2003. **48**: p. 3295-3305.

9. Prevost, J., H. deBoer, J. Poll, P. Voet, and P. Levendag, *Analysis of the motion of oropharyngeal tumours and consequences in planning target volume determination.* Radiother Oncol, 2008. **87**: p. 268-273.

10. ICRP, *Prevention of accidents to patients undergoing radiation therapy.* 2000, International Commission on Radiological Protection (ICRP).

This page is intentionally blank

Publication #4

Evaluation of larynx-sparing techniques with IMRT when treating the head and neck

Published in International Journal of Radiation Oncology, Biology & Physics, 72(2), 2008 p617-622

Authors:

Webster GJ

Rowbottom CG

Ho KF

Slevin NJ

Mackay RI

16 pages

Abstract

Purpose

Concern exists that widespread implementation of whole-field intensity-modulated radiotherapy (IMRT) for the treatment of head-and-neck cancer has resulted in increased levels of dysphagia relative to those seen with conventional planning. Other investigators have suggested an alternative junctioned-IMRT (J-IMRT) method, which matches an IMRT plan to a centrally blocked neck field to restrict the laryngeal dose and reduce dysphagia. The effect on target coverage and sparing of organs at risk, including laryngeal sparing, in the optimization was evaluated and compared with that achieved using a J-IMRT technique.

Methods and Materials

A total of 13 oropharyngeal cancer whole-field IMRT plans were planned with and without including laryngeal sparing in the optimization. A comparison of the target coverage and sparing of organs at risk was made using the resulting dose–volume histograms and dose distribution. The nine plans with disease located superior to the level of the larynx were replanned using a series of J-IMRT techniques to compare the two laryngeal-sparing techniques.

Results

An average mean larynx dose of 29.1 Gy was achieved if disease did not extend to the level of the larynx, with 38.8 Gy for disease extending inferiorly and close to the larynx (reduced from 46.2 and 47.7 Gy, respectively, without laryngeal sparing). Additional laryngeal sparing could be achieved with J-IMRT (mean dose 24.4 Gy), although often at the expense of significantly reduced coverage of the target volume and with no improvement to other areas of the IMRT plan.

Conclusion

The benefits of J-IMRT can be achieved with whole-field IMRT if laryngeal sparing is incorporated into the class solution. Inclusion of laryngeal sparing had no effect on other parameters in the plan.

Introduction

Studies have recently been published that suggest a benefit to using a junctioned intensity-modulated radiotherapy (J-IMRT) technique over a standard whole-field (WF)-IMRT plan for treatment of the head and neck. J-IMRT consists of inverse planning for the superior section of the plan, matched to a conventional anterior neck field, with a central block to reduce the spinal cord dose, usually at approximately the level of the hyoid. The purpose of this technique is to allow sparing of the larynx, which is typically blocked from the anterior field, resulting in very low doses and low toxicity. IMRT planning tends to result in a greater dose to the larynx and corresponding greater levels of dysphagia (1).

In addition to reducing the delineation time for clinicians (because the lower neck can be treated with an open field), one advantage of the J-IMRT technique is that it combines the benefits of IMRT, in terms of greater conformality to the target volumes and increased sparing of the contralateral parotid gland and other normal tissues, with an established method to spare the larynx. It is also claimed that the smaller volume required for IMRT planning with J-IMRT could lead to quicker planning and treatment times, as well as possibly improving the quality of the resulting IMRT plan by allowing the optimization to focus on aspects of the plan other than sparing the larynx (2).

One disadvantage of the J-IMRT technique is that the additional step required in the planning process (conventional field plus IMRT compared with IMRT alone) creates a junction at the match line that could lead to significant over- or underdosage in the event of slight errors in machine setup, as well as a significant reduction in the conformality and coverage in the lower neck. A study by Chao *et al.* (3) found that 5 of 17 locoregional failures occurred in the lower neck region when this technique was used, possibly as a result of poor dose coverage of the deep-lying nodes.

A recent review of this topic by Amdur *et al.* (2) focused on solving the problem of the match line error by feathering the junction (*i.e.,* creating three match lines with the conventional field and

forcing the IMRT plan to optimize to the resulting dose distribution) (4). This was found to reduce the effect of a 1-mm error in the collimator position from 20% to <5% error in dose. The investigators only recommended the J-IMRT technique for plans in which the primary planning target volume (PTV) does not extend beyond the match line. These recommendations were based on an earlier report by the same group that demonstrated a significantly increased dose to the larynx using WF-IMRT compared with when it was covered by a matched anterior neck field (1). This finding was supported by Fua *et al.* (5), who compared WF-IMRT and J-IMRT for 28 patients and found an increase in Grade 3 dysphagia (19 of 20 vs. 5 of 8) for a mean dose of 55.2 and 27.2 Gy, respectively. Dabaja *et al.* (6) found similar results.

All of these investigators have acknowledged that if the larynx is incorporated into the optimization, the larynx dose can be reduced significantly from a quoted mean dose of approximately 50 Gy, typically found when laryngeal sparing is not attempted, to 25–40 Gy. However, Amdur *et al.* (1) and (2) were reluctant to reduce the larynx dose in this way, suggesting that it could compromise the dose distribution elsewhere. A study by Lee *et al.* (7) appeared to contradict this claim by demonstrating, for a limited number of plans, that other dose constraints would not be compromised by reducing the larynx dose to $\sim$30 Gy when the disease is superior to the match line. The investigators proceeded to suggest guidelines for which technique is more suitable for particular disease types (7).

The present study aimed to demonstrate the achievable extent of larynx sparing with WF-IMRT and to provide a comparison with a J-IMRT technique, in terms of the larynx dose, the target coverage, and the effect on the quality of the inverse-planned section of the dose distribution.

Methods and Materials

Extent of larynx sparing achievable with WF-IMRT

A total of 13 clinical WF-IMRT plans incorporating larynx sparing were used in this study, with the larynx defined as superiorly covering the arytenoid cartilages and inferiorly including the

thyroid cartilage. It might be useful to increase the spatial extent of dose sparing in this region by expanding the defined volume superiorly or inferiorly, although the proximity of the disease superiorly and the poorer coverage of the prophylactic lymph nodes resulting from inferior extension will require compromises.

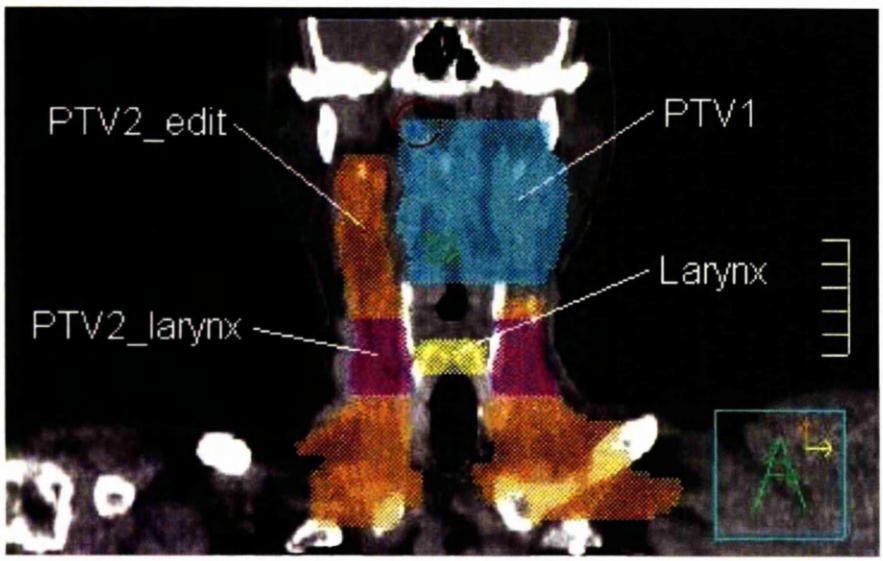


Fig. 1. Volumes used to achieve larynx sparing with intensity-modulated radiotherapy (IMRT). Planning target volume (PTV) 1 does not extend to level of larynx; PTV2 split into PTV2_edit and PTV2_larynx.

Of the 13 patients, 4 had neck disease extending inferiorly past the larynx and 9 did not. The planning method differed from those of previous studies in that the prophylactic PTV volume (PTV2) was split into two parts: PTV2_edit and PTV2_larynx (Fig. 1). The PTV2_larynx was defined as the volume of PTV2 within 1 cm (superiorly or inferiorly) of the larynx, and the PTV2_edit volume was the remainder, situated >1 cm away from the larynx (Fig. 1). The previous clinical constraints used to optimize the PTV2_edit were copied to the PTV2_larynx and modified to relax the minimal dose constraints, effectively allowing a small proportion of the volume adjacent to the larynx to be underdosed. A dose reduction to the larynx could be achieved using either dose–volume or equivalent uniform dose constraints; however, for this

study we used a maximal equivalent uniform dose constraint to the larynx of 2,800 cGy for patients in whom the primary disease did not extend to the larynx (for the 4 patients in whom the primary disease extended to the level of the larynx, the maximal equivalent uniform dose was set to 3,500 cGy). The plans required seven coplanar step-and-shoot IMRT beams of 6 MV. The prescription dose was 65Gy in 30 fractions to PTV1 and 54 Gy to PTV2, with, typically, a total of 50–60 segments. The plans were then created without attempting to spare the larynx. In these cases, the constraints on the larynx and PTV2_larynx were removed and PTV2_edit was redefined to include the PTV2_larynx.

The quality of different aspects of the dose distributions were quantified using dose–volume histogram parameters: maximal and 1-cm^3 doses to the spinal cord (tolerance of 48 and 46 Gy, respectively, in line with the PARSPORT trial protocol) (8) and brainstem (tolerance of 50 and 48 Gy, respectively); mean dose to the contralateral parotid, larynx, and volume of oral cavity outside the PTV; and percentage of volume receiving <90% and <95% of the prescribed dose for PTV1 and PTV2, respectively (tolerance of <1% and <5%, respectively).

Comparison of WF-IMRT and J-IMRT

For comparison with the J-IMRT technique, the treatment plans of the 9 patients in whom the neck disease did not extend to the larynx were each replanned using three variations of the J-IMRT technique at 6 MV. Method 1 used a conformal anterior field with central blocking (2 cm width) along the full length of the field. Method 2 used an anterior field with central blocking extending only to protect the larynx; and Method 3 used an anterior field with central blocking extending only to protect the larynx and a matched posterior field with full central spinal blocking. In each case, the prophylactic dose of 54 Gy was prescribed to a point approximately 2 cm lateral of any blocking and at a depth of 2.5 cm from the anterior skin surface. A comparative analysis was then performed on the dose–volume histograms using the parameters outlined in the previous section. Because previous work on feathering the junction was done to reduce the

sensitivity to machine errors rather than to improve the quality of the plan, it was not thought necessary for the purposes of the present study.

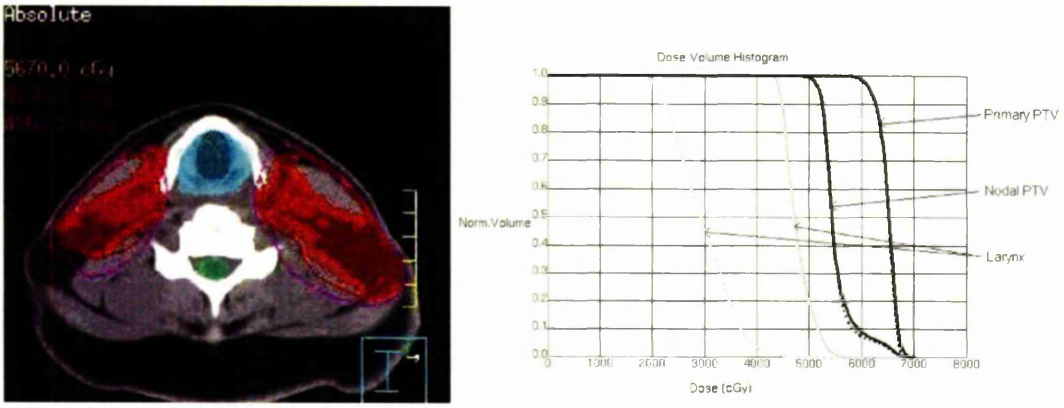


Fig. 2. (a) Coverage of planning target volume 2 at level of larynx using whole-field intensity-modulated radiotherapy and (b) resulting dose–volume histograms (incorporation of laryngeal sparing illustrated by dashed lines).

Results

Extent of larynx sparing achievable with WF-IMRT

When planned with WF-IMRT, good target coverage and organ-at-risk sparing was achievable with a mean larynx dose of 29.1 Gy (range, 26.3–30.4) for the 9 patients with superior disease and 38.8 Gy (range, 36.0–43.1) for the 4 patients with disease extending inferiorly to the level of the larynx. An example of the resulting coverage at the level of the larynx is shown in Fig. 2a, indicating regions of compromise in target coverage at the anterior-medial edges of the target volumes. Repeating the planning process without attempting to spare the larynx resulted in improvements in target coverage focused on the region adjacent to the larynx and a significant increase to the larynx mean dose to 46.2 Gy (p = 0.000) for superior disease and 47.7 Gy (p = 0.004) for inferior disease.

Table 1 and Fig. 2b show that other areas of the plan were not compromised by the addition of laryngeal sparing to the optimization, resulting in no systematic effect on the mean dose to the contralateral parotid, oral cavity, or PTV1 coverage. A disadvantage of this technique is that it can lead to greater doses in the posterior neck at the level of the larynx, because the delivery of radiation from the anterior has been limited. In these cases, a compromise was made with the larynx dose by creating a "Post_neck" volume and applying a maximal dose–volume histogram constraint.

Table 1. Target coverage and OAR doses for WF-IMRT with and without larynx sparing

	Interpatient average dose (Gy)			
	Superior disease (n = 9)		Inferiorly extending disease (n = 4)	
Variable	Spared	Unspared	Spared	Unspared
PTV1 <90% dose	0.4	0.3	0.5	0.5
PTV1 <95% dose	4.6	4.3	6.1	5.3
PTV2 <90% dose	1.3	0.9	0.1	0.1
PTV2 <95% dose	4.5	3	3	3.3
Spinal cord 1 cm^3	44.7	45.1	43	42.7
Maximal spinal cord	47	45.9	44.3	44.4
Brainstem 1 cm^3	44.7	45.8	43.6	44.8
Maximal brainstem	49.2	48.3	48.3	50
Mean parotid	26.1	26.1	31.8	31.8
Mean larynx	29.1	46.2	38.8	47.7
Mean oral cavity	41.6	42	40.3	40.3

Abbreviations: OAR = organ at risk; WF = whole field; IMRT = intensity-modulated radiotherapy; PTV = planning target volume

Comparison of WF-IMRT and J-IMRT

From Table 2, a mean larynx dose of 29.1 and 24.4 Gy was planned for WF-IMRT and J-IMRT using Method 3, respectively (p = 0.008). When comparing other aspects of the plans, no significant effect was found on the dose to the contralateral parotid, oral cavity, or PTV1 coverage. However, with Methods 1 and 2, in which no matched posterior field was used, a marked reduction occurred in target coverage using J-IMRT in the inferior part of the plan, a moderate example of which is given in Fig. 3a–d, where the medial and posterior edges of PTV2 were underdosed relative to that achieved with IMRT. This loss of coverage could become more

significant depending on the specific patient contours and the lateral and posterior extent of the lymph nodes (Fig. 3e,f). The second J-IMRT method, in which no spinal blocking was present inferior to the larynx, also resulted in a significant increase in the spinal cord 1-cm^3 volume such that the plan exceeded tolerance (p = 0.024).

Table 2. Comparison of target coverage and OAR doses for WF-IMRT and J-IMRT techniques

	Interpatient average dose (n = 9)			
Variable	WF-IMRT	J-IMRT 1	J-IMRT 2	J-IMRT 3
PTV1 <90% dose	0.4	0.5	0.5	0.6
PTV1 <95% dose	4.6	6.2	6.2	6
PTV2 <90% dose	1.3	3.3	2.5	1.4
PTV2 <95% dose	4.5	9.8	10.1	5.5
Spinal cord 1 cm^3	44.7	44.2	47.5	44
Maximal spinal cord	47	46.2	48.3	46.8
Brainstem 1 cm^3	44.7	45	44.8	43.8
Maximal brainstem	49.2	48.4	48.1	48.3
Mean parotid	26.1	24.8	25	25.1
Mean larynx	29.1	24	24.2	24.4
Mean oral cavity	41.6	41.6	41.6	41.6

Abbreviations: J = junctioned; J-IMRT 1,2,3 = J-IMRT methods 1,2, and 3, respectively; other abbreviations as in Table 1

The data in Table 2 demonstrate that the addition of the posterior beam in Method 3 led to target coverage comparable to that achieved with WF-IMRT (p = 0.339). However, the individual plans demonstrated that although WF-IMRT gave good coverage in all situations, in 2 of 9 cases, Method 3 of J-IMRT led to severe underdosage of large regions of PTV2, even after the relative weightings of the neck fields were optimized.

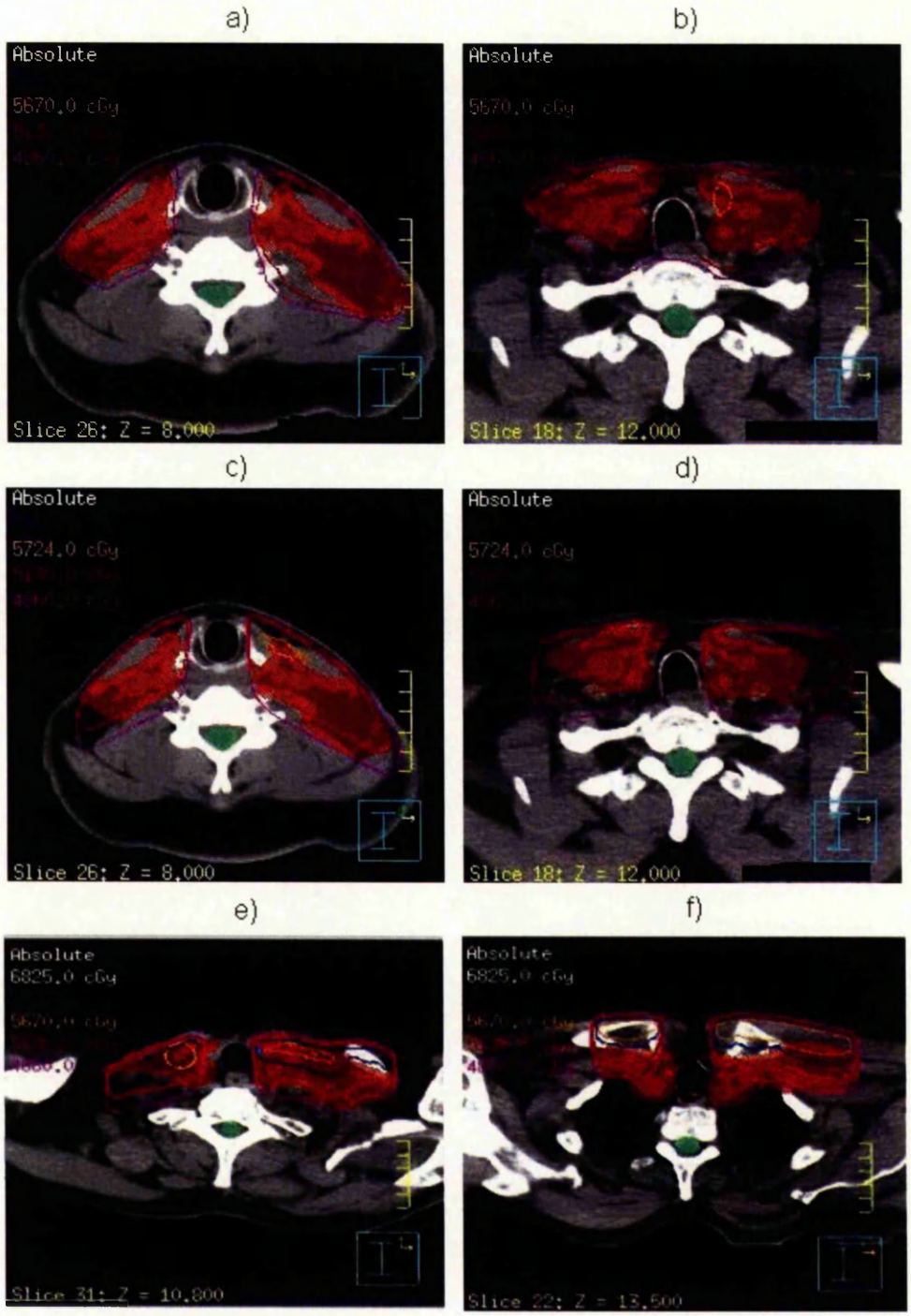


Fig.3. Comparison of target coverage with (a,b) whole-field intensity-modulated radiotherapy (IMRT) and (c,d) junctioned IMRT. Better coverage was achieved with whole-field IMRT at medial and posterior edges of nodes. (e,f) Extreme example of limitations of junctioned IMRT also shown.

Discussion

The mean laryngeal doses can be significantly reduced with WF-IMRT, including in cases with inferiorly extending disease, without significantly compromising target coverage or the achievement of dosimetric goals elsewhere in the plan. The Pinnacle treatment planning system (ADAC, Milpitas, CA) used in this study incorporates direct machine parameter optimization software, which improves the efficiency of the optimization process and indicates what can be achieved with currently available commercial software. Other planning systems might not achieve as good a compromise between larynx sparing and acceptable target coverage, although any future trend toward improved optimization or delivery efficiency would act to reduce the discrepancy in the mean larynx dose we have demonstrated between WF-IMRT and J-IMRT.

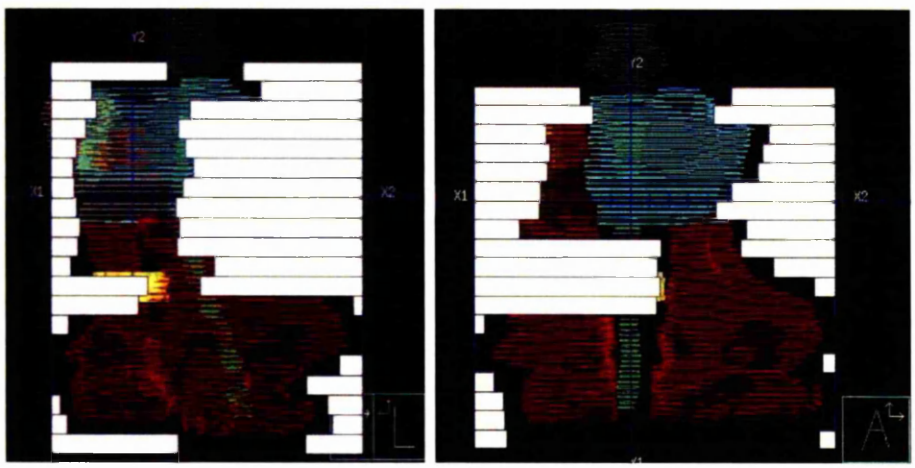


Fig.4. Segments illustrating increase in segment irregularity with incorporation of larynx sparing into intensity-modulated radiotherapy optimization

Care must be taken with the segments used to deliver the plan because the shielding of the larynx is often achieved by extending two or three leaves across large fields (Fig. 4), thus creating very irregular shapes that will be less accurately calculated by the Pinnacle treatment planning system. In this situation, the more extreme segments (*i.e.*, those with extensive blocking or elongation) have either been removed or manually adjusted to expose the larynx.

This solution is effective because the alteration of a limited proportion of segments will only cause a moderate increase in the larynx dose that is unlikely to be of clinical significance. The WF-IMRT plans used in the present study were all clinically verified using an absolute dose tolerance of 3% in PTV1 and an individual fluence verification tolerance of <5% pixel failure within the 20% isodose at 4%/4 mm.

Although the J-IMRT technique advocated by other investigators allows for greater laryngeal sparing, we do not believe it is the best option, given the results of the present study. Several investigators have recommended a mean laryngeal dose threshold of 50 Gy on the basis of the observed functional outcome (9) and (10). Early data from a study by Feng *et al.* (11) appear to reinforce this by suggesting that a mean dose threshold of 50 Gy to the larynx correlates with dysphagia, although they found no evidence of a similar effect at a dose <40 Gy. Sanguineti *et al.* (12) recommended a mean laryngeal dose threshold of 43.5 Gy and a volume receiving 50% of prescribed dose of <27% to minimize laryngeal edema. All these constraints were comfortably achieved in this study for all patients. At present, no evidence has shown that the mean larynx dose values observed for WF-IMRT and J-IMRT (29.1 vs. 24.4 Gy) are likely to result in a significant increase in toxicity. The randomized controlled trial necessary to establish any clinical difference was beyond the scope of this work.

Whole-field IMRT also allows more conformal coverage of the prophylactic nodal volumes in the lower neck compared with the underdosage on the medial and posterior edges that are common with J-IMRT. J-IMRT is unsuitable when neck disease extends inferiorly past the larynx; thus, a WF-IMRT technique is recommended for these cases (7). Our study has demonstrated that WF-IMRT can achieve considerable larynx sparing. It is our experience that the extra time required to add larynx sparing to the optimization is more than offset by the gains, particularly compared with the necessary development of feathered junctions for J-IMRT.

Conclusion

The results of our study have shown that it is possible to reduce the larynx dose to clinically acceptable levels using a WF-IMRT technique that incorporates larynx sparing into the optimization without significantly compromising target coverage or other normal tissue constraints. Although the J-IMRT technique advocated in other studies allows for greater laryngeal sparing, the difference is unlikely to be clinically significant.

References

1 R. Amdur, J. Li and C. Liu *et al.*, Unnecessary laryngeal irradiation in the IMRT era, *Head Neck* **26** (2004), pp. 257–264

2 R. Amdur, C. Liu and J. Li *et al.*, Matching intensity-modulated radiation therapy to an anterior low neck field, *Int J Radiat Oncol Biol Phys* **69** (2007), pp. S46–S48

3 K. Chao, G. Ozyigit and B. Tran *et al.*, Patterns of failure in patients receiving definitive and postoperative IMRT for head-and-neck cancer, *Int J Radiat Oncol Biol Phys* **55** (2003), pp. 312–321

4 J. Li, C. Liu and S. Kim *et al.*, Matching IMRT fields with static photon field in the treatment of head and neck cancer, *Med Dosim* **30** (2005), pp. 135–138

5 T. Fua, J. Corry and A. Milner *et al.*, Intensity modulated radiotherapy for nasopharyngeal carcinoma: Clinical correlation of dose to the pharyngo-laryngeal axis and dysphagia, *Int J Radiat Oncol Biol Phys* **67** (2007), pp. 976–981

6 B. Dabaja, M. Salehpour and I. Rosen *et al.*, Intensity modulated radiation therapy (IMRT) of cancers of the head and neck: Comparison of split-field and whole-field techniques, *Int J Radiat Oncol Biol Phys* **63** (2005), pp. 1000–1005

7 N. Lee, J. Mechalakos and D. Puri *et al.*, Choosing an intensity modulated radiation therapy technique in the treatment of head-and-neck cancer, *Int J Radiat Oncol Biol Phys* **68** (2007), pp. 1299–1309

8 M.T. Guerrero, C.H. Urbanoa and C. Clarka *et al.*, Target Volume Definition for Head and Neck Intensity Modulated Radiotherapy: Pre-clinical Evaluation of PARSPORT Trial Guidelines PARSPORT trial management group, *Clin Oncol* **19** (2007), pp. 604–613.

9 K. Dornfeld, J. Simmons and L. Karnell *et al.*, Radiation doses to structures within and adjacent to the larynx are correlated with long-term diet- and speech-related quality of life, *Int J Radiat Oncol Biol Phys* **68** (2007), pp. 750–757

10 A. Eisbruch, P. Levendag and F. Feng *et al.*, Can IMRT or brachytherapy reduce dysphagia associated with chemoradiotherapy of head and neck cancer? The Michigan and Rotterdam experiences, *Int J Radiat Oncol Biol Phys* **69** (2007), pp. S40–S42.

11 F. Feng, H. Kim and T. Lyden *et al.*, Intensity-modulated radiotherapy of head and neck cancer aiming to reduce dysphagia: Early dose-effect relationships for the swallowing structures, *Int J Radiat Oncol Biol Phys* **68** (2007), pp. 1289–1298

12 G. Sanguineti, P. Adapala and E. Endres *et al.*, Dosimetric predictors of laryngeal edema, *Int J Radiat Oncol Biol Phys* **68** (2007), pp. 741–749

This page is intentionally blank

Publication #5

Design and implementation of a head-and-neck phantom for system audit and verification of intensity-modulated radiation therapy

Published in Journal of Applied Clinical Medical Physics, 9(2), 2008, p46-56

Authors:

Webster GJ

Hardy MJ

Rowbottom CG

Mackay RI

14 pages

This page is intentionally blank

JOURNAL OF APPLIED CLINICAL MEDICAL PHYSICS, VOLUME 9, NUMBER 2, SPRING 2008

Design and implementation of a head-and-neck phantom for system audit and verification of intensity-modulated radiation therapy

Gareth J. Webster,[a] Mark J. Hardy, Carl G. Rowbottom, and Ranald I. Mackay

North Western Medical Physics, Christie Hospital NHS Foundation Trust, Manchester, U.K.
Gareth.Webster@physics.cr.man.ac.uk

Received 6 August 2007; accepted 13 December 2007

The head and neck is a challenging anatomic site for intensity-modulated radiation therapy (IMRT), requiring thorough testing of planning and treatment delivery systems. Ideally, the phantoms used should be anatomically realistic, have radiologic properties identical to those of the tissues concerned, and allow for the use of a variety of devices to verify dose and dose distribution in any target or normal-tissue structure.

A phantom that approaches the foregoing characteristics has been designed and built; its specific purpose is verification for IMRT treatments in the head-and-neck region. This semi-anatomic phantom, HANK, is constructed of Perspex (Imperial Chemical Industries, London, U.K.) and provides for the insertion of heterogeneities simulating air cavities in a range of fixed positions. Chamber inserts are manufactured to incorporate either a standard thimble ionization chamber (0.125 cm^3: PTW, Freiburg, Germany) or a smaller PinPoint chamber (0.015 cm^3: PTW), and measurements can be made with either chamber in a range of positions throughout the phantom. Coronal films can also be acquired within the phantom, and additional solid blocks of Perspex allow for transverse films to be acquired within the head region.

Initial studies using simple conventional head-and-neck plans established the reproducibility of the phantom and the measurement devices to within the setup uncertainty of ±0.5 mm. Subsequent verification of 9 clinical head-and-neck IMRT plans demonstrated the efficacy of the phantom in making a range of patient-specific dose measurements in regions of dosimetric and clinical interest. Agreement between measured values and those predicted by the Pinnacle^3 treatment planning system (Philips Medical Systems, Andover, MA) was found to be generally good, with a mean error on the calculated dose to each point of +0.2% (range: –4.3% to +2.2%; $n = 9$) for the primary planning target volume (PTV), –0.1% (range: –1.5% to +2.0%; $n = 8$) for the nodal PTV, and +0.0% (range: –1.8% to +4.3%, $n = 9$) for the spinal cord. The suitability of the phantom for measuring combined dose distributions using radiographic film was also evaluated.

The phantom has proved to be a valuable tool in the development and implementation of clinical head-and-neck IMRT, allowing for accurate verification of absolute dose and dose distributions in regions of clinical and dosimetric interest.

Key words: IMRT, verification, audit

PACS numbers: 87.53.-j, 87.53.Xd, 87.56.Fc

[a] Corresponding author: Gareth Webster, North Western Medical Physics, Christie Hospital NHS Trust, Manchester, U.K.; phone: +44-161-446 3536; fax: +44-161-446-3545; email: Gareth.Webster@physics.cr.man.ac.uk

46

I. INTRODUCTION

The head and neck is a challenging site for treatment with intensity-modulated radiotherapy (IMRT), and thorough testing of planning techniques and treatment delivery systems is required for safe introduction of that technique into clinical practice.[1,2] The improved conformity of the dose delivered with IMRT to target structures and the sharp gradients between target and critical tissues can allow for dose escalation to the treatment volume while acceptable levels of radiation damage are maintained. The result can be increased proximity of high-dose regions to critical structures.

The potential for a large dosimetric error between planned and delivered treatments means that several stages in the treatment process must be performed to a higher level of accuracy than is typically conventionally achieved. For example, the beam modeling by the treatment planning system (TPS) is likely to require improved accuracy as compared with the accuracy acceptable for conventional treatments, in which penumbral and out-of-field doses are less significant to the evaluation of the dose distribution. This issue arises because the contribution of out-of-field doses to clinically significant areas is much higher for IMRT than for conventional radiotherapy. For critical structures such as the spinal cord and brainstem, errors in such contributions to the calculated dose could misleadingly overestimate or, critically, underestimate the delivered dose to a sensitive tissue. Direct measurement in such regions can confirm the accuracy of the planned dose.

To provide an accurate representation of the patient contour, verification of IMRT techniques at both the commissioning and clinical stages should use phantoms that can provide a realistic representation of the clinical site being treated. Ideally, these phantoms should be anatomically realistic, have radiologic properties that are identical to those of the tissues concerned, and allow for a variety of measuring devices to be used to verify dose and dose distribution in a number of key positions throughout the target and normal-tissue volumes. In addition, given the many links in the radiotherapy chain, it is important to minimize systematic errors in the process from imaging through treatment planning to the fractionated delivery of the radiation.[1] A phantom that can test the various components of the system before clinical implementation of IMRT is vital to ensure confidence in the new technique.

To meet those requirements, a semi-anatomic head-and-neck phantom was designed and manufactured at the Christie Hospital. The phantom can be used to verify clinical IMRT treatments in the head-and-neck region and to ensure that systematic errors in the process are minimized. The present paper details the considerations involved in the design of the phantom and the current successes and limitations resulting from its use.

II. MATERIALS AND METHODS

Table 1 summarizes the design considerations for the head-and-neck phantom and the features incorporated into the design to address those considerations. The approximately anatomically realistic shape of the phantom comprises a head section that links to a shoulder section (Fig. 1); together, the two parts approximate the contours of a typical patient. The choice of Perspex (Imperial Chemical Industries, London, U.K.) as the phantom material was made because of considerations of ease of availability and the similarity of that material's density to the average of tissue and bone in the human head. The validity of the density assumption was checked by comparing the average density of the phantom, including oral cavity and esophageal heterogeneities, to the average density of 7 head-and-neck IMRT patients from the lung apex to approximately the superior border of the spinal cord. The mean patient density was found to be 1.073 g/cm^3 (range: 1.018 – 1.236 g/cm^3) as compared with 1.076 ± 0.003 g/cm^3 for the phantom. The alternative option of manufacturing the phantom using a water-equivalent material

that allowed for the insertion of a comprehensive range of bone- and air-equivalent heterogeneities, while being more anatomically realistic, was felt to compromise the simplicity and flexibility of the phantom. Given those limitations, the use of a material of mean tissue density was felt to provide a more practical patient representation. The solution developed here allows for an assessment of dose calculation accuracy at depths and obliquities similar to those for a patient.

Fig. 2 shows the structure of the head section of the phantom: a range of movable slabs and removable blocks that run craniocaudally through the head section. One of the blocks includes a cylindrical insert that can be removed to allow insertion of an ionization chamber for absolute dose measurement. Chamber inserts were designed with scalloped ends to tightly fit a thimble ionization chamber (0.125 cm^3: PTW, Freiburg, Germany) and a PinPoint chamber (0.015 cm^3: PTW), as illustrated in Fig. 3(a). Flat-ended spacers are used as necessary to alter the superior–inferior position of the chamber for measurements. The shoulder section is fitted with a similar system of versatile measurement points, so that a three-dimensional (3D) grid of 156 measurement points is available throughout the head and neck region. The relevant point can be selected at positions of clinical or dosimetric interest.

The ability to incorporate simple heterogeneities into the phantom not only permits a more realistic representation of a patient, but also allows for verification of heterogeneity correction in the TPS. Some heterogeneities can be simulated simply by removing existing inserts, for example in the lung (Fig. 2). However, it is often more anatomically realistic for an air cavity

TABLE 1. Design considerations for a verification phantom for intensity-modulated radiation therapy of head and neck

Design requirement	*Solution*
Approximate patient shape	Separate sections were made for the head and the shoulders (Fig. 1) to generate a realistic surface contour.
Flexible design	Movement of various components is accommodated throughout the phantom's volume.
Stable geometric phantom with repeatable setup	The required flexibility of measurement positions limits the complexity of the phantom, and so a simple geometric design is used.
Radiologically realistic material	The phantom was manufactured from Perspex (Imperial Chemical Industries, London, U.K.), a material with a density comparable to that found to be the average of tissue and bone in the head-and-neck region.
An array of repeatable dose measurement points throughout the phantom volume	Both sections of the phantom have a slab structure (Fig. 2; head: two slabs; shoulder: one slab) containing movable blocks. The blocks, one of which can contain an ionization chamber, can be moved laterally, and the slabs can be interchanged. Spacers provide the superior–inferior flexibility needed to accurately position the ionization chamber.
Ability to use standard thimble ionization chambers (0.125 cm^3: PTW, Freiburg, Germany) and PinPoint chambers (0.015 cm^3: PTW)	Various chamber inserts were manufactured to securely hold both chamber types (Fig. 3).
Facility to include clinically relevant heterogeneities	Removable lung sections are included inferiorly. Balsawood inserts of density 0.07 g/cm^3 [Fig. 3(b)] simulate the oral cavity and trachea. For flexibility, other materials and different positions can be used.
Facility to acquire coronal films of combined dose distributions	Films can be positioned between certain slabs extending through both the head and the shoulder section of the phantom.
Facility to acquire transverse films of combined dose distributions	A separate phantom consisting of two Perspex slabs with surface contours identical to those of the head section of the original phantom was made (Fig. 4).

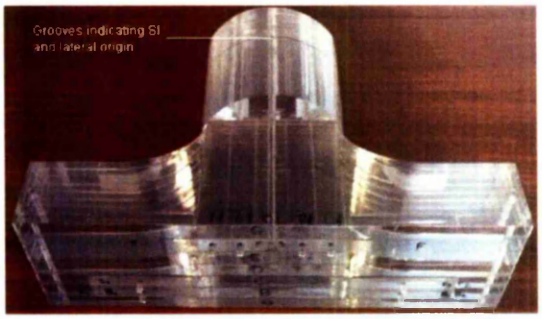


FIG. 1. Semi-anatomic Perspex (Imperial Chemical Industries, London, U.K.) phantom for verification of head-and-neck treatment delivery (HANK). SI = superior–inferior.

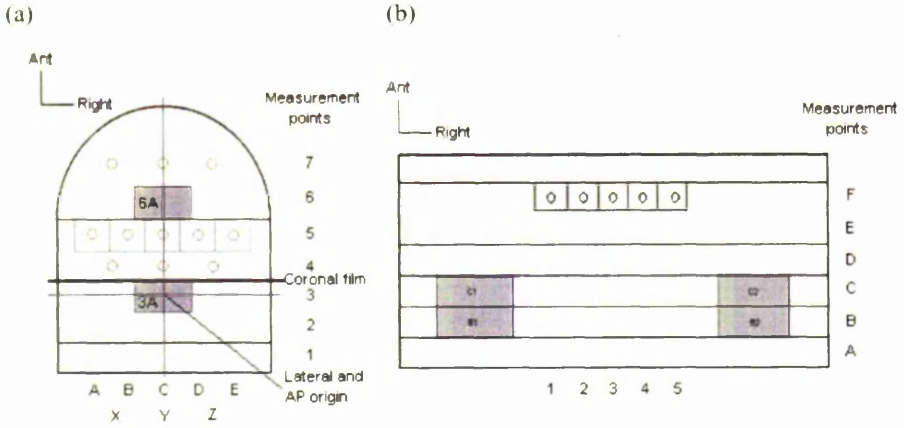


FIG. 2. (a) Head section: The blocks containing slices 2 and 3, and 4 and 5 can be interchanged. (b) Shoulder section: Measurement points can be placed in slices C to F. The tracheal heterogeneity can be positioned at the center of the measurement slice. Blocks B and C are removable blocks simulating lung heterogeneities. Ant = anterior; AP = anterior–posterior.

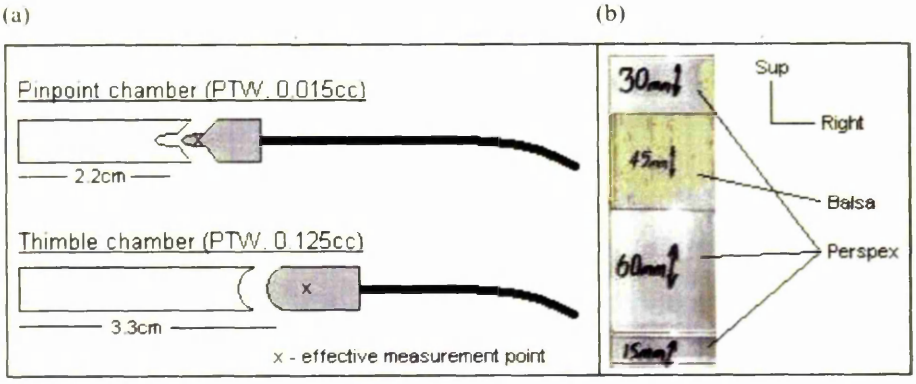


FIG. 3 (a) Schematic of chamber inserts for the thimble (0.125 cm^3: PTW, Freiburg, Germany) and PinPoint (0.015 cm^3: PTW) ionization chambers. (b) Repeatable setup of the balsawood heterogeneity representing the oral cavity, viewed in the coronal plane of heterogeneity [see Figs. 2 and 6(a)]. Sup = superior.

to be of a more limited length—for example, the oral cavity, which requires the air cavity to be of a known and repeatable length and position. Inserts have therefore been made to simulate typical heterogeneities. A balsa wood insert can replace a Perspex insert in the shoulder section of the phantom to simulate the trachea. Blocks of Perspex and balsa have also been made to simulate the oral cavity in the configuration shown in Fig. 3(b), which is positioned in the coronal plane in phantom section 6A [see Fig. 2(a)]. Balsa wood is a good material for this purpose because it has a low electron density, close to that of air (0.07 g/cm^3). The balsa wood inserts allow for repeatable positioning of the Perspex blocks for precise measurements near to the inhomogeneity.

The structure of the phantom permits radiographic film to be inserted between slabs 3 and 4 in the head, extending into the shoulder section if necessary, to measure the dose distribution in the coronal plane [see Fig. 2(a)]. No such capability exists for measuring transverse slices, because the phantom would then have to be split into several parts, rather than the current two, which would make the phantom impractical to use and would reduce its stability. A second phantom was therefore designed (Fig. 4). This simplified replica of the head section contains no heterogeneities or slab structure (the latter being unnecessary when measuring transverse slices). To measure the dose distribution of any transverse plane, a film can be placed between the slabs before the phantom is positioned and the treatment delivered.

Head-and-neck phantoms for radiotherapy have been developed elsewhere. Examples include the Rando phantom (originally produced by Alderson Research Laboratories, Stamford, CT) and the Radiological Physics Center (RPC) phantom,[3] both of which provide an anatomically realistic external contour of a typical patient. The Rando phantom is constructed from a natural human skeleton cast inside material that is radiologically equivalent to soft tissue. The slab structure allows for the use of transverse films and a 3D array of thermoluminescent dosimeters (TLDs) for dosimetric verification. The RPC phantom consists of an anthropomorphic Perspex cast, inside which a cubic dosimetry insert able to hold TLDs and film can be positioned; the surrounding volume is then filled with water. Neither of these phantoms allows for the insertion of ionization chambers. The HANK phantom described in this work provides increased flexibility over other phantoms in terms of the available number and position of measurement points with either ionization chamber or film. The design of the phantom also means that it can be easily modified to incorporate inserts for other dosimeters such as microMOSFETs (metal oxide semiconductor field effect transistors) and fiberoptic detectors.

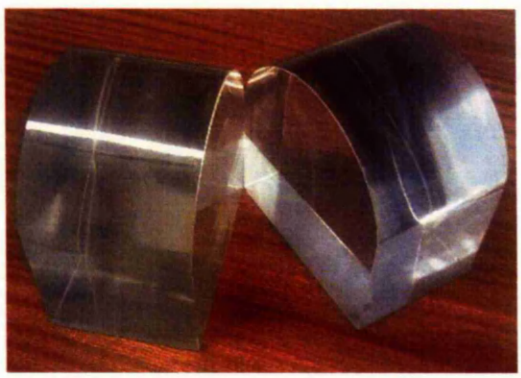


FIG. 4. Head phantom for transverse film measurements.

A. Conventional/3D-conformal plan tests

We used the HANK phantom to carry out an audit on the general method used for treatment planning at our institution. A PinPoint chamber (0.015 cm^3) was used for all measurements. Low dose gradient measurements were repeated with a thimble ionization chamber (0.125 cm^3). The ionization chambers were calibrated against a Farmer-type chamber (30002: PTW), which in turn was traceable to the U.K. national standard. Plans were calculated on the Pinnacle[3] TPS (version 7.4f), which incorporates a collapsed cone convolution dose algorithm, and were delivered on an Elekta Precise (Elekta, Crawley, UK) linear accelerator (LINAC).

To assess the reproducibility of measurement position achievable with the phantom, we carried out 3 investigations to establish a positional uncertainty:

- A PinPoint chamber was placed within the phantom at the edge of a field. Baseline reproducibility of delivery and measurement was established by 6 identical repeats without moving any component of the phantom or LINAC.
- To determine the repeatability of chamber positioning, the same field was delivered 6 times, each time removing and replacing the chamber.
- To ascertain positional uncertainty in the phantom setup, a field was delivered 6 times, resetting the phantom in 1 cardinal couch direction to the original position. This test was carried out in all 3 orthogonal couch directions, and the field was chosen so that the chamber was in jaw penumbra in the relevant direction.

The standard deviation (SD) in the dose measurement was related to a positional uncertainty by determining the dose gradient for each point from the TPS. The uncertainty was calculated as ±2 SD so as to give a 95% confidence level. The coefficient of variation (CoV) was also calculated (CoV = SD / mean).

To verify the validity of using the phantom for all types of head-and-neck treatment, 9 other standard conventional treatments, listed in Fig. 5, were delivered to the phantom, including heterogeneous plans and conformal treatments. The treatments were delivered and assessed in order of increasing complexity so that the cause of any significant error could be isolated. The

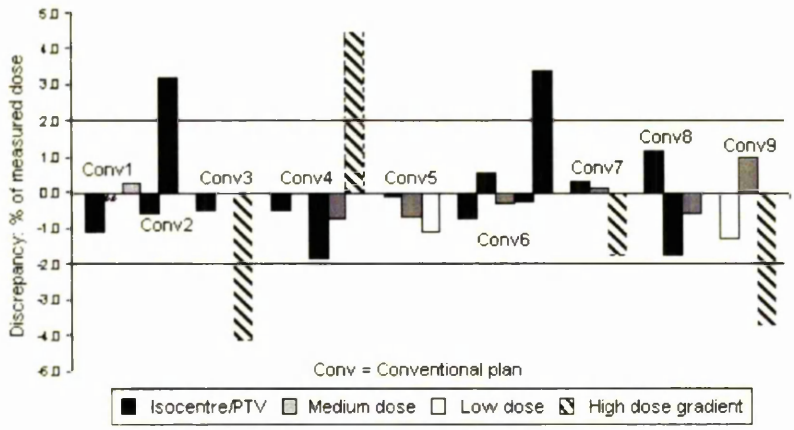


FIG. 5. Results for simple treatments delivered to phantom, required to 2% accuracy. For all plans, isocenter dose was 200 cGy. All high dose gradient measurement doses were within 2 mm of the point dose within the planning system. Doses: isocenter/PTV (planning target volume) = 180 – 220 cGy; medium = 110 – 160 cGy; low = 5 – 40 cGy. Conventional plans: Conv1 = homogeneous superiorly-oblique wedged pair; Conv2 = homogeneous orthogonal wedged pair; Conv3 = parallel opposed pair (isocenter only); Conv4 = single anterior field; Conv5 = three-field brick (pituitary); Conv6 = heterogeneous superiorly-oblique wedged pair; Conv7 = heterogeneous orthogonal wedged pair; Conv8 = heterogeneous conformal nasal cavity; Conv9 = heterogeneous wedged pair (orbit).

range of measurements performed was designed to establish whether the phantom would be capable of successfully auditing an IMRT program for head-and-neck patients. The measurement points depended on the particular plan and were chosen to be in high- and mid-level dose regions and regions of sharp dose gradient.

The acceptability criteria of 2% or 2 mm set out by the International Commission on Radiation Units and Measurements were used to assess the suitability of the phantom for all non-IMRT treatments.[4] The 2-mm acceptance dose range was determined by using points of interest in the planned dose distribution to calculate the maximum dose range 2 mm in each cardinal direction.

B. IMRT plan tests

Head-and-neck IMRT was recently clinically implemented at the Christie Hospital. The procedure to verify the absolute dose is outlined here to highlight the utility of the HANK phantom. All cases were planned in line with the PARSPORT Trial protocol, using 5 beams at 6 MV (62 – 80 segments in total) to deliver 65 Gy to the primary PTV and 54 Gy to the nodal PTV in 30 fractions.[5]

A computed tomography scan of the phantom (containing heterogeneities to represent the oral cavity and the trachea) was imported into the TPS. The clinical patient plan was transferred to the phantom, and the dose distribution was calculated. A sample of absolute dose values was taken from the 156 available measurement points (coordinates noted in the departmental protocol), chosen from regions of clinical or dosimetric interest within the dose distribution—that is, in the primary PTV, the nodal PTV, and the spinal cord [highlighted in Fig. 6(a,b)]. The points chosen were in stable areas of the combined dose distribution, avoiding steep dose gradients, so that the verification results were less likely to be erroneous as a result of small setup errors. Such a selection may not always be possible, particularly with a dose point representative of the spinal cord region, and so a range of potentially suitable points should be evaluated to choose one that is least affected by penumbrae.

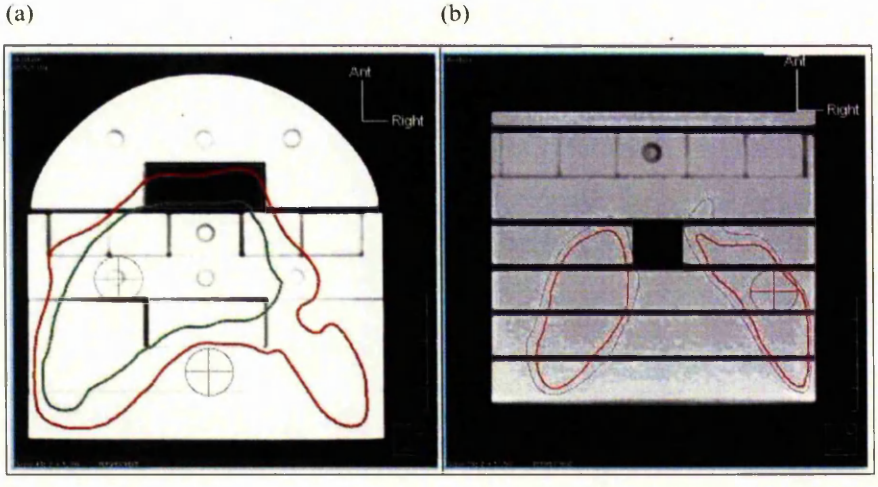


FIG. 6. Computed tomography scan of phantom in the treatment planning system, illustrating (a) head section with simulated oral cavity and typical dose distribution with clinically relevant primary planning target volume (PTV, red) and spinal cord (blue) points of interest, and (b) shoulder section with nodal PTV measurement point (red) and trachea insert. Ant = anterior.

The plan was delivered to the phantom, and a calibrated PinPoint chamber was used to take measurements at the chosen points. Overall dose action levels were chosen as 3% for the primary PTV and spinal cord, and 4% for the nodal PTV. The action level for individual beam errors was set at 5%.

Verification of the combined dose distribution was determined from the combination of all treatment beams involved in the delivery of a full fraction of a treatment to radiographic film positioned within the phantom. Transverse slices were calculated in the TPS as described earlier, using the phantom shown in Fig. 4, in which bisected films (resealed in a dark room) were then acquired 15 mm superior and 15 mm inferior to the isocenter for 3 patients. Coronal films were acquired within the main phantom for 3 patients.

III. RESULTS

A. Conventional/3D-conformal plan tests

Table 2 shows the results of the repeatability of measurement and reproducibility of phantom setup; the findings demonstrate that the phantom can be set up and the dose to a point measured with a spatial precision of approximately ±0.5 mm. The chamber can be positioned to ±0.03 mm precision. Those results were calculated from CoVs of 0.005 and approximately 0.05 for chamber positioning and phantom setup respectively, values that compare to a CoV of 0.001 for repeated measurements without moving the phantom, chamber, or LINAC components. These findings demonstrate that the phantom setup has the precision required to measure dose in sharp field gradients, an attribute essential for IMRT commissioning.

Fig. 5 shows the results for absolute dose verification for all measured points; the findings indicate good agreement for most points (measurements to within ±2% of TPS prediction). The errors exceeding tolerance were either in high dose gradients or measured at beam penumbrae that entered the phantom through a sharp change in shape. The TPS is known not to model some sharp changes in shape well, because there is no requirement in clinical practice to model them. However, for a single-field treatment, the TPS did model a beam with the central axis passing through a sharp change in phantom shape to within 2% and 2 mm of the measured dose. It can therefore be concluded that the problem lies in modeling sharp changes in phantom shape within the beam penumbra. Measurement points were subsequently chosen to avoid measurements within the beam penumbrae that passed through such changes in shape. This problem was not considered significant with regard to the clinical use of the TPS, because such changes in shape are not encountered clinically.

TABLE 2. Reproducibility results[a]

Uncertainty	*Coefficient of variation (±)*	*Positional uncertainty (±mm)*
Chamber repositioning	0.004	0.03
Longitudinal setup	0.051	0.30
Vertical setup	0.031	0.27
Lateral setup	0.029	0.27
Root mean square setup uncertainty		0.49

[a] Positional uncertainty was determined from the treatment planning system, assuming a linear gradient from the measurement point to 2 mm from the point. Approximately 65 cGy was delivered from 200 MU in the center of the phantom.

B. IMRT plan tests

Fig. 7 shows the absolute dose verification results for 9 patient plans. The results for individual dose points have a mean error on the calculated dose to each point of +0.2% (range: –4.3% to +2.2%, $n = 9$) for the primary PTV, –0.1% (range: –1.5% to +2.0%, $n = 8$) for the nodal PTV, and 0.0% (range: –1.8% to +4.3%, $n = 9$) for the spinal cord, indicating no systematic errors in measurement.

The measured differences for individual beams were found to exceed the 5% action level for, on average, 3 of the 15 measurements for each plan (3 measurements for each of 5 beams). In particular, discrepancies were found for measurements of the spinal cord point, an expected finding because of the high proportion of beam penumbrae that make up the dose in this region, making that point particularly sensitive to setup error or subtle errors in beam modeling. Tolerance was also exceeded for individual beams for measurements within the primary PTV.

We investigated these discrepancies further by determining the error on individual segments and found them to correlate closely with heavily-weighted segments in which the measurement point was found to be in or near the field penumbra and therefore susceptible to setup or delivery error, and also to correlate with subtle errors in the modeling of the beam penumbra by the TPS. Fig. 8 shows examples of these segments for the IMRT 9 plan.

These results suggest that the phantom and the beam model are both suitable for IMRT planning and verification but that, because of the inability of point measurement techniques to determine distance-to-agreement, plan verification must include methods to measure the dose distribution. Coronal films were acquired for 3 clinical plans between slabs 3 and 4 of the head section (extending to between slabs C and D of the shoulder section), corresponding approximately to a plane 1 – 2 cm anterior to the spinal cord, in which the dose distribution is considerably modulated in the attempt to minimize dose to the spinal cord. Gamma analysis showed good agreement with calculation, with a mean pixel failure of 1.6% (range: 0.5% – 3.6%) within the 50% isodose at 4% and 4 mm.

Fig. 9 shows a typical transverse gamma map for a combined treatment and demonstrates reasonably close agreement between calculation and measurement, although not as close as that seen with the coronal films. Results showed a mean pixel failure of 5.1% (range: 2.3% – 9.4%)

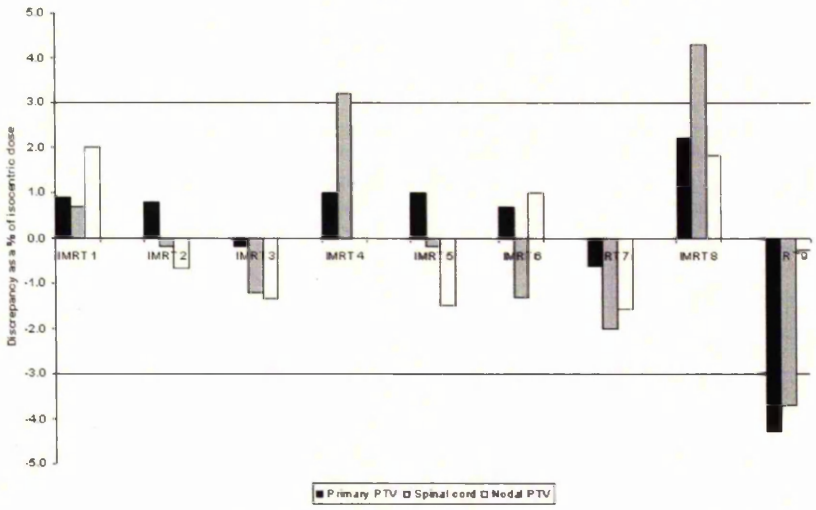


FIG. 7. Results for absolute dose verification of clinically and dosimetrically relevant points for a sample of head-and-neck intensity-modulated radiation therapy plans. PTV = planning target volume.

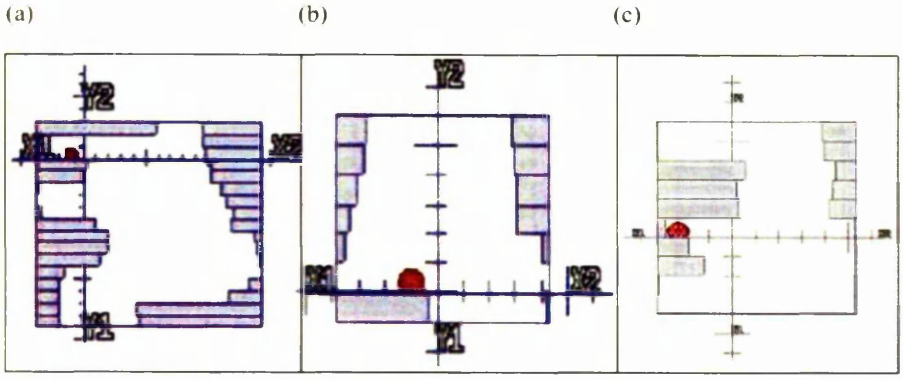


FIG. 8. Illustration of segments causing measurement error in (a,b) primary planning target volume (PTV) point from the left anterior oblique beam, and (c) spinal cord point from the right anterior oblique beam. The red spheres represent the measurement points.

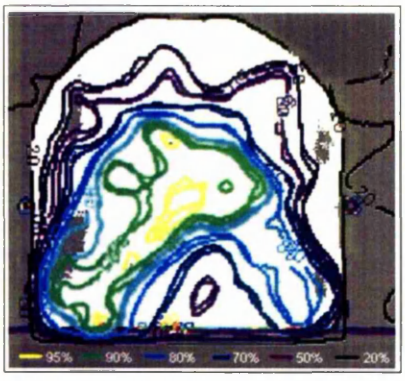


FIG. 9. Typical gamma evaluation map at 5% and 5 mm for an axial slice at approximately the level of the parotids.

within the 50% isodose at 4% and 4 mm. The image does not appear to suffer from any artifacts as a result of cutting the film, but it does show areas of poor agreement between calculation and measurement toward the phantom surface, perhaps because of the known breakdown of the TPS dose calculation accuracy in that region.[6]

IV. DISCUSSION

Our investigations into setup reproducibility and verification accuracy have provided confidence that the phantom developed in-house is suitable for audit of conventional treatments with and without the use of multileaf collimators, and with homogeneous and heterogeneous plans. They have also served to audit the entire patient pathway for simple head-and-neck treatments from computed tomography scan through TPS to delivery. Our findings provided the necessary foundation to confidently use the phantom for IMRT audit purposes.

The HANK phantom has enabled systematic audit of the radiotherapy chain and development of IMRT verification procedures. It has highlighted a penumbra measurement problem inherent to the verification of IMRT plans for complex sites such as the head and neck, and as

a result, absolute dose measurement points are, where possible, chosen in regions of low dose gradient so as to reduce the potential impact of setup errors in the measurement. Problems can arise from choosing measurement points in apparently stable areas within the full treatment plan dose distribution. Care must be taken when selecting measurement points to avoid positioning them in or near the penumbrae for several segments.

The additional measurement of the relative dose distribution with film is essential because of the lack of spatial information provided by point measurements of absolute dose using this phantom. The additional phantom shown in Fig. 4 can be used to acquire transverse slices through the head, although the accuracy of the resulting films appears limited as compared with the coronal films.

V. CONCLUSION

A semi-anatomic Perspex phantom was developed and is now in routine use for the verification of IMRT for the head and neck. The phantom incorporates heterogeneities to represent the oral cavity and trachea and can be used with ion chambers or film to measure dose points or dose distributions. It has a non-uniform array of 156 possible measurement positions, from among which points of clinical interest can be chosen. The phantom can also be used to measure coronal and transverse dose distributions with radiographic film.

This simple phantom provides the accuracy and flexibility required for both IMRT technique development in the head and neck and for a thorough pre-treatment verification program.

ACKNOWLEDGMENTS

The authors acknowledge the early work carried out on the HANK phantom by J. Fairfoul, as well as the manufacture of the phantom and subsequent modifications carried out by the NWMP Workshop staff at the Christie Hospital.

REFERENCES

1. Galvin JM, Ezzell G, Eisbrauch A, et al. Implementing IMRT in clinical practice: a joint document of the American Society for Therapeutic Radiology and Oncology and the American Association of Physicists in Medicine. Int J Radiat Oncol Biol Phys. 2004;58(5):1616–1634.
2. Bedford JL, Childs PJ, Nordmark Hansen V, Mosleh-Shirazi MA, Verhaegen F, Warrington AP. Commissioning and quality assurance of the Pinnacle^3 radiotherapy treatment planning system for external beam photons. Br J Radiol. 2003;76(3):163–176.
3. Molineu A, Followill DS, Balter PA, et al. Design and implementation of an anthropomorphic quality assurance phantom for intensity-modulated radiation therapy for the Radiation Therapy Oncology Group. Int J Radiat Oncol Biol Phys. 2005;63(2):577–583.
4. International Commission on Radiation Units and Measurements (ICRU). Prescribing, recording, and reporting photon beam therapy. ICRU Report 62. Supplement to ICRU Report 50. Bethesda (MD): ICRU; 1999.
5. International Standard Randomized Controlled Trial Number (ISRCTN) Register. A multi-centred randomised study of parotid sparing intensity modulated radiotherapy (IMRT) to reduce xerostomia and increase quality of life in head and neck cancer [Web page]. No. 48243537. Bethesda (MD): Biomed Central; 2002. [Available at: www.controlled-trials.com/ISRCTN48243537/48243537; cited 07/02/08]
6. Papanikolaou N, Battista JJ, Boyer AL, et al. Tissue inhomogeneity corrections for megavoltage photon beams: report of Task Group No. 65 of the Radiation Therapy Committee of the American Association of Physicists in Medicine. Madison (WI): Medical Physics Publishing; 2004.

Publication #6

Accuracy and precision of an IGRT solution

Published in Medical Dosimetry, 34(2), 2009, p99-106

Authors:

Webster GJ

Rowbottom CG

Mackay RI

16 pages

Abstract

Image-guided radiotherapy (IGRT) can potentially improve the accuracy of delivery of radiotherapy treatments by providing high-quality images of patient anatomy in the treatment position that can be incorporated into the treatment setup. The achievable accuracy and precision of delivery of highly complex head-and-neck intensity modulated radiotherapy (IMRT) plans with an IGRT technique using an Elekta Synergy linear accelerator and the Pinnacle Treatment Planning System (TPS) was investigated. Four head-and-neck IMRT plans were delivered to a semi-anthropomorphic head-and-neck phantom and the dose distribution was measured simultaneously by up to 20 microMOSFET (metal oxide semiconductor field-effect transmitter) detectors. A volumetric kilovoltage (kV) x-ray image was then acquired in the treatment position, fused with the phantom scan within the TPS using Syntegra software, and used to recalculate the dose with the precise delivery isocenter at the actual position of each detector within the phantom. Three repeat measurements were made over a period of 2 months to reduce the effect of random errors in measurement or delivery. To ensure that the noise remained below 1.5% (1 SD), minimum doses of 85 cGy were delivered to each detector. The average measured dose was systematically 1.4% lower than predicted and was consistent between repeats. Over the 4 delivered plans, 10/76 measurements showed a systematic error > 3% (3/76 > 5%), for which several potential sources of error were investigated. The error was ultimately attributable to measurements made in beam penumbrae, where submillimeter positional errors result in large discrepancies in dose. The implementation of an image-guided technique improves the accuracy of dose verification, particularly within high-dose gradients. The achievable accuracy of complex IMRT dose delivery incorporating image-guidance is within ± 3% in dose over the range of sample points. For some points in high-dose gradients, submillimeter errors in position can lead to errors > 3%. The precision of the delivery system was demonstrated to be within the experimental noise of the detector system of 1.5% (1 SD).

Introduction

The increasing complexity and more specifically-targeted nature of modern radiotherapy treatment places greater emphasis on the need for precise and accurate radiation delivery compared to traditional treatment techniques. Recent developments, such as on-board kV-imaging facilities, now allow improved verification of treatment delivery accuracy by providing high quality images of the precise position of the patient and their internal anatomy [1,2]. The availability of this knowledge also provides an opportunity to adjust the patient position prior to treatment or to monitor the accuracy of dose delivery throughout treatment, a process termed image-guided radiotherapy (IGRT). These and other applications for this new technology are under widespread investigation, with various authors quantifying the inaccuracies expected from current techniques that do not incorporate an image-guided approach. There are currently no published results of investigations into the accuracy and precision of dose delivery achievable with IGRT for modern treatment techniques.

Traditional concepts for measuring the accuracy and precision of linear accelerators in terms of a series of simple mechanical and dosimetric parameters becomes a redundant concept when considering recent advances in IGRT techniques. Ultimately what is required is the ability to deliver planned dose gradients to a certain accuracy and within a certain precision.

Intensity modulated radiation therapy (IMRT) provides the ability to reliably and accurately deliver highly complex dose distributions on modern linear accelerators. IMRT therefore provides a more appropriate test for accuracy agreement with measurements and precision repeatability of measurements of IGRT than traditional tests. The higher dose gradients and improved conformality available with IMRT require increased levels of accuracy and precision. The ICRP recommend that an accuracy of 2% and precision of 1% should be the target requirement for IMRT delivery using IGRT.[3] This can only be achieved by having a high level of accuracy at

every stage of the process, including treatment planning, imaging and registration, dose delivery and the geometric calibration of the imaging linear accelerator and MLC.

The Christie Hospital is equipped with an Elekta Synergy linear accelerator capable of acquiring volumetric kV X-ray images (XVI). Previous work has established that microMOSFETs are well-suited for the purpose of dosimetric verification of IGRT as they are clearly visible on XVI images, produce minimal artefacts and have a very small active volume [4]. It was demonstrated that the IGRT system is capable of delivering dose to an imaged object with submillimeter accuracy. The aim of this work was to implement an integral system test of the accuracy and precision of the delivery of complex plans using IGRT. This aimed to quantify the mechanical aspects of the process and therefore excluded factors such as patient motion and inter-fractional changes in patient anatomy. To provide as rigorous a test as possible, head and neck IMRT plans were used, which are amongst the most complex plans in clinical use, delivering high conformality dose distributions between target and critical structures using highly modulated beams.

Materials and methods

A Perspex head and neck phantom (HANK) developed in-house for verification of IMRT techniques was modified for use in this work and is shown in Figure 1a [5]. The original design allowed a standard PTW30001 ionisation chamber (PTW, 0.125cm^3) or PinPoint chamber (PTW, 0.015cm^3) to be positioned in any of 32 positions within the head and 80 positions within the shoulders. Reconfiguring the phantom allowed a further 32 positions to be measured in the head section. For the purposes of this study, the number of available measurement points in the head was increased to 80, of which 20 can be used at any one time, which are all accessible simultaneously with the same phantom configuration. Perspex inserts of varying width were designed to allow precise lateral positioning of certain measurement points, at the expense of reducing the number available to 19. Suitable inserts were also made to securely and repeatedly

position the microMOSFET detectors and are shown in Figure 1c). A notation system had been developed to facilitate the easy use of the phantom and was extended to incorporate these changes. Figure 1b) illustrates the available axial positions and their notation, while the detectors can be positioned in predetermined fixed positions on the z-axis. In this paper, a measured point will be referred to by the plan number and the position within the phantom. So, for example, a measurement of the dose delivered to point 3B (see Fig.1b) by Plan 4 would be noted as 4:3B.

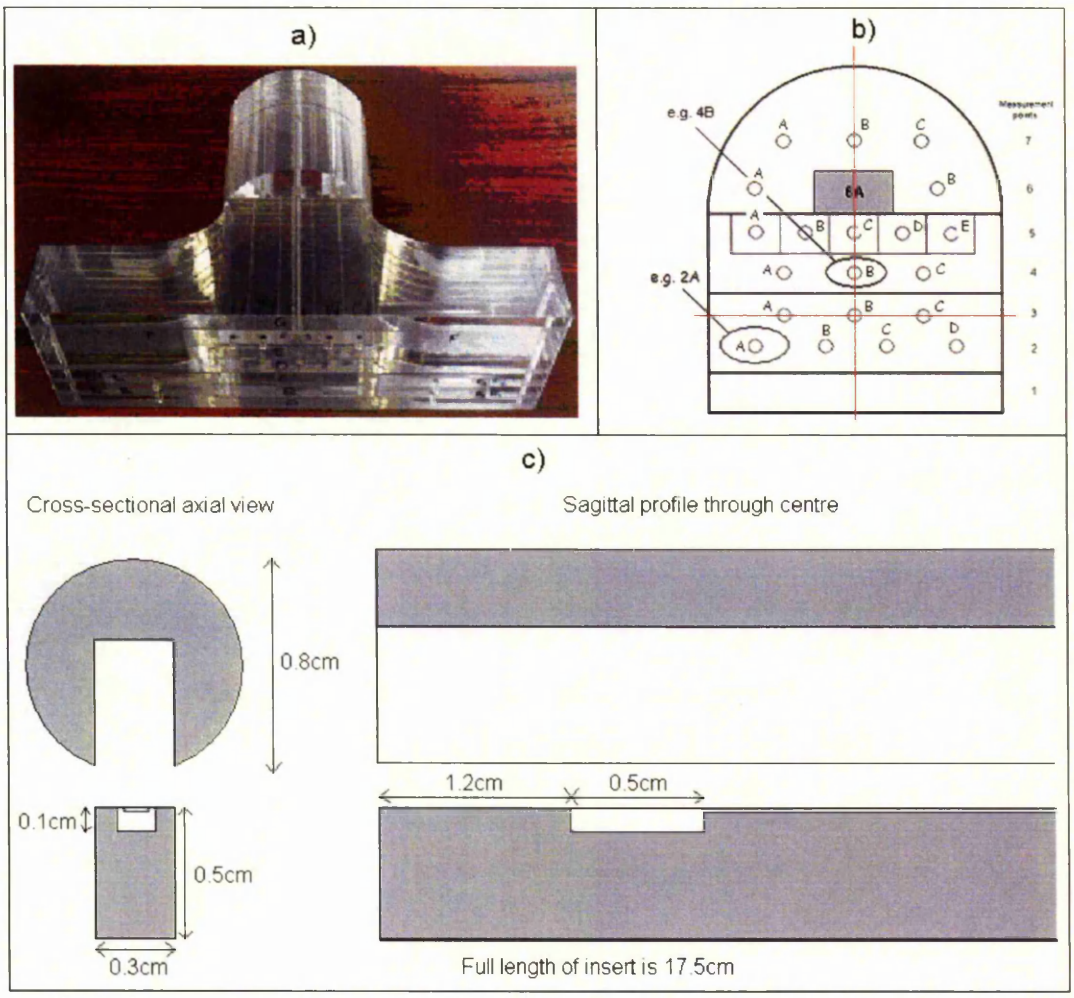


Fig.1. a) The HANK phantom and b) the notation system implemented to facilitate clinical use. The phantom can now measure 20 points simultaneously in any of 160 available points. Inserts developed to accurately and repeatedly position microMOSFET detectors within the phantom (c)

The measurement points were chosen to sample the dose distribution in high dose, low dose and high dose gradient regions. Four clinical head and neck IMRT dose distributions intended to allow parotid sparing were used for this study, each consisting of 5 6MV step-and-shoot beams with a total of between 64 and 68 segments delivered on an Elekta Synergy linear accelerator. In all cases the gantry angles were 0, 65, 144, 216 and 295 degrees. The treatment planning was carried out using Philips Pinnacle v7.4f with a beam model incorporating the effects of rounded leaf ends and transmission through the tongue and groove of the MLC leaves [6]. A 0.2cm resolution dose calculation grid was used for all plans and the contents of the phantom were overridden to the mean phantom density of 1.12g/cm^3 so as to remove the effect of density interpolation errors in air gaps on the CT scan of the phantom. All measurements were carried out on a Sinmed IMRT board, which was outlined from the fused image and incorporated into the dose calculation assuming a uniform density distribution. A mean value of 0.5g/cm^3 was chosen as this corresponded to 2% attenuation through the centre of the board in line with commissioning measurements.

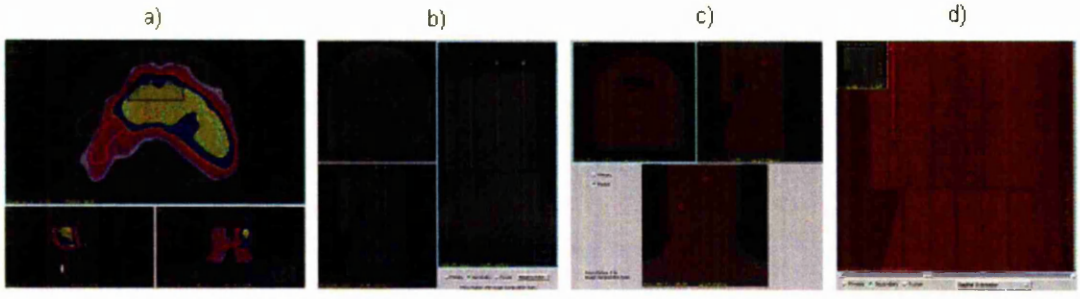


Fig.2. Procedure for testing the IGRT system. a) Initial IMRT dose distribution copied to verification phantom b) the XVI image of the phantom clearly showing the microMOSFETs c) the image registration within Pinnacle allows the removal of setup error when comparing planned and delivered doses d) a sagittal view of the phantom showing how the microMOSFET positions are obtained

The verification process is illustrated in Figure 2 and begins with positioning the HANK phantom appropriately on the couch and positioning each of the detectors in their predetermined positions. The phantom is then scanned using the XVI imager and the image transferred to the

Treatment Planning System where it is fused with the original phantom scan on which the dose distribution has been calculated. The effects of setup error can be removed at this stage by using the fused image to determine the exact position of the delivery isocentre and recalculating the dose distribution using this point. This process also removes errors in the positioning of the detector as it allows precise localisation of each microMOSFET (see Figure 2c) by identifying the tip and moving 0.8mm superiorly to the active centre [4]. The plan is then delivered and the measured readings from each detector are obtained at the end of each fraction for comparison with the calculated results.

The microMOSFETs are subject to significant uncertainties due to noise, which are particularly prevalent at low doses. They can be used in either standard or high sensitivity settings, which allow a compromise between detector life (the microMOSFETs expire at 20,000V) and signal-to-noise ratio (SNR). The detectors used here were investigated to determine the low-dose threshold at which the SNR ratio begins to increase. This was done by repeatedly delivering a range of doses with a 10x10 field at 6MV to several detectors in standard calibration conditions (using a specifically-designed calibration insert) and quantifying the noise to determine the minimum dose threshold. These results would then indicate the dose required to calibrate each detector. This calibration was done using two repeat deliveries of 100MU in standard calibration conditions on four occasions over the course of the investigation period and using the mean readings to establish a calibration factor for each detector. The delivered dose prescriptions could ultimately be delivered with higher MU if required to ensure that all measured doses had an optimal SNR. To ensure that there is no error in the overall reading due to latency in measurement at low MU (i.e. for individual segments within a beam) the cumulative dose was compared for ten 10MU deliveries and one 100MU delivery using both the microMOSFETs and a calibrated Farmer chamber (PTW, 0.6cm^3).

The same readings were obtained on three separate occasions over 3 months with different detectors in each of the measurement points. The measured readings were subject to a daily

output correction, established using a Farmer chamber traceable to the NPL standard (PTW, 0.6cm^3). The machine was not set up specifically for this test and so the results represent a clinically realistic delivery. This process ensured that any errors were not due either to significant random errors in beam delivery or subject to a faulty detector, each of which would be identified by a high standard deviation on the three readings.

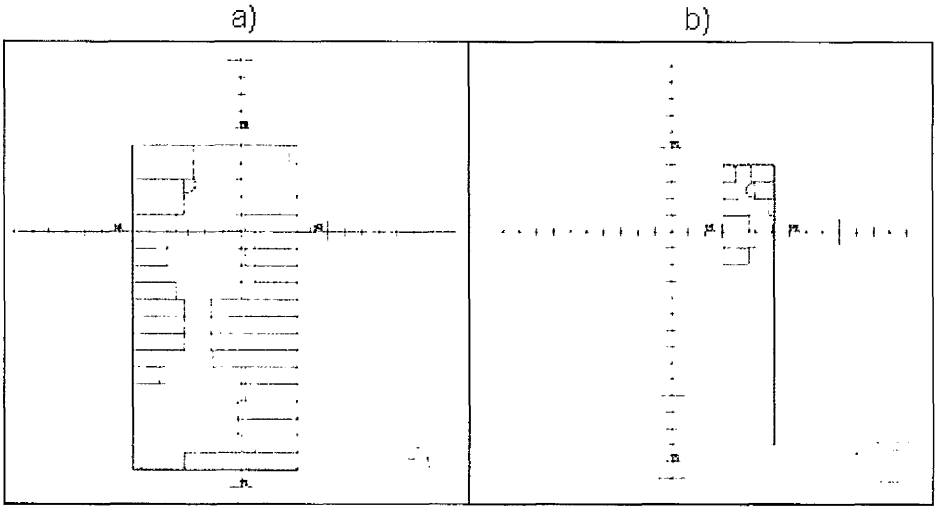


Fig. 3. Examples of segments with measurement points that would be expected to be less accurately calculated in Pinnacle

The possibility of errors due to sub-optimal beam modelling was explored by making fine adjustments to sensitive modelling parameters in the Pinnacle v7.4f beam modelling software and quantifying their impact on the overall agreement for all points. The groove width and radius of curvature of the MLC leaves were originally set according to manufacturer recommendations at 0.1cm and 15cm respectively. The doses were therefore recalculated with alternative groove widths (0.0 and 0.15cm) and leaf radius of curvature (5 and 25cm). 0.5mm shifts in the optimised leaf offset table were also investigated. A limitation of the TPS is the presence of interpolation errors in penumbra modelling due to the finite resolution of the model calculations (0.2cm). This effect was quantified by recalculating the dose to the three measurement points with poorest

agreement by shifting the dose grid in 0.5mm increments across the penumbra of the overall dose distribution. This allows the impact of reducing the dose grid to 0.5mm to be quantified.

Measurement points that revealed errors above the investigation level (here defined as 2 x SD. of the detector) were remeasured using a PinPoint chamber (PTW, 0.015cm^3). This chamber has very good spatial resolution and allows the dose to be measured from individual segments, which would require impractically high MU delivery with the microMOSFETs. The resulting 'Pinpoint error' was then compared to a 'Corrected error'. This 'Corrected error' was calculated by removing segments in which the measurement point falls within a beam penumbra, defined here as 0.5cm either side of the field edge, which would thus be expected to result in a significant error (>5%) for even a submillimetre error in position, examples of which are given in Figure 3. The 'Corrected error' is therefore the error on the remaining segments that the TPS could be reasonably expected to calculate accurately, or the error due to delivery and measurement rather than calculation. The overall impact of the image-guided process was quantified for these same points by comparing the error using the predicted delivery isocentre with that using the isocentre determined with the XVI.

Results

Initial experiments demonstrated the relationship between the coefficient of variation and dose on the microMOSFET. The shapes of the curve for both standard and high sensitivity bias settings were similar with a minimum dose of 85cGy required to produce a consistently low level of experimental noise. When using microMOSFETs with the standard bias the variation above 85cGy was approximately 3% of the measured dose. For the high sensitivity bias this was reduced to 1.5%. Figure 4 shows the curve for the high sensitivity bias. Measuring less than 85cGy led to a marked increase in the detector noise. For this reason, subsequent measurements were designed to coincide with planned doses >85cGy for a single fraction. An increase in the prescribed MUs of 1.5x was found to be sufficient to ensure that the lowest

measured dose points exceeded this level and so all plans were delivered with a 50% increase in MU. The latency in the measurement of low MU segments was approximately 0.5% using the microMOSFETs and <0.1% using the calibrated Farmer chamber (PTW, 0.6cm^3).

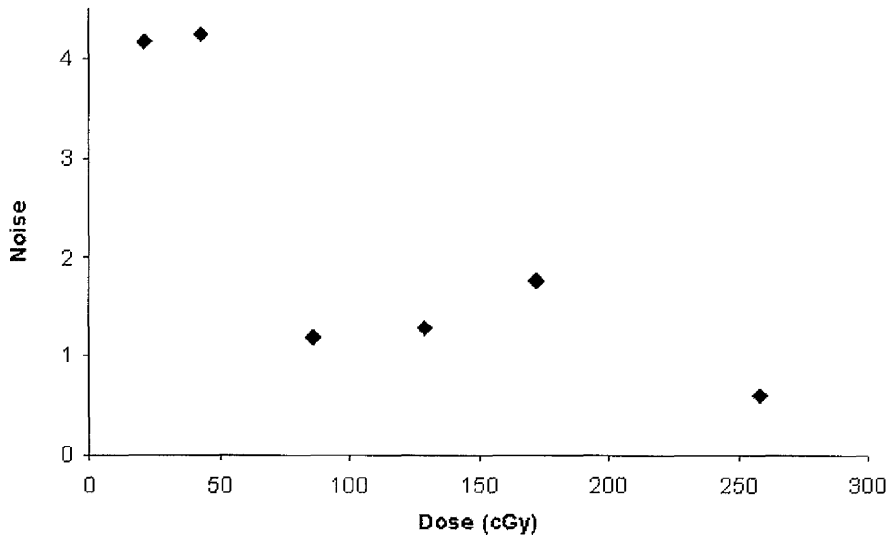


Fig. 4. Detector noise versus measured dose over the range 20cGy to 260cGy. Detector noise is defined as the coefficient of variation (COV) where COV(%) = 100 x 2SD/mean)

Figure 4 demonstrates that differences between measured and planned doses of <1.5% will be difficult to resolve with the present system. The stability of individual microMOSFETs have also been investigated by repeat measurement over several days with a mean standard deviation of ~1.5% for a high sensitivity bias microMOSFET, the same order of magnitude as the coefficient of variation on readings for a single experimental session.

The plan measurement results are shown in Figure 5 for all points in all plans and indicate a majority of measurements within 3% of calculated values. Of the 76 points measured over the 4 delivered plans, ten exceeded the 3% investigation level, with only three exceeding a 5% error. These 10 points were further investigated. The repeated measurements at each point were

largely consistent, with only 5/76 points having a coefficient of variation (COV) over three readings of >6%. These points were also found to exceed the investigation level and so underwent further analysis. This consistency between detectors when their position is alternated indicates that there are no faulty detectors. The repeatability of measurement also confirms consistency between deliveries and suggests a reduced likelihood of random errors.

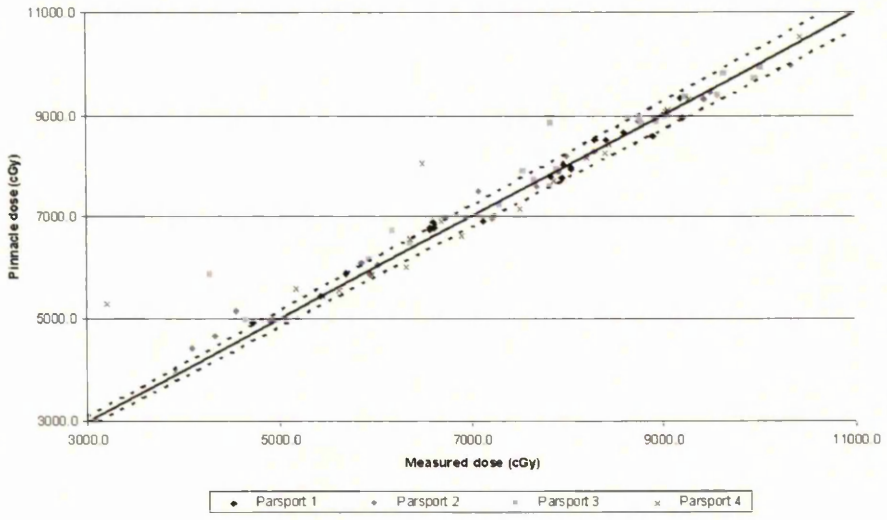


Fig. 5. Plot of average measured against calculated dose for all points for all 4 clinical plans. Solid line represents perfect agreement while the dotted lines indicate the extent of the 3% tolerance

As illustrated in Figure 6, changes to the beam model within the Pinnacle TPS did not make a significant or consistent improvement to the calculation accuracy. Shifting the dose grid to counter interpolation errors in the penumbra modelling was also found to give no better agreement with the measured dose so the points exceeding this investigation level were remeasured using the PinPoint chamber (PTW, 0.015cm^3) to determine the error in the dose calculation of each segment. The discrepancies between the 'Pinpoint' and 'MOSFET' errors seen in Figure 7 is due to there being a slight difference in the effective point of measurement, which is systematically 1.6mm inferior for the microMOSFET due to a difference in the manufacture of the phantom inserts used to position the two different detectors.

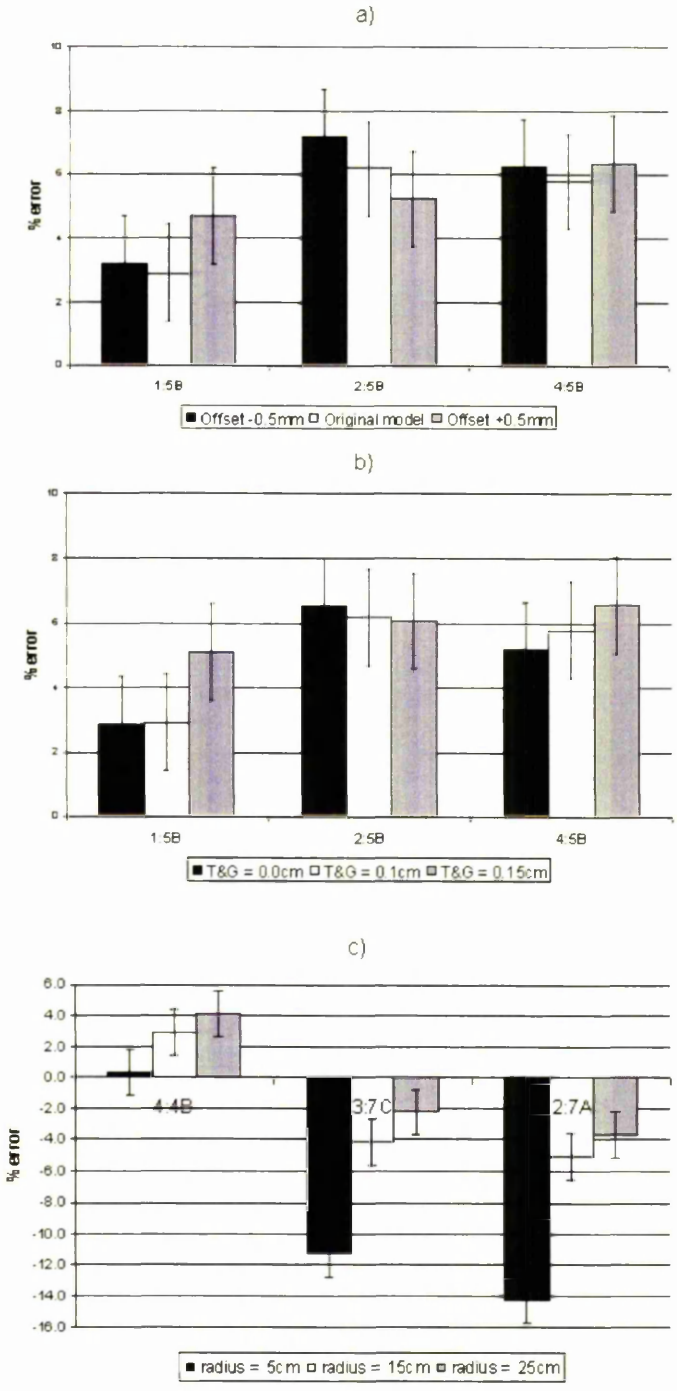


Fig. 6. Effect of varying a) MLC offset, b) groove width and c) leaf end radius of curvature

To test the 'Corrected error' the segments within 0.5cm of the field edge, a mean of 12 segments per measurement point (range 4 to 18) from the original 64-68 segments, were selected from each plan. The results in Figure 7 indicate that the error on each point is reduced to within the investigation level when segments measuring penumbra (within 0.5cm of the field edge) are not included in the comparison. The impact of the IGRT process is illustrated in Figure 8, showing that for 5/10 points the effect was to improve agreement, pulling the error within tolerance, in 1 case the error increased beyond tolerance and in 4 there was no significant impact.

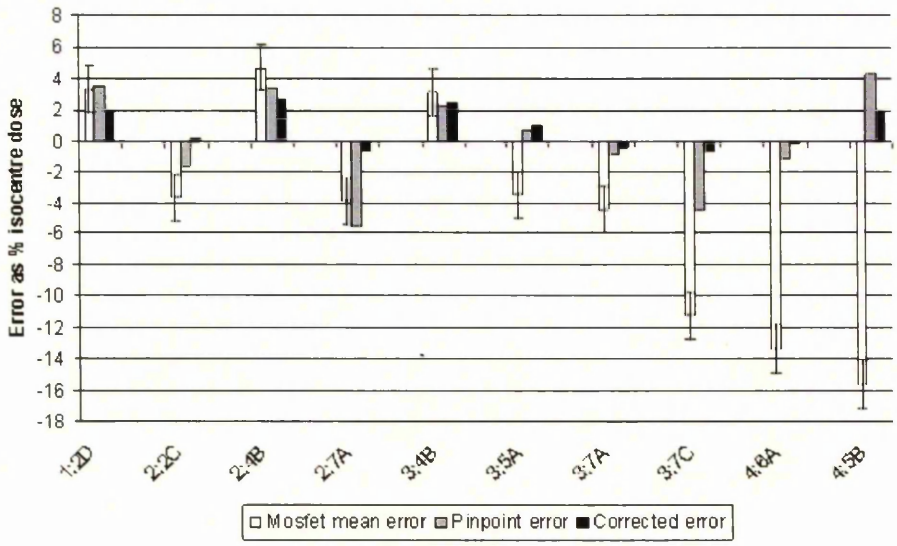


Fig.7. Remeasurement with PinPoint chamber (PTW, 0.015cm^3) of points exceeding investigation level. 'Corrected error' is the residual error after points within 0.5cm of the field edge have been removed from the measurement and calculation

Discussion

Four clinical head and neck IMRT plans were delivered and a total of 76 points measured on two or three occasions, depending upon the lifetime of the individual microMOSFETs, some of which reached their 20,000V limit during the study. Only 10 of these measurements were found to

exceed the investigation level of 3% error in absolute dose. Several possible sources of error were investigated and found to be insufficient to explain these errors.

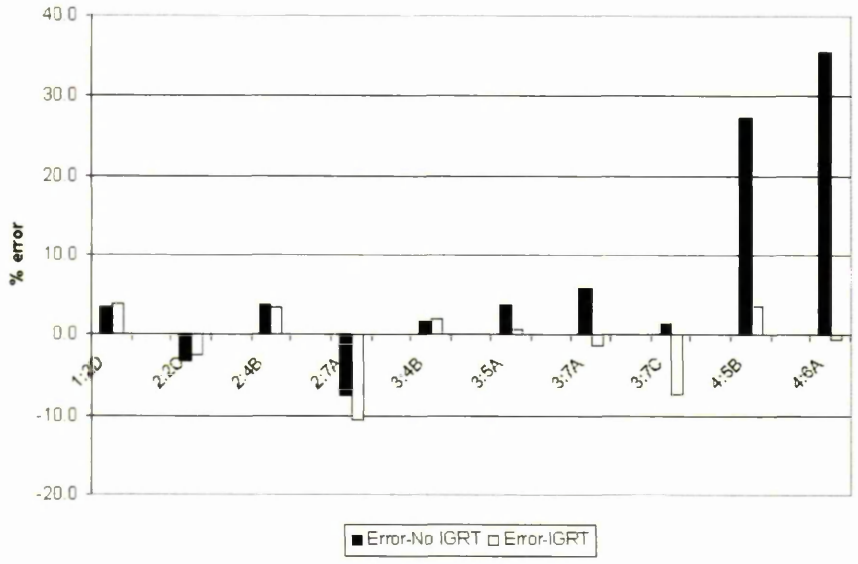


Fig. 8. Impact of IGRT on agreement between measurement and calculation.

Repeat measurements with the microMOSFETs reduced the random errors for all measurement points, while the possibility of a false-negative result due to a faulty detector was removed by changing their position for each repeat measurement. The repeated measurements over a period of three months also ensured that beam delivery errors, such as MLC positioning or beam symmetry, for example, would not be a consistent cause of error. The risks of setup or detector-positioning errors were addressed by the image-guided approach that ensured precise knowledge of the isocentre, chamber and phantom positions at the time of delivery.

Possible errors in the dose calculation due to sub-optimal beam modelling were investigated by altering groove width and MLC positions but were not found to improve the agreement. Ultimately, it was demonstrated that the errors in the 10 measurement points were due to their

precise position, since when points measured in positions susceptible to sub-millimetre positional errors were not considered, the errors were reduced to within tolerance.

This work represents the quantification of the accuracy and precision of the technical aspects of IGRT. As such, it represents the optimal achievable accuracy and precision before patient-induced errors are taken into account. The results are perhaps applicable to rigid bony anatomy registration but would be subject to errors from the relative positions of bony anatomy within the patient, for example jaw motion within an immobilisation shell for a head and neck patient, or the position of soft tissue relative to the bony anatomy used for registration. The extent of the errors due to the latter problem will be dependent upon treatment site; a recent study by Prevost et al suggests that, even for head and neck, this could be relevant, finding oropharyngeal target shifts of 1-2mm due to breathing motion.[7]

Conclusion

A system has been developed to test the accuracy and precision of the technical aspects of an IGRT technique. The main rationale for such a test was to establish the clinically relevant accuracy and precision for delivering a highly complex dose distribution via IMRT. MicroMOSFETs are suitable dosimeters for this test, although with the doses used the experimental variation is ~1.5%. Different dosimeters would be required to detect accuracy and precision below this level.

It has been confirmed that the achievable accuracy of a complex IMRT dose delivery incorporating an image guided approach using Elekta Synergy and the Pinnacle TPS (v7.4f) is within +/-3% in dose over the range of sample points including within high dose gradients. The precision of the delivery system was also demonstrated to be within the experimental noise of the system of 1.5% (1SD) for high-sensitivity microMOSFETs.

References

1. Marchant, T.E.; Amer, A.M.; Moore, C.J.; Stratford, J. Comparison of set up errors measured in 2D and 3D using EPI and cone beam CT. *Radiotherapy & Oncology* (Oral presentation at 8th Biennial ESTRO Meeting on Physics and Radiation Technology for Clinical Radiotherapy, Lisboa, September 2005) *abstracted in* **76** (S2) S128; 2005

2. Borst, G.R; Sonke, J-J.; Betgen, A.; Remeijer, P.; van Herk, M.; Lebesque, J.V. Kilo-Voltage Cone-Beam Computed Tomography Setup Measurements for Lung Cancer Patients; First Clinical Results and Comparison With Electronic Portal-Imaging Device. *Int. J. Radiat. Oncol. Biol. Phys.* **68**:555-561; 2007

3. Prevention of accidental exposures to patients undergoing radiation therapy: ICRP Publication 86 (2001)

4. Rowbottom, C.G.; Jaffray, D.A. Development of an integral system test for image-guided radiotherapy. *Medical Physics,* **31**(12):3500-3505; 2004

5. Webster, G.J.; Hardy, M.J.; Rowbottom, C.G.; Mackay, R.I. Design and Implementation of a Head & Neck Phantom (HANK) for System Audit and Verification of IMRT. *J App Clin Med Phy.* 9(2), 2008

6. Webster, G.J.; Rowbottom, C.G.; Mackay, R.I. Development of an optimum photon beam model for head and neck IMRT. *J App Clin Med Phy.* **8**(4); 2007

7. Prevost J.B.; de Boer, H.; Poll, J.; Voet, P.; Levendag, P. Analysis of the motion of oropharyngeal tumours and consequences in planning target volume determination. *Radiotherapy & Oncology* (2008), doi:10.1016/j.radonc.2007.12.029

3 Verification of IMRT dose delivery to the patient

3.1 Introduction

The increased complexity of IMRT for the head and neck renders conventional methods of treatment verification and monitor unit (MU) checking, such as the manual dose calculation to a stable point within the dose distribution, inadequate: the highly -modulated nature of delivery and the high incidence of small, off-axis segments means that large regions of the dose distribution are due to the cumulative dose from several beam penumbra, which are not adequately checked by conventional methods. The particular complexity of the head and neck, with often greater heterogeneity and variation in external contour than other sites, means that patient-specific verification of all clinically important regions of individual treatments is critical [38; 39].

Detailed recommendations for pre-treatment verification are given in IPEM96: Guidance for the Clinical Implementation of Intensity-Modulated Radiation Therapy [22]. A commonly used verification process for IMRT, and that used locally, involves 2 stages: (i) the plan is transferred to a phantom and the dose in the high dose region measured and compared with calculation; (ii) some form of 2D dose verification is carried out, either using a film within a phantom to measure the combined dose distribution in a particular position and orientation, or by verifying individual beam fluences. This could be carried out using a semi-anthropomorphic phantom, which may have a realistic external contour or internal medium that would provide a reasonable representation of the actual patient [40; 41], although it is often performed within less representative block phantoms or, increasingly, using a 2D dosimetric array to obtain verification of both absolute dose and relative dose distribution. This process can be very time-consuming and, in centres where IMRT is routinely used, could represent the limiting factor to wider implementation.

To speed up the verification process, attempts have been made to write IMRT-capable MU verification software [42]. These rely on a sophisticated beam model that incorporates head

scatter and phantom scatter. For simplicity, they are carried out on a phantom or homogeneous representation of the patient contour. These significantly speed up the verification process with Georg et al estimating a factor of 4 to 5 improvement in efficiency. However, they do not give a sense of the accuracy of the full dose distribution, instead providing a check of the monitor units impacting on a particular point in the plan.

This inability to assess the full dose distribution is also a major limitation of both the absolute point dose measurement and fluence map verification methods as they provide only a limited 2D sample of the dose distribution, which could easily fail to highlight an error in either the delivery or calculation of some aspect of the plan. Acquisition of several measured slices, in order to build up a more complete picture of the dose distribution would exacerbate the inefficiency of the process while still only providing a limited sample of the dose distribution.

Ideally what is needed is a true 3D measurement of the dose distribution. Gel dosimetry has often been studied in relation to this but has been associated with problems regarding resolution and noise for verification of more complex plans [43], as well as logistical problems due to the requirement for magnetic resonance imaging of the irradiated phantom. A recent study by Oldham et al shows promise for the use of this application in the verification of complex IMRT plans (albeit not a realistic clinical distribution) by using an optical CT system to scan an irradiated cylindrical gel phantom. The results showed good agreement with calculation, finding a 96% pass rate for a tolerance of 3%/3mm. If similar results are acquired for clinical plans then this method could be very useful, although availability of the scanning hardware may be a limiting factor [44].

All of the current solutions to the problem of IMRT verification are further limited by the fact that they are applied to a phantom prior to calculation and measurement. This means that the precise location of any discrepancies, and therefore their likely clinical impact, must be inferred by approximating the location of the errors on the patient dataset and estimating their overall impact. Ideally, pre-treatment verification would directly demonstrate the impact of any delivery or planning errors on the dose distribution to the actual patient [38]. This has

been attempted using back-projection of EPID images to multiple planes within the patient for each beam [23; 45]. The EPID images are deconvolved to remove the effects of lateral scatter within the EPID before being modified account for the increased distance of the EPID measurement surface relative to the patient, scatter from the patient to the EPID and attenuation through the patient using an assumption of homogeneity within the external contour of the patient. The authors demonstrate good agreement for 9 rectum patients (within gamma tolerance of 2%/2mm) except where large gas build-up is present. This inability to accurately account for inhomogeneities, integral to the technique due to the lack of a robust algorithm for the dose calculation, limits the application of this technique to sites such as the head and neck.

This work aims to improve on the above work by converting measured film fluences to corresponding in-air fluences that can be input into the TPS and projected through the patient dataset, leading to a calculation of the delivered dose distribution to the patient geometry with the dose algorithm contained within the TPS. This would permit a direct evaluation of the accuracy of IMRT delivery in the patient for the first time, allowing clinically relevant tolerances, probably involving both the magnitude and location of any errors, to be established.

3.2 Materials and methods

The proposed methodology, outlined in Fig 3.1, aims to directly calculate the delivered dose in the patient dataset by converting a measured fluence film into a suitable in-air fluence and projecting this through the density representation of the patient. Individual beam fluence films are measured at d_{max} in a solid water slab phantom using Kodak EDR2 film, scanned at a resolution of 0.4mm. The uploading of the in-air fluence to the TPS is achieved using a simple Unix script that also defines the size and resolution of the uploaded array. The clinical plans investigated incorporated seven 6MV step-and-shoot beams, optimised using Direct Machine Parameter Optimisation (DMPO) software (Pinnacle v7.4f), with typically a total of 50-60 segments delivered on an Elekta Synergy linear accelerator.

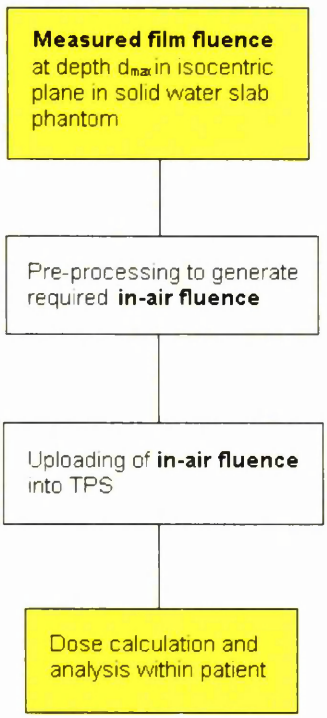


Figure 3.1. Process for feedback of measured fluence maps into Pinnacle TPS

3.3 Analytical methodology

When the in-air fluence is projected through a CT dataset, a number of corrections are applied prior to calculation of the resulting dose distribution. These corrections are incorporated into a user-optimised beam model and are intended to provide simplified representations of the interaction processes that occur as the beam passes through the treatment head and into the patient. If the measured film fluence was uploaded directly to the TPS to represent the in-air fluence, these corrections, the 'real' effect of which will be present in the film, would effectively be applied twice if the clinical beam model were used. This effect has been confirmed by uploading a calculated dosemap (defined here as the dose calculated at d_{max} in solid water, which in clinical practice is verified against the measured film fluence) as the in-air fluence and then calculating the resulting dosemap. This would ideally result in the subsequently calculated dosemap being identical to the original, but the

result illustrated in Fig 3.2 indicates a blurring of the dose distribution and disagreement in off-axis dose.

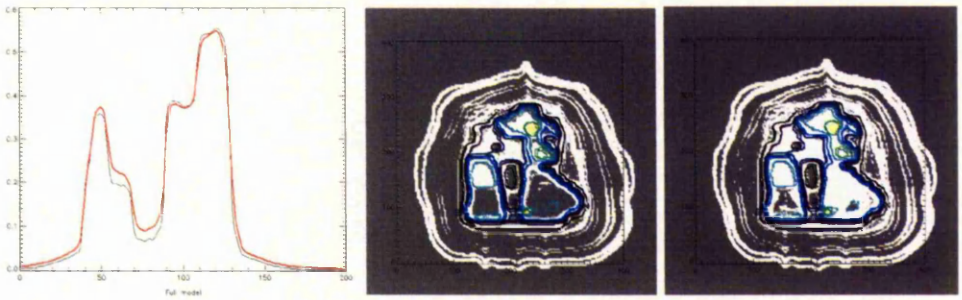


Figure 3.2. a) Central axis overlay of feedback of original calculated dosemap (black) uploaded as in-air fluence and recalculated as a subsequent dosemap (red). Corresponding gamma analyses at b) 2%/2mm and c) 4%/4mm (grey areas indicate regions exceeding tolerance)

This confirms that the removal of the corrections applied during the dose calculation is necessary prior to projection of the in-air fluence into the patient. The impact of each of these corrections must be understood if they are to be safely removed from the beam model and are therefore described in detail in the next section.

3.3.1 Dose calculation algorithm

The beam model within the Pinnacle TPS used for this work has been specifically tuned to optimise the accuracy for head and neck IMRT calculations, as outlined in Publication #2 (Webster et al, 2007). A detailed description of the dose calculation algorithm has been provided by Starkschall et al [46; 47] but the aspects of the model relevant for the purposes of this work are again outlined here and in Fig 3.3.

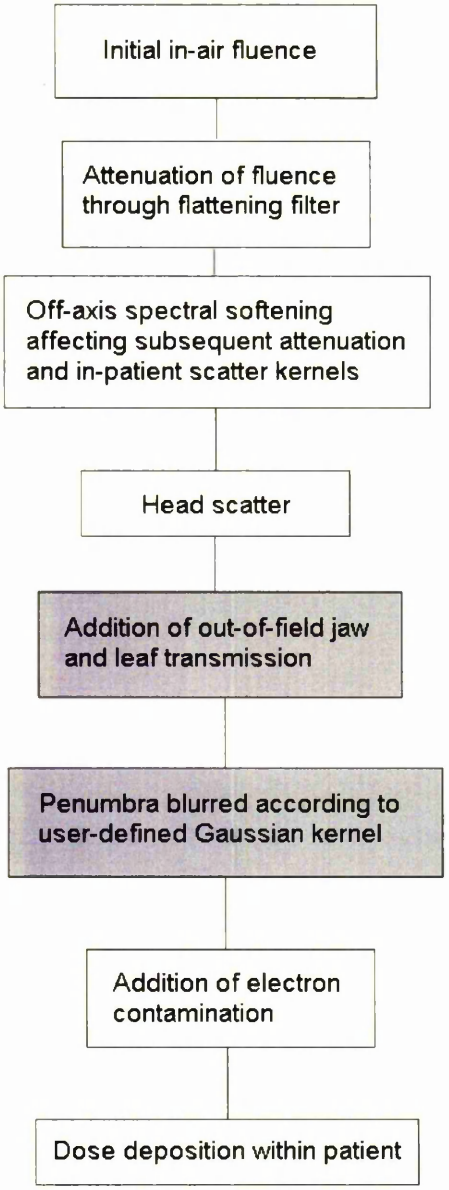


Figure 3.3. Flowchart showing corrections applied to the in-air fluence by the TPS to calculate dose distribution. Steps highlighted in grey are not applied in 'intensity modulation' mode

The collapsed cone convolution algorithm is used, which describes each element of the dose distribution as the convolution of the total energy released per unit mass (terma, T(r)) with a polyenergetic dose-spread kernel (A(r-r')) that represents the relative energy deposition in the vicinity of the primary interaction at r. The calculated kernel is a weighted average of monoenergetic dose-spread kernels established using Monte Carlo modelling, over the

energy spectrum of the beam fluence at r. The dose at a point r, D(r), due to photon scatter within the patient is therefore given by the formula

$$D(r) = \int_{V} d^3 r T(r') \overline{A}(r, r - r')$$

where the terma is defined as

$$T(r') = \int_{E} dE \frac{\mu}{\rho}(r', E) \frac{d\Psi(r', E)}{dE}$$

It is apparent from this formula that both the mass attenuation coefficient (μ/ρ) and the energy fluence (dΨ) are dependent on both position and energy of the interaction. In large part, the energy fluence is a function of the radial off-axis position due to the effects of the flattening filter. The flattening filter is a radially-symmetric stainless steel cone that preferentially attenuates the central axis of the beam ejected from the waveguide to give a roughly flat beam at a clinically relevant depth in the patient. The TPS models this variable attenuation by applying a simple correction factor, which is a function of radial distance from the central axis, to the initial fluence. An unwanted side-effect of the flattening filter is that the more highly-attenuated central axis component of the beam has had a greater proportion of the low-energy spectral component filtered out than the less attenuated peripheral components of the beam. This leads to an off-axis softening of the beam at greater distances from the central axis, modelled in the TPS by a power-law component that alters the energy spectrum as a function of the off-axis distance, which in turn alters the mass attenuation coefficient and the polyenergetic scatter kernel.

The additive contribution of scatter from the flattening filter to the beam fluence is modelled by a convolution of a Gaussian distribution of user-defined parameters, intended to represent the source of scatter, with a unit mask defining the field edge. Penumbral blurring is similarly modelled using a Gaussian distribution to define the focal spot blurring function and the field edges to determine the exposure of each photon interaction point to this source. Out-of-field

doses, resulting from transmission through jaws and MLC, are again modelled as additive contributions based on user-specified transmission factors.

A series of complex equations are used to model the dose due to beam contamination from electrons produced by photon interactions in the treatment head. The relevant parameters used in the current beam model were not thoroughly verified as part of the beam commissioning process as it had minimal impact at depths of clinical interest. However, for fluence maps measured at the depth-dose maximum, this effect may be significant, particularly the off-axis factor, $F_{OA}(r)$, which will introduce a non-uniformity into the calculation

$$F_{OA}(r) = e^{-A\theta^2}$$

where θ is the angular displacement of the line to point r from the central axis and A is a user-defined parameter. The value of A used in the current beam model is 0, suggesting no off-axis variation.

Several of the required correction factors would vary as a function of depth, but a limitation of the Pinnacle TPS is that each must be a compromise based on a best fit to the data measured at a range of clinical depths. In practice, where compromise is necessary, it will tend to be the depths of likely clinical interest for which best agreement is sought (for 6MV this would tend to be at depths of 5-10cm).

3.3.2 Impact of beam model corrections

3.3.2.1 Jaw and leaf transmission; focal spot size

The above description of the dose calculation algorithm is applicable for all calculations carried out using DMPO-optimised beams, which define the jaw and MLC specifications required for delivery. However, the ability to override the existing in-air fluence is available only with 'intensity modulation' beams. These correspond to an earlier stage of optimisation in which no leaf and jaw parameters have been set. Referring to Fig 3.3, when dose calculations within the patient are subsequently applied in 'intensity modulation' mode, the in-air fluence does not undergo the corrections highlighted in grey, since the jaw and leaf positions that inform them are not established.

The absence of these corrections was confirmed by calculating the dosemap resulting from a flat 10x10cm in-air fluence with both the optimised beam model, followed by subsequent beam models with each of these corrections in turn removed from the beam model. The results shown in Fig 3.4 demonstrate no impact from removing corrections for either jaw transmission or focal spot size in 'intensity modulation' mode.

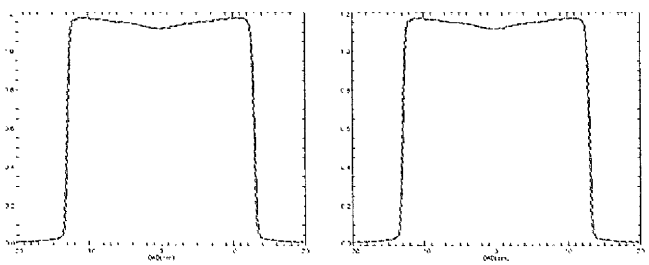


Figure 3.4. Profile overlays of dosemap calculation at d_{max} in solid water slab phantom of 10x10cm flat in-air fluence with (black) and without (red) corrections for a) jaw and MLC transmission, b) penumbral blurring due to focal spot size

3.3.2.2 Electron contamination; attenuation from flattening filter; head scatter

Fig 3.5a) suggests a uniform low-level additive effect of electron contamination, although this will vary as a function of depth. However, since the effect is negligible at clinically-relevant depths (<0.01Gy beyond 2cm) it can be removed from the beam model without adversely affecting the accuracy of the calculation.

The corrections for attenuation of the beam by the flattening filter and the generation of head scatter are illustrated in Fig 3.5b)&c) respectively, both of which show the radially-symmetric nature of the corrections, with the former being of a larger magnitude. Since these corrections do not alter the energy spectrum of the beam, and are merely correction masks applied to the in-air fluence prior to interaction with the patient, they can be safely removed from the beam model without adversely affecting the accuracy of the resulting dose calculation. The accuracy of these corrections in terms of their agreement with the physical reality is irrelevant for the purposes of this study as the 'real' effect is already present in the measured film fluence, while the correction, which will be associated with an inaccuracy, is not applied.

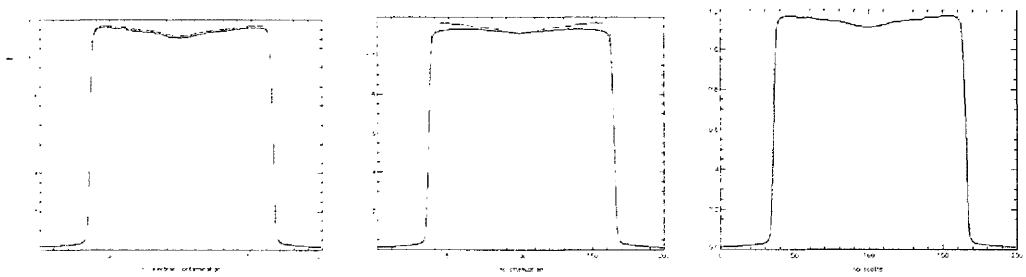


Figure 3.5. Profile overlays of dosemap calculation at d_{max} in solid water slab phantom of 10x10cm flat in-air fluence with (black) and without (red) corrections for a) electron contamination, b) flattening filter attenuation and c) head scatter

3.3.2.3 Off-axis softening

The softening of the energy spectrum away from the central axis results in two effects: variable attenuation and variable scatter deposition as a function of off-axis distance (the

latter is discussed in section 3.3.3). Based on the changing energy of the beam off-axis quantified within the beam model, the corresponding mass attenuation coefficients in water [48] suggest a variation in attenuation of <1% at depth d_{max} within the allowed 40x40cm field. This error is sufficiently small that no correction for changing attenuation is required.

However, this variation in attenuation across the field leads to a changing depth dose curve that becomes more significant with depth. The off-axis function of the energy spectrum therefore forms an integral part of the dose calculation, rather than merely providing a modification to the in-air fluence, and cannot be removed from the beam model without compromising accuracy. The off-axis softening correction is therefore maintained within the beam model

3.3.3. In-patient scatter effects

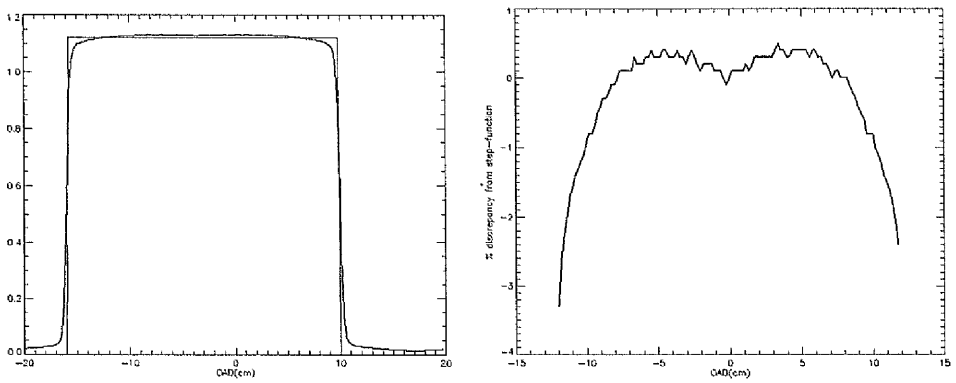


Figure 3.6. Central-axis profile of a) calculated dosemap (red) resulting from flat 10x10cm in-air fluence (black). b) shows the deviation of the resulting profile from the intended function

Fig 3.6 shows the effect of setting the beam model as described above and calculating the dosemap resulting from the flat 10x10cm in-air fluence. It is apparent that the resulting dose profile is a good approximation to the step-function that we would ideally like, with errors <0.5% in low-gradient regions, although this increases markedly as the penumbra are approached. This will likely be due to the blurring effect of the scattered dose deposition

within the phantom. Two approaches to correcting for in-phantom (or in-patient) scatter have been explored in this work and are discussed in detail here.

3.3.3.1 Sharpening of ODM penumbra

Code has been written in IDL to define the regions of gradient in the image, in both the horizontal and vertical axes, using Sobel operators in conjunction with a threshold above which a gradient is defined (an example gradient mask is shown in Fig 3.7a). Working on each axis in turn, the gradients are then sharpened according to the conditions given in Fig 3.7b, in which each pixel within the gradient is assigned to either the higher or lower intensity value depending on the relative proximity of the edges of the gradient region.

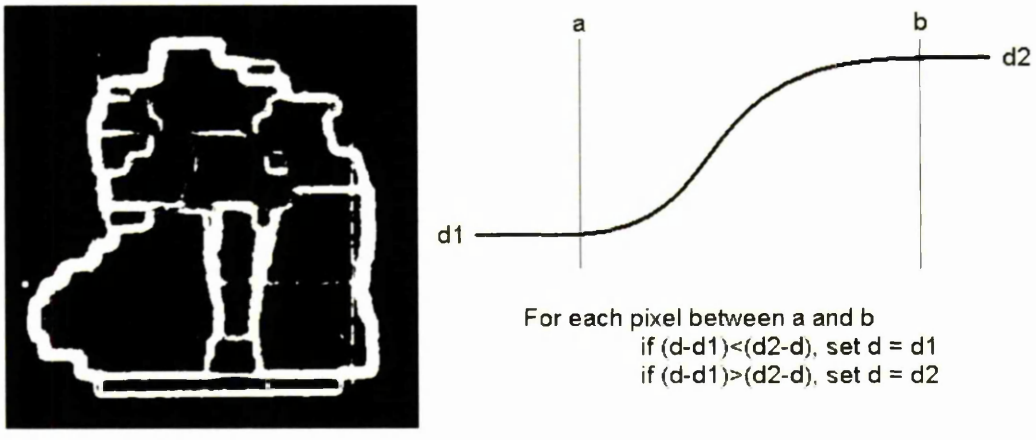


Figure 3.7. a) gradient mask over which correction is applied, b) assignation of intensity within gradient regions based on relative proximity to high and low intensity plateaus

Using this methodology, together with the beam model outlined in section 3.3.2, resulted in some benefits in terms of more closely replicating the required beam penumbra. However, although this approach could be further optimised by changing the threshold above which a gradient is defined, or by altering the assignation parameters for the pixels within gradient regions, it is fundamentally incapable of removing the blurring introduced to peaks and

troughs within the image. These can be of considerable importance in highly modulated fields and so an alternative methodology was sought.

3.3.3.2 Deconvolution

The in-air fluence (f) is subject to a uniform, isotropic blurring kernel (k) that is applied equally to all points to give the resulting dose distribution (d). Conversely, this Gaussian blurring kernel can be removed from the resulting image of the dose distribution by deconvolution according to the formula

$$d = f \otimes k$$

which can be expressed in Fourier terms as a simple arithmetic calculation to give the Fourier Transform of the required in-air fluence (F)

$$D = F \times K \qquad \text{or} \qquad F = {}^{D}\!/_{K}$$

where D, F, and K are the Fourier Transforms of d, f and k respectively. An inverse Fourier Transform can then be applied to give the in-air fluence, f, deconvolved to remove the effects of in-patient scatter.

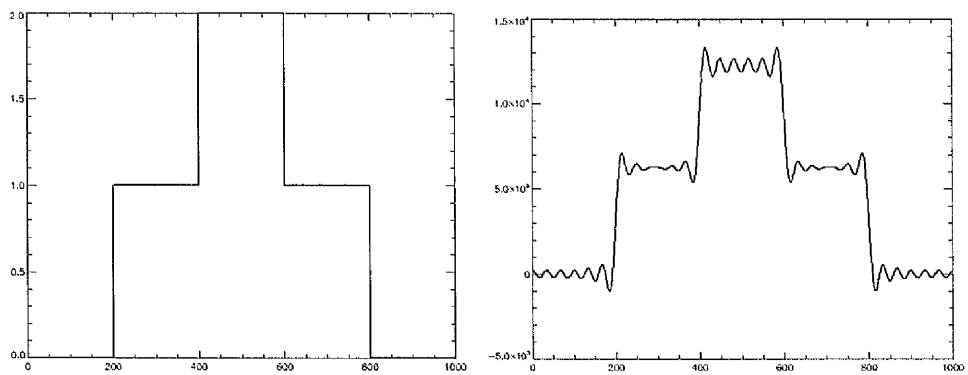


Figure 3.8. a) a simple step-function b) is convolved and then deconvolved using the same Gaussian kernel

It was observed at an early stage that this process introduces noise into low-gradient areas of the image, as can be seen in Fig 3.8 in which a simple step-function has been convolved and then deconvolved using the same blurring kernel. This would ideally result in the original function being recreated but it is apparent that the information lost during convolution cannot be restored using deconvolution. A mask could be applied to restrict the deconvolution to regions of high gradient, thus sharpening the image while leaving the low-gradient regions unaffected, although from Fig 3.6b we see that a low-level impact of these changes of scatter conditions at penumbra can be felt beyond the range at which artefacts become problematic.

However, a more fundamental problem with this approach is that our assumption relies on a uniform scatter kernel across the image. As discussed above, the off-axis softening caused by the flattening filter means that this is not the case. Due to the loss of spatial information when using the Fourier Transform to convert the image into the frequency domain, it is not possible to integrate a spatially-variant scatter kernel into the deconvolution process and so any inaccuracies that result must be accepted.

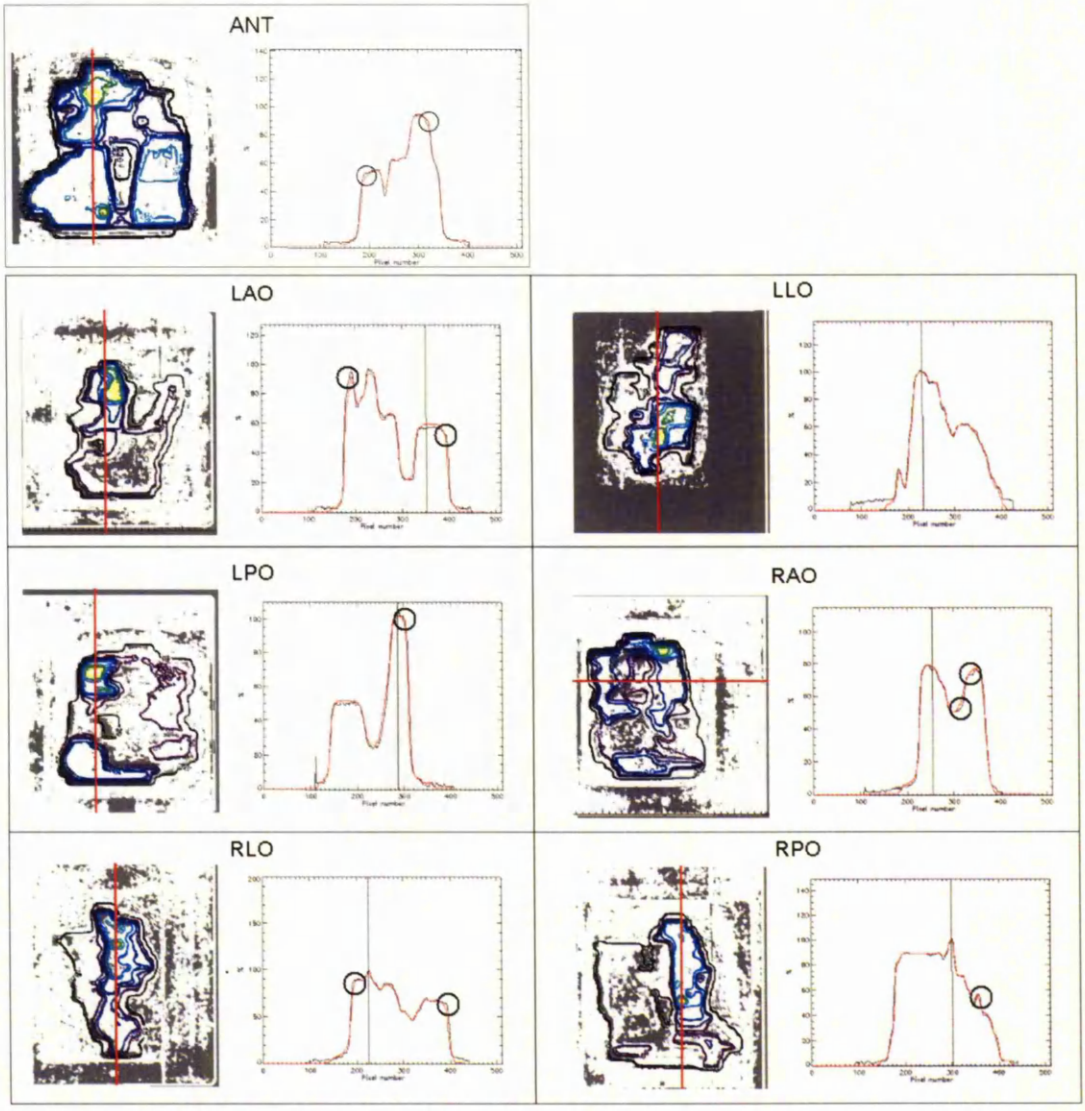


Figure 3.9. gamma analyses (2%/2mm, grey regions indicate regions exceeding tolerance) and comparison of profiles along red lines for recalculated dosemap (red) with original measured film fluence (black) based on optimised in-air fluence. These images highlight discrepancies in penumbral regions.

This inherent inaccuracy was quantified by applying the deconvolution (using an iteratively-optimised kernel) to a clinical film fluence measured at d_{max} in the solid water slab phantom. The resulting in-air fluence after the outlined corrections was then projected through the beam model established in section 3.3.2 and the dosemap calculated. The resulting comparison of the dosemap with the measured film fluence, which would ideally show identical images, is illustrated in Fig 3.9, showing good agreement in low dose regions (mostly <2% error) but with discrepancies in the highlighted penumbral regions. These disagreements are likely due to the restriction of the region of the gradient mask over which

the deconvolution acts, which cannot be extended as deconvolution artefacts will then be introduced.

Since these errors are located within the penumbra it is likely that these will be areas of particular clinical concern such as the high gradient region near the spinal cord. For this reason, it is these areas that require the highest precision for any dose verification technique and so these errors are not acceptable for clinical implementation.

3.4. Conclusions and Future work - iterative optimisation of the ODM

The current work has aimed to generate the in-air fluence array required to generate a calculated dosemap that is identical to the measured film fluence, with the intention of projecting these in-air fluences through the patient dataset to calculate the delivered dose distribution. Removal of the various corrections within the beam model has served to obtain good agreement within low-gradient regions. However, the removal of in-patient scatter, integral to the TPS, has proved more difficult to achieve. The first attempt to achieve this, by automatically identifying the gradients and using a binary intensity allocation, was insensitive to peaks and troughs within the fluence, while the second, deconvolving the measured film fluence with a uniform scatter kernel, required a compromise between the range over which the deconvolution was applied and the induced artefacts that was ultimately found to result in too great an inaccuracy.

Future work on this project will investigate the efficacy of an iterative approach to generating the required in-air fluence array. This approach will be analogous to the iterative reconstruction technique used in many medical imaging applications: the intensity values that make up the in-air fluence will be iteratively adjusted to give optimal agreement between the resulting calculated dosemap at d_{max} in the solid water slab phantom and the measured film fluence. The efficiency of the iterative loop will be increased by the use of Unix scripting

to link the processing code within IDL to the dose calculation and dosemap export within the TPS.

The changes to the beam model outlined in section 3.3.2 appear to be sufficient to achieve the required accuracy in low-gradient regions and so the iterative changes will focus on the high-gradient regions, incrementally altering the intensity values of the in-air fluence where errors are located when comparing the resulting dosemap and the measured film fluence, summarised in the following process

i. Initial in-air fluence produces calculated dosemap at d_{max}
ii. In-air fluence iteratively adjusted (high gradient regions only) to improve agreement between calculated dosemap and measured film fluence
iii. In-air fluence used to calculate dose to patient
iv. Clinically relevant QA highlighting changes to dose distribution from delivery of IMRT fields

4 Risk of radiation-induced secondary malignancies

4.1 Introduction

It is known that radiotherapy can cause the induction of secondary malignancies in surviving patients. The combination of increasing number of radiotherapy patients, improved outcomes and lower morbidity of modern treatments means that secondary malignancies could become an increasing problem, particularly in view of the suggestion by some authors that modern techniques such as intensity-modulated radiotherapy (IMRT) may increase the risk of second cancer induction [8]. In large cohorts of patients treated with radiotherapy, the incidence is found to be significant [49]. It is therefore necessary to develop methodologies to enable the prediction and, where possible, control of this outcome.

Due to the long-term follow-up required to establish baseline incidence data for secondary malignancies, the data available is currently limited. However, large-scale studies have been carried out that allow a comparison between patients treated with radiotherapy and an identical group treated by some other means. Boice et al found that for a large international cohort of 182,040 cervical cancer patients treated either by surgery or radiotherapy, the relative risk of a secondary malignancy increased after radiotherapy by approximately 9%, particularly in nearby organs where high doses were often received, such as the bladder and rectum. The risk was observed to be highest among long-term survivors and women treated at an early age [35]. Travis et al carried out a matched study into breast cancer incidence in survivors of Hodgkin's disease, finding a relative risk of 3.2 for patients with breast tissue treated to >4Gy. Risk was observed to increase with dose (8x risk for 41-61Gy) but no age-dependency was observed [50]. A study by Brenner et al involved a large cohort of prostate cancer patients treated either with radiotherapy or surgery. The results showed increased secondary malignancy risk after radiotherapy: 6% overall with significantly larger risks for bladder and rectum, both likely to receive a high dose due to proximity to the target volume [51]. Other studies demonstrated supportive results: Constine et al finding significantly increased secondary malignancy incidence in children treated for Hodgkin's disease with

radiotherapy, particularly females and particularly after large delivered doses [52]; recently, an age-dependency has been demonstrated for the risk of secondary malignancy in the contralateral breast >5 years after primary breast radiotherapy treatment, finding a significantly increased risk for women <40 years old at the age of exposure [53]. A recent report of the Committee on the Biological Effects on Ionizing Radiation (BEIR) attempted to comprehensively evaluate the available cancer incidence data based primarily on Japanese atomic bomb data, but also on various in vitro, animal and medical radiation studies like those described previously [27].

The concern over this issue due to increases in long-term cancer survival is compounded by concerns that IMRT, an increasingly popular treatment option, could increase the risk relative to conventional radiotherapy techniques. This is due to the expectation by some investigators that cell-killing becomes the dominant process at higher doses, resulting in reduced cell mutation and carcinogenesis. This model [54], with it's implication of high risk at low doses and low risk at high doses, would suggest that the increased volume of healthy tissue irradiated to a low dose by IMRT would increase cancer risk relative to conventional treatments. However, the studies outlined above appear to contradict this, showing high risk in high dose, peripheral regions of the treatment (e.g. increased bladder risk after prostate radiotherapy) [51].

Sachs & Brenner have attempted to explain this discrepancy between the standard dose response model and the observed data by considering repopulation effects within normal tissue, which would counteract the cell killing and increase carcinogenesis. The resulting fit of their analytical model to the empirical data gives a promising indication that this could accurately quantify the risks once empirical data on individual organs is confidently established [55].

A further concern regarding IMRT, mentioned previously in reference to the study by Verellen et al, is that the increased MUs required by the segmented delivery lead to an increased leakage dose to the patient outside the field. At low doses this would lead to

increased risk to the untreated volume [56; 57]. Few studies have attempted to quantify this effect, although Hall & Wuu, assuming a factor of 2-3 increase in monitor units (MUs) and using a single dose measurement outside the field, concluded that IMRT results in a 0.25% increased risk due to leakage radiation [58].

Several authors have used variations on these models to quantify relative risk between techniques. Hall & Wuu used the changing dose distribution throughout the scanned volume to quantify the relative risk between 3D-CRT and IMRT for a range of dose-response curves, finding an increased risk of 0.75% [59; 58]. This study, like all other such theoretical calculations, is inherently limited by a lack of knowledge of the dose-response curve. The accuracy is further restricted by no consideration being given to individual organ sensitivities.

Kry et al measured organ equivalent doses in a phantom for a range of treatment techniques and applied NCRP risk coefficients for individual organs to calculate a whole-body equivalent dose [60]. They found an approximate doubling of the risk from IMRT, although the methodology used is flawed, as indicated by Schneider: no account is taken of dose inhomogeneities within tissues, such as sharp gradients found in tissue neighbouring the target (e.g. bladder dose from a prostate treatment), and the NCRP coefficients were established for use in radiation protection and are inappropriate for use with individual patients [61]. An earlier study used a similar whole-body-dose method, this time incorporating leakage dose by measuring body doses with TLDs and considering the MUs required for IMRT delivered with the NOMOS system relative to conventional treatments (given as 3630MU/585MU = 6.2) [62]. The study found an 8x greater risk for IMRT, although several values used, particularly the delivered MU, seem very high based on local experience.

Schneider et al have proposed an alternative method that accounts for dose gradients across organs by defining the Organ Equivalent Doses (OED) to be the uniform dose to a particular organ that would result in the same risk as that caused by the actual dose distribution [63]. This method is again limited by a lack of confidence in the carcinogenic

dose response for individual organs, with the authors applying data from Hodgkin's disease to all patients [64]. Using this method, IMRT is found to produce broadly equivalent risks to conventional treatments. In a subsequent publication the authors demonstrate that for both adult patients and children the risk of secondary malignancies due to modern radiotherapy techniques such as IMRT and proton therapy is not significantly increased, stressing the importance of including both age-dependent sensitivity factors and the entire dose distribution in the calculation [65]. A promising study by Dasu et al again attempted to analytically define the dose-response curves and clearly demonstrated that linear extrapolation of the secondary malignancy risk from that observed at low doses is incompatible with clinically-observed result, effectively confirming the reductive effect of cell killing at higher doses [66; 67]. The study was less convincing at higher doses where agreement with observed induced malignancy rates in the bladder and rectum was not found, possibly highlighting the current limitations of such an analytical approach.

The epidemiological data suggests that the risk of inducing a secondary malignancy should be carefully considered by the clinician when evaluating treatment options, particularly regarding patients expected to survive 10 years or more or those being treated at a relatively young age, where the risk is increased by both longevity and increased radiosensitivity. The benefits of using IMRT to reduce long-term toxicities in paediatric patients are apparent and there is much current work around this issue. However, there is inconsistency in the widespread implementation due to the lack of confidence currently available in cancer risk calculations: some centres are introducing paediatric IMRT treatments apparently without explicit consideration of the potential for increased risk [68; 69], while others are conservatively opting not to exploit the tissue-sparing benefits of IMRT in these patients [70-73]. Until confidence exists in our ability to predict the associated risks of secondary malignancies for specific patients, treatment options will be either to put at risk the long-term health of a patient or to potentially forego the advantages of IMRT to patients that would benefit.

This work describes a proposed methodology for calculating the risk of secondary malignancies from radiotherapy. It aims to improve on current methods by incorporating patient-specific information on age, gender and relative weight-adjusted organ sensitivites into a risk per unit mass approach to the calculation. The approach used here also permits visualisation of the distribution of risk, which is a necessary step in incorporating these considerations into the treatment process.

4.2 Methodology

4.2.1 Calculation of risk

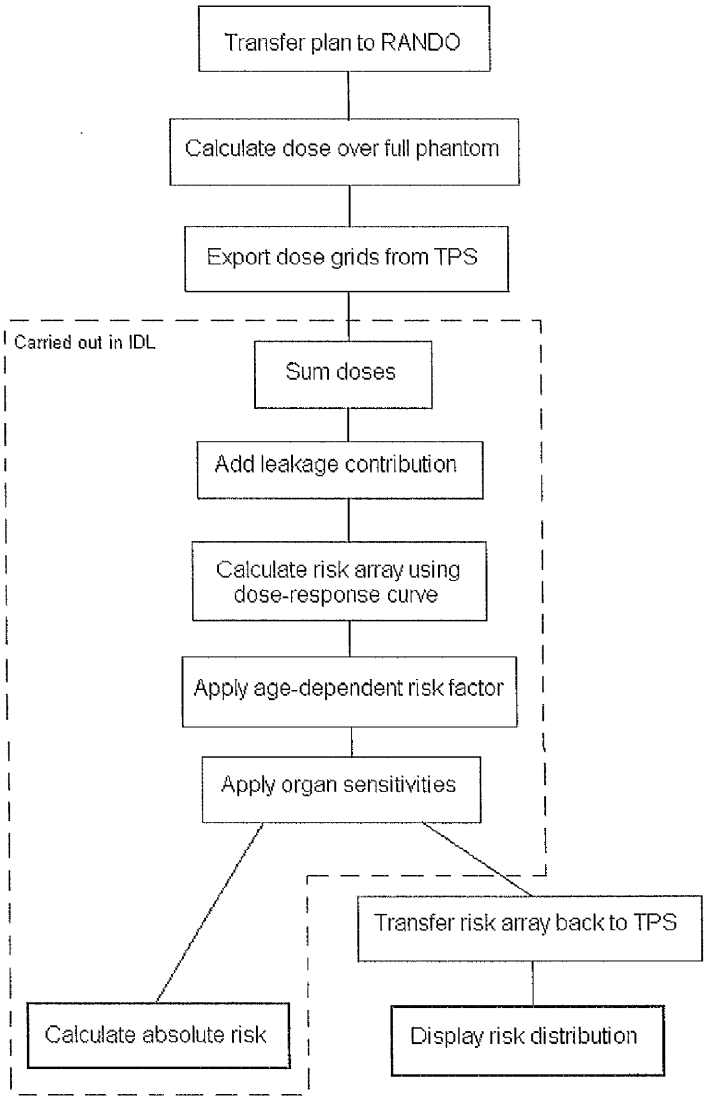


Figure 4.1. Flowchart showing function of risk calculation code

Brenner, in a recent paper criticising the use of the effective dose concept as a tool to calculate risk of cancer incidence, stated the need for a new 'effective risk' comparison measure to be determined, which should provide an easily interpretable indication of the risk of inducing a secondary malignancy in a specific patient [74]. The current work aims to achieve this by calculating absolute risk to individual dose points (r_i) within the patient and

then integrating over the full patient dataset (containing N dose points) to give the combined absolute risk (R).

$$R = \sum_{i=1}^{i=N} \frac{r_i(D_i)s_i}{N}$$

where the risk at a particular dose point (r_i) is a function of dose (D_i). Each dose point situated within a radiosensitive organ will then be multiplied by an appropriate sensitivity factor (s_i). Code has been developed in IDL (ITT Solutions, Boulder, CA) to carry out this work, the process is outlined in Figure 4.1.

4.2.2 Patient and plan simulation

Planning and visualisation aspects of the work were carried out using the Pinnacle Treatment Planning System (TPS), v8.0m (Philips, Madison, WI). A CT scan of the RANDO phantom (originally produced by Alderson Research Laboratories, Stamford, CT) acquired at 5mm slice thickness and modified in IDL to conform to adult dimensions specified in ICRP89 was used to simulate the typical patient [75]. Simplistic legs and arms were also added before the dataset was imported back into the TPS. Adult female and paediatric phantoms were similarly generated.

Relevant organs were incorporated into the dataset either by (i) outlining them directly onto the phantom (lung), (ii) copying outlines of normal clinical organs onto the phantom and manually adjusting to fit the appropriate anatomy (stomach, colon, liver, bladder) or (iii) by taking 'average' volumes from ICRP89, assuming realistic geometric shapes and then appropriately positioning them within the phantom (bladder, thyroid, prostate), which is a process similar to that proposed by the ICRP in a recent draft publication [76]. A typical target volume for extensive oropharyngeal disease and the relevant critical structures were outlined on RANDO for the head and neck and used to create plans for comparison; the IMRT plan used 7 step-and-shoot 6MV beams (57 segments in total); a conventional plan

was created using a 2-phase lateral parallel-pair superior to the hyoid bone and a matched anterior-posterior parallel pair with spinal shielding on the posterior. The dose distribution was calculated over the full dataset at 4mm resolution.

4.2.3 Calculation of dose

The dose at a given point in the patient (D_i) is a combination of the scatter dose resulting from the interaction of the primary radiation beam within the patient, the scatter from the treatment head and the leakage produced in the collimator. i.e.

$$Dose, D_i = Scatter\_dose, D_S + Leakage\_dose, D_L$$

where D_S is given by the Pinnacle TPS. D_L can be calculated from leakage@1m established at acceptance testing (measured locally at 0.03%, leading to a value of ~0.0195Gy for the 65Gy prescription routinely used for head and neck IMRT treatments). The dose due to leakage at any point i in the patient can then be given by the formula

$$D_{L,i} = Leakage@1m \times \left(\frac{100}{d_i}\right)^2 \times \exp(-\mu x_i) \times MU_{factor}$$

where d_i is the distance from the radiation source to dose point i, x_i is the thickness of patient through which the beam travels before reaching i, μ is the mass attenuation coefficient (taken as 0.0582cm^{-1}) and MU_{factor} is the proportional increase in delivered MUs relative to a standard treatment, again taken as a conventional 2-phase head and neck plan. After averaging over 8 patients treated with the local IMRT technique a MU_{factor} of 1.6 was established after taking account of differences in prescribed dose and fractionation.

4.2.4 Dose-response curve

The risk of inducing secondary malignancies is known with reasonable precision at low doses (<2Gy) from the results of the atomic bomb survivor studies [77]. However, as discussed above, beyond this limit there is little confidence in the shape of the dose-response curve. Acknowledging this uncertainty, this study uses the approach proposed by Hall & Wuu [58], whereby 3 dose-response models (shown in Figure 4.2 and selected for this study to broadly represent likely scenarios) are evaluated.

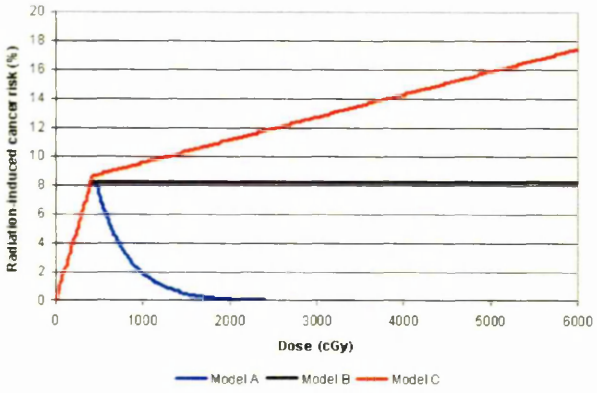


Figure 4.2. Dose-response curves used in this study: Model A assumes that cell-killing quickly becomes the dominant process; Model B assumes a higher impact from cell recovery and repopulation effects, resulting in a plateau of risk; Model C assumes that cell-killing is a less important process and assumes increasing risk with dose, albeit at a lower rate than the <4Gy region. A rounded DDREF factor of 2 is used

4.2.5 Age-specific, gender-specific and organ-specific risk sensitivities

The dose-response curves described above are intended to represent an average population. In reality, the evidence suggests significant variations in the risk associated with particular sub-sets of the population.

The BEIR VII data (Table 12D-1) gives the estimated lifetime risks for patients exposed to a single dose of 0.1Gy. The data is given for males and females separately, and is categorized by age at exposure. Overall risks to the entire body are given, as well as the risks to specific

organs [27]. This data has been used here to derive factors to scale the dose-response curve appropriately depending on the age and gender of the patient in question, as well as for the relative sensitivities of the organs in which risk is being calculated. These factors are used here at discrete ages of 5, 25, 50 and 70 years.

The age-dependent factor (A) was determined by calculating the gender-averaged risk of inducing a malignancy in 'other' tissue and normalising to the risk at the median age of exposure. The values for 'other' tissue were chosen as this would remove the effect of variation in relative organ sensitivity with age and not allow distortion due to the overwhelming contribution from individual organs such as the breast or prostate, which obviously do not give rise to most of the risk for particular sub-populations or treatment sites. Gender-specific factors (G_m and G_f) are derived for each age using the equation

$$G_m = \frac{(2 \times Male\_'other'\_risk)}{(Male\_'other'\_risk + Female\_'other'\_risk)}$$

and similarly for G_f.

When calculating relative organ sensitivities, the risk data has been converted to Risk/kg using standard mass volumes taken from ICRP 89, normalised to the overall risk to the patient. The appropriate organs and masses are given in Table 4.1.

Table 4.1. Calculation methodology for relative organ sensitivities from incidence data given in BEIR VII for a 5 year old male

Tissue	Mass (kg)	Absolute risk (from BEIR VII)	Risk/kg	Relative sensitivity
Body	76.00	508	6.68	1.00
Other' organs	72.42	140	1.93	0.29
Stomach	0.15	25	166.67	24.93
Colon	0.37	113	305.41	45.69
Liver	1.80	19	10.56	1.58
Lung	1.17	101	86.32	12.91
Prostate	0.02	33	1650.00	246.85
Bladder	0.05	76	1520.00	227.40
Thyroid	0.02	1	50.00	7.48

4.3 Results

4.3.1 Patient-specific sensitivity parameters

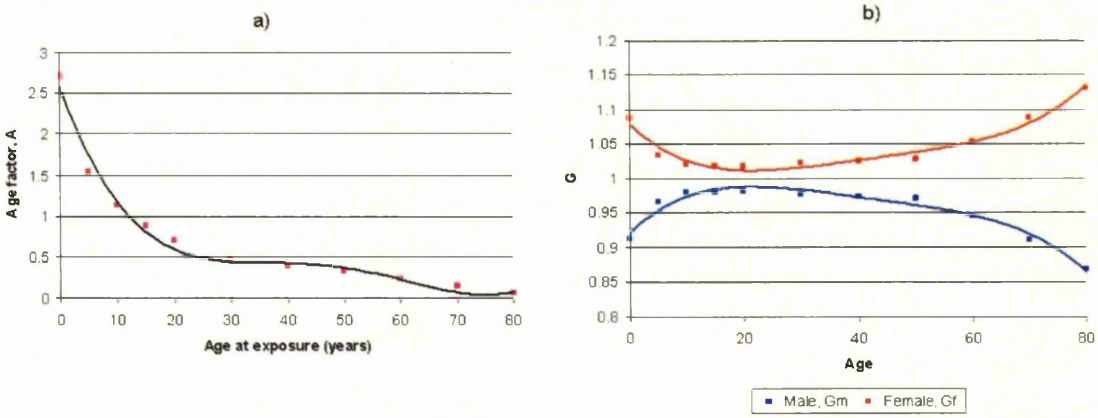


Figure 4.3. Variation of a) age factor A and b) gender factor G as a function of age. Data taken from BEIR VII

The age sensitivity factor (A) is shown in figure 4.3a), normalised to the age of median risk (13.8 years).

The gender-sensitivity factors are shown in Figure 4.3b), which demonstrate elevated female sensitivity as a function of age. A large variation in relative organ sensitivities (s_i) has been seen between organs, with little consistency observed between males and females. The variation in s_i as a function of age is shown in Figure 4.4.

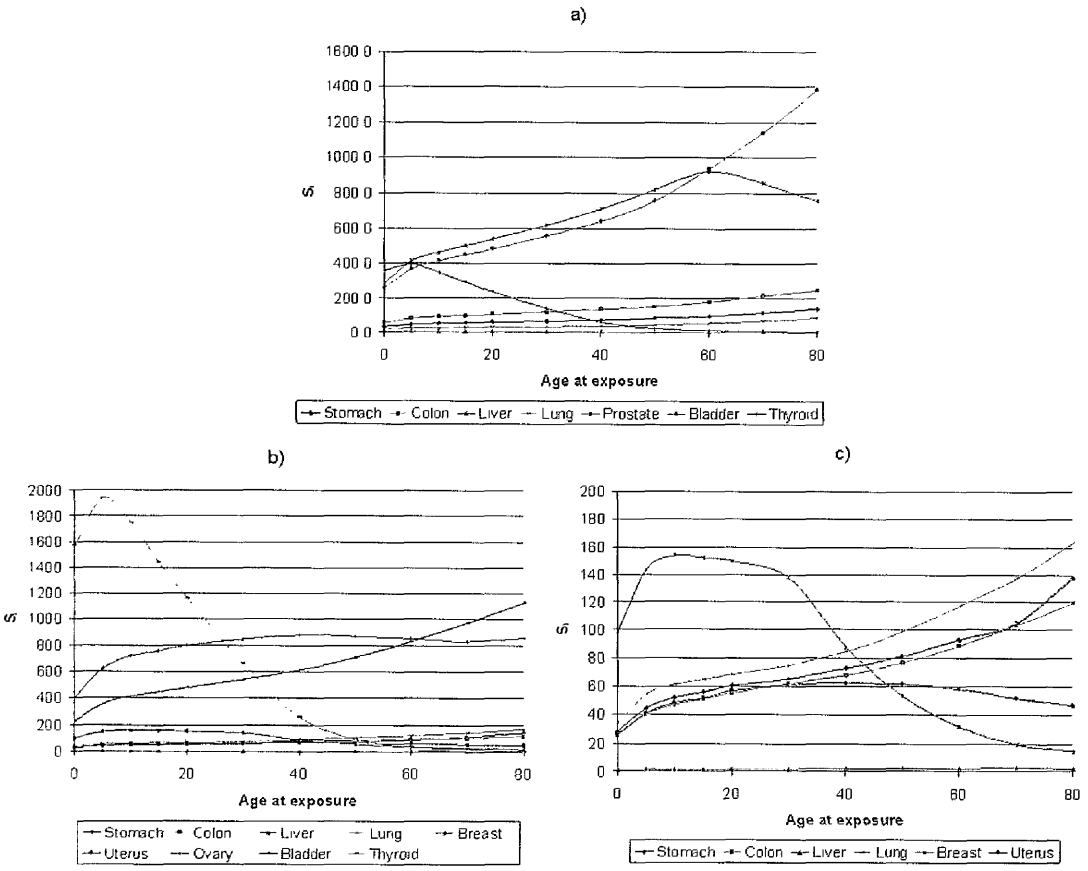


Figure 4.4. Relative organ sensitivity variation with age for a) male and b&c) females (Figure 4.4c is at a different scale to better visualise the trends for less sensitive tissues). Data taken from BEIR VII and modified as illustrated in Table 4.1

4.3.2 Risk calculations

All results given here are using either the male or paediatric phantoms and refer to a comparison of IMRT and 3D-conformal techniques for the head and neck.

The results shown in figure 4.5, which demonstrate the expected decline in risk with age at exposure for both treatment modalities, show no markedly increased risk associated with the IMRT technique compared with 3D-conformal for adult patients. A slight increased risk is predicted for paediatric patients (0.1-0.5%, depending on the dose-response model used). The risk can be significantly altered by inclusion of organ radiosensitivity parameters,

although the magnitude and direction of changes is dependent upon the dose-response model used in the calculation (see fig.4.5d-f).

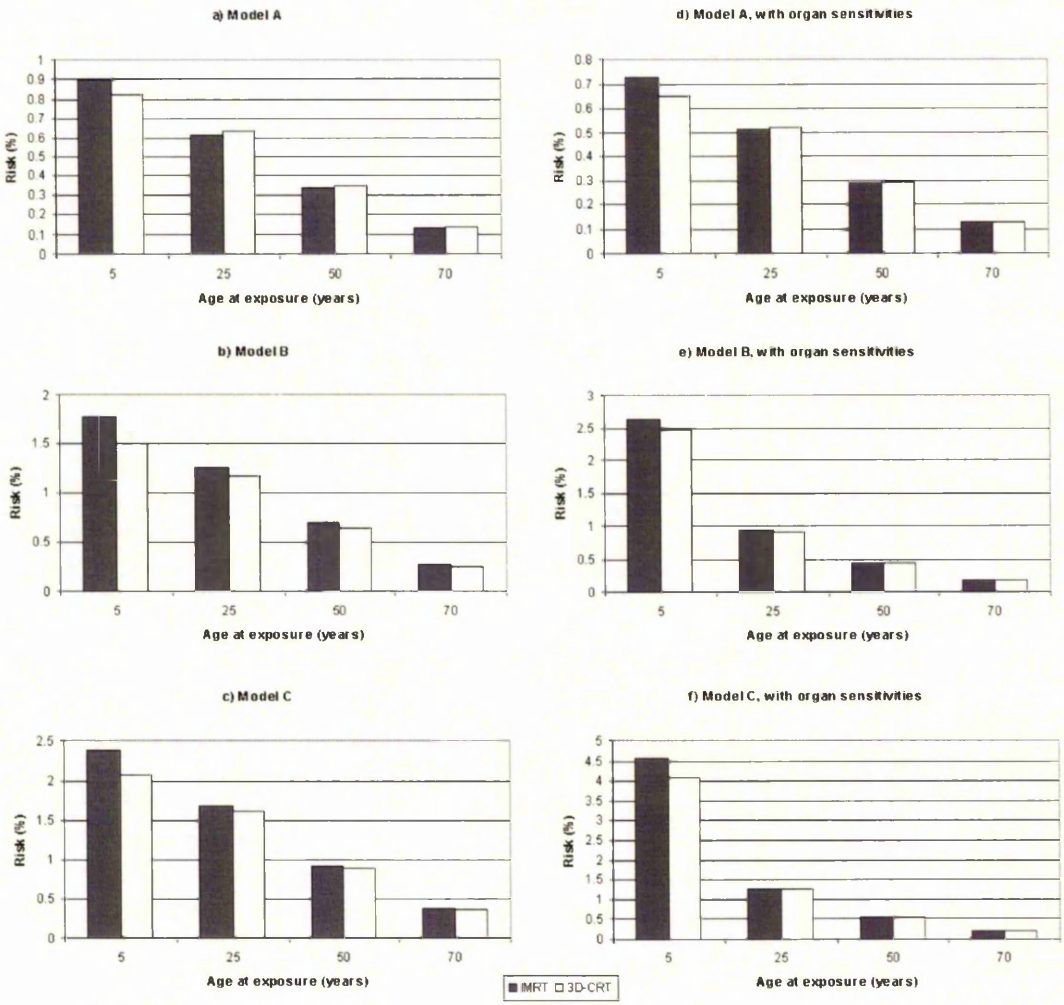


Figure 4.5. Absolute risk values for IMRT and 3D-conformal treatment techniques for the different dose-response curves

The risk distribution for a 5 year old male using the IMRT technique is illustrated for the different dose-response models in Figure 4.6. Beyond the broad similarity between treatment techniques, this image suggests that the inclusion of organ sensitivities is important in accurately determining risk, with the lungs, which receive mainly low and moderate doses, exhibiting high concentrations of risk. It can also be seen that highly sensitive structures such as the prostate contribute to the risk, although the small volume and distance from the high dose region means that it is found to have minimal impact on the overall risk.

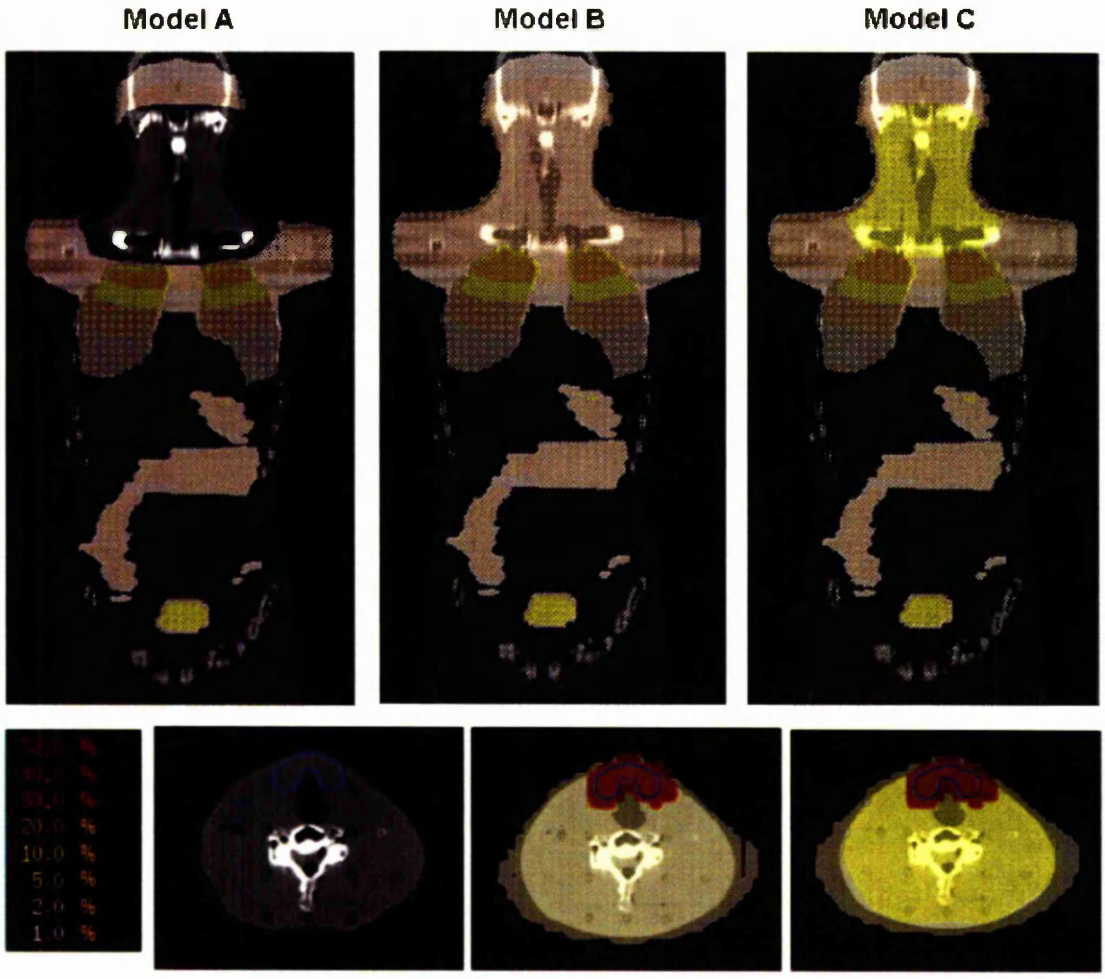


Figure 4.6. Risk distributions from head and neck IMRT treatment for different models (risk at any given point corresponds to the absolute cumulative risk if this were present throughout the entire dataset). 3D conformal treatment showed a similar distribution

The primary source of risk is found to vary depending upon the dose-response model used: when using Model A, in which the risk is independent of the high dose volume, the dominant source of risk (35% of the total) is found to be due to the sensitivity of the lung; when using Model B the sensitivity to the moderate and high dose range translates to a sensitivity from both the lung (14%) and the thyroid (55%); the increased response to high doses using Model C leads to a dominance of risk from the thyroid (68%) and less impact from the lung (8%).

The component of the risk due to leakage dose was found to be <0.1% in all situations, demonstrating that any associated errors in the calculation would be unlikely to have a significant impact. The discrepancy in risk between treatment techniques in this regard is entirely due the measured MU_{factor}.

4.4 Discussion

The work presented here demonstrates the efficacy of an alternative approach to calculating the risk of secondary malignancies following radiotherapy. This approach models the patient as an array of dose points that can each be assigned a risk of malignancy. When combined, this produces the cumulative risk to the patient. This methodology improves on that of previous approaches in several ways: population-based sensitivities; relative organ sensitivities derived directly from epidemiological data; flexible and easy to modify in light of newly-acquired data or patient-specific information; whole-body dose evaluated; analytical calculation of leakage dose. There is currently insufficient established data to reliably deviate from assumptions of uniformity in issues such as intra-organ radiosensitivity, dose-response curves for particular organs or underlying genetic conditions for individual patients and so broad assumptions of uniformity have been assumed for the current work. The voxel-based approach used here will allow easier future incorporation of factors such as non-uniform risk within an organ than current organ-based approaches.

The scatter dose calculated outside the treatment field by Pinnacle is of limited accuracy, which could impact on the overall risk calculation. This was investigated and revealed that a factor of 2 error in the dose measured >5cm beyond the treatment field resulted in <0.5% difference in the overall risk (largely within 10cm of the high dose region), for both IMRT and 3D-CRT techniques. This was intended to represent a worst-case scenario and, whilst the potential error is not insignificant, the low magnitude and inter-technique uniformity of this additional risk would not alter the conclusions of this study.

This current methodology has been used to compare treatment techniques for the head and neck with regard to the risks associated with treatment for patients of different ages, but focuses primarily on paediatric patients. The calculated risk of secondary malignancy after treatment of paediatric patients varies depending on the model used, ranging from approximately 0.7% for Model A to 2.5% for Model B and 4.5% for Model C (referring to Figure 4.5d-f), while values for adult patients are significantly lower. Relevant extensive studies have not been carried out involving head and neck patients, until which time it is difficult to validate modelling accuracy or confidently establish the correct dose-response. It is reasonable to expect that any observed discrepancy in risk between 3D-conformal and IMRT techniques for the head and neck could be less than that predicted in the prostate: the proportionally lower volume of tissue conventionally irradiated to very low dose is less for the standard parallel-pair technique for the head and neck than the standard 3 or 4-field techniques used to treat the prostate. This means that the comparative volume receiving a significant dose with IMRT, in which the dose is spread throughout the entire axial slice, would be expected to be greater for the prostate, leading to a greater increase in risk of an induced malignancy.

Long-term, extensive epidemiological studies could indicate that the assumption of inter-organ uniformity of dose response curves, and indeed that relative organ sensitivity is independent of dose, is invalid. This possibility is supported by data suggesting different site-dependent extrapolations of the atomic bomb data beyond 4Gy, with different curves observed for stomach and bladder, for example [78]. More data on post-RT cancer induction at specific sites could inform organ-specific dose-response curves in the future. The possibility of including fractionation effects (as discussed by Sachs & Brenner and Jones) into the risk calculation would also be beneficial once radiobiological constraints are known with sufficient precision [55].

The risk due to leakage, which has been calculated to include the increased MUs delivered by IMRT as well as corrections for inverse-square law and attenuation throughout the patient, is found to be <0.1%. The calculated leakage risk is greater for IMRT compared to

3D-CRT by a factor equal to the relative number of delivered MUs and, since leakage dose is significantly lower than the 4Gy dose level at which the dose-response models diverge, is independent of the model used. Our locally-established MU_{factor} for IMRT appears to be lower than similar values quoted in the literature, possibly due to use of step-and-shoot rather than dynamic delivery and our practice of limiting the smallest segments to $9cm^2$. The use of many smaller segments or dynamic delivery could significantly increase the required MUs and therefore the risk due to leakage.

Previous work has assumed that increases in leakage dose and volume of low dose irradiation encountered using IMRT would result in increased risk. However, the results shown here in Figure 4.5 demonstrate that the effect is very small (<0.5%) for paediatric patients and negligible for adults.

Interest has recently been shown in strategies to reduce the risk of secondary malignancies prior to treatment [79]. To demonstrate the potential of photon treatments to reduce risk, the paediatric IMRT plan was reoptimised to reduce dose to lung and thyroid until the target coverage was significantly compromised. Absolute risk reduction of only 0.1% was achieved, although the efficacy of this approach is highly dependent on both the clinical site in question and the dose-response curve used for the tissues involved (e.g. a steep dose-response curve for thyroid could allow significantly reduced risk to the patient). This suggests that the peripheral dose due to the inherent scatter from photon treatments may be difficult to reduce and that alternative modalities such as protons may produce more benefits [80]. It also suggests that for the head and neck the potential improvements to conformality available with arc-modulated techniques, Tomotherapy or dynamic-IMRT may be of limited benefit for this purpose.

4.5 Conclusion

A novel approach to the calculation and display of the risk of secondary malignancies due to radiotherapy has been described. Dose calculations on adapted phantom images, together with organ sensitivities calculated from the BEIRVII data, were used to calculate the risk for each unit mass of tissue and the cumulative absolute risk. This approach allows the flexibility to incorporate sensitivity factors for both the individual organ and patient and to visualise the risk distribution.

IMRT results in a comparable risk of secondary malignancy to the patient as 3D-CRT, while leakage dose through the treatment head has only a small overall impact. Our ability to conform dose away from sensitive structures to reduce this risk appears limited, although much uncertainty exists over the correct dose-response curve, without which such considerations cannot be confidently assessed.

Further work is needed to confidently establish the dose-response curves for individual organs and patient populations. Due to the scarcity of relevant data for the head and neck, the work described in this chapter will be repeated for prostate radiotherapy, benchmarked against the data published by Brenner et al [51] to determine the likely dose-response curve for the affected tissue such as bladder, rectum and lung.

5 Summary and discussion

The treatment of head and neck malignancy with radiotherapy has the potential to benefit greatly from the adoption of modern complex treatment techniques such as IMRT, both in terms of improved tumour control and reduced rates of toxicity. The clinical application of the technique is currently expanding, with the desire for implementation currently exceeding the required facilities and expertise within the UK, suggesting that it will ultimately be in widespread clinical use as these hurdles are overcome at individual institutions.

The optimal IMRT treatment process has yet to be established, with much uncertainty existing at each stage. This work has aimed to strengthen each aspect of the process and address areas of concern regarding the implementation of IMRT.

Delineation

In Publication #1, the design and implementation of an imaging protocol to allow the incorporation of high-quality MR images acquired in the treatment position into the target delineation process was described and the associated precision of the technique quantified. The simple, practical methodology used superficial RF coils placed close to the anterior surface of the patient using a simple Perspex frame. The imaging protocol, which is tolerated well by patients, has been adopted into clinical practice and is expected to be of significant benefit in the staging and delineation of nasopharyngeal tumours, in which intra-cranial spread is uncertain, and in the delineation of oropharyngeal tumours that are obscured by dental artefacts on CT scans. The Syntegra image fusion software within the TPS uses the normalised mutual information algorithm, which is found to be adequate for the purposes of image fusion. However, it is advisable to restrict the volume over which the fusion is evaluated to a volume surrounding the target region, avoiding shoulder and other surface contours from which small distortions in patient anatomy will have a significant detriment on the fusion accuracy.

Accuracy and precision of dose calculation and delivery

Publication #2 confirms the suitability of the collapsed cone convolution algorithm for use in head and neck IMRT, demonstrating good agreement with measured fluence films within a clinically-acceptable 4%/4mm gamma tolerance. The importance of precise modelling of the rounded leaf ends was highlighted as being critical to the accuracy of the dose calculation: the proximity of multiple beam penumbrae to critical structures such as the spinal cord can mean that slight errors in leaf modelling can mask significantly increased doses.

The impact of dental artefacts was demonstrated in Publication #3 to result in statistically, although perhaps not clinically, significant errors when part of the standard clinical practice that saturates all CT densities at a low level of approximately 1.8g/cm^3. In reality, the localised impact of these errors is likely to be exaggerated by our assumptions of, for example, no random errors in patient setup, treatment delivery and patient motion, all of which would act to reduce the observed errors. However, the study showed that for all but the most extreme cases, these inaccuracies could be reduced by extending the Hounsfield Unit (HU) conversion to account for the higher densities in the initial calculation, reducing the observed errors to acceptable levels. In the extreme cases, identified as patients in whom the target region extends into a severely affected region, a simple bulk density correction of the teeth and soft tissue is sufficient to improve the accuracy. In all cases therefore, dose calculation accuracy need not be significantly compromised by the presence of dental artefacts.

The design and implementation of an audit phantom for use in both routine clinical verification and technique audit has been described in Publication #5. Additional to describing further refinements of the phantom, Publication #6 has confirmed that the achievable dosimetric accuracy of the IGRT process is within 3%, including within high-gradient regions. Good delivery precision was observed, with repeatability found to be within the experimental noise of the detector of 1.5%.

This work focussed entirely on technical aspects of treatment delivery and assumed a rigid, unchanging patient. The reality of inter and intra-fraction movement of the patient will create additional uncertainty. However, the work in Publication #6 confirms that even over the course of several months, over which time the performance of the linear accelerator may be expected to drift, high precision and accuracy in the delivery of complex head and neck IMRT plans is achievable.

Subsequent to Publications 5&6, the HANK phantom has been routinely used for patient-specific dose verification to establish confidence in the local treatment technique. It is also in routine use as a commissioning tool for new linear accelerators to confirm the accuracy of delivery and calculation for new linear accelerators and planning system software upgrades respectively. Work is currently underway within the Christie Hospital to develop an improved phantom that combines increased flexibility in multiple measurement positions with novel detectors and a more realistic patient contour. This is intended for use in the setup verification and audit of advanced techniques such as IMRT, IGRT and Volumetric Intensity Modulated Arc Therapy (VMAT). The latter development has not been discussed in detail in this work but offers the potential for further improvements to radiotherapy treatments through improved dose conformality and treatment efficiency.

There is much current interest in the implementation of in-vivo dosimetry to allow timely detection of gross treatment errors in radiotherapy [81]. Although the Mosfet detectors described in Publication #6 represent a plausible option for this application, they are unlikely to represent an optimal solution: although their reasonable noise characteristics, good spatial resolution and instant readout would be advantageous, the cumbersome nature of the required electronics, their limited lifespan and the need for precise positioning for measurement in highly modulated fields make them an inferior alternative to the use of portal imagers.

The attempt to calculate the impact of residual delivery errors received by the patient by projecting measured film fluences through the patient CT dataset, described in Chapter 3,

was unfortunately unsuccessful. Whilst the beam model in the TPS could be manipulated to obtain good agreement between calculated and measured fluence in low gradient regions, accuracy in high gradient regions is felt to be insufficient for clinical use. This inaccuracy is due to attempts to remove the effects of in-patient scatter from the measured film fluence (prior to it being input as the delivered in-air fluence), which would otherwise be accounted for twice in the calculation of the delivered dose distribution as it is both contained within the measured film fluence and is inherent in the dose calculation algorithm. This removal of blurring has been attempted both by a forced sharpening of the penumbra, which would not account for peaks and troughs within the dose distribution, and by deconvolution of the measured film fluence with a uniform scatter kernel intended to represent the spread of in-phantom scatter. Future attempts to establish the required in-air fluence will most likely require an iterative approach to the problem. It is hoped that the completion of this work will allow the impact of IMRT delivery errors to be directly assessed in terms of their impact on the patient, providing the first real indication of the dosimetric tolerances required to maintain a clinically acceptable plan.

Ultimately, it is anticipated that this will lead to a more patient-specific mode of dose verification that will allow measured delivery errors to be quantified in terms of their direct impact on the dose distribution to the actual patient anatomy.

Practicalities of treatment plan optimisation

The results of Publication #4 demonstrate the flexibility of the IMRT planning process, clearly showing the extensive level of laryngeal sparing that can be achieved. This work, alongside that of Publication #3, stands as an illustrative example of how simple solutions can often be found to practical problems encountered when implementing IMRT.

Current local practice achieves a high level of sensitive tissue sparing in clinical head and neck IMRT planning, with many structures receiving significantly less dose than would be received with 3D-CRT treatments (i.e. oral cavity, SMG, parotid gland, cochlear) and others

being maintained at conventional levels (i.e. larynx, other uninvolved tissues) while no compromises are required to target coverage.

Risk of secondary malignancies

Results of the work outlined in Chapter 4 suggest only a limited increase in the risk of secondary malignancies from IMRT compared to 3D-CRT for the head and neck. Investigations also suggest that negligible reduction in risk would result from the inclusion of extra shielding within the treatment head to reduce leakage dose, and that attempting to spare sensitive structures as part of the planning process is unlikely to lead to significant benefits for head and neck treatments.

Much work is required in this area to gather data to more accurately inform the models currently used. In particular, the dose-response curve of different organs and their relative sensitivity as a function of dose, are not known with good precision.

6 Conclusion

It is essential to the implementation of IMRT that due caution is exercised at every step of the development and treatment process. However, many of the problems that arise when implementing this complex technique are not insurmountable and efforts should be made to identify simple, practical solutions where possible. In this way, compromises to patient treatment may be avoided, as illustrated in Publication #3 in reference to the detrimental impact of dental artefacts and, particularly, in Publication #4 in response to observed increases in dysphagia. Other, more long-term risks, such as the potential for increased incidence of secondary malignancies after IMRT, are currently difficult to quantify. However, despite earlier studies predicting this increase, more recent, and arguably more scientifically rigorous, work has reinforced the contrasting conclusions drawn in chapter 4: that IMRT is not associated with a significantly increased risk of secondary malignancies. In light of this it is perhaps expedient to continue to exploit the potential benefits of IMRT where radiotherapy is an essential part of a patient's treatment, while obviously taking particular care to consider other options in high-risk patient sub-populations. In the meantime, data should continue to be gathered to more accurately quantify the modelling parameters required to accurately predict this effect.

To conclude, early indications suggest that patients are likely to benefit greatly from the introduction of IMRT for the treatment of head and neck cancer due to the associated reductions in toxicity. Confidence in the safe delivery of IMRT, will encourage the exploitation of the expected gains in local control expected from dose escalation, possibly in combination with chemotherapeutic agents. Widespread implementation of IMRT should proceed cautiously, whilst striving to achieve the full benefits of this technique.

The advances made in head and neck radiotherapy by the development of IMRT complement continued advances in other related fields such as chemotherapeutic agents (e.g. cisplatin), which are in widespread use as radiosensitising agents that may improve local control and therefore overall survival. More recently the development of targeted

therapeutic agents such as Cetuximab, which binds to the tumour epidermal growth factor receptors to reduce cell division, offer significant promise for many patients in terms of equivalent or greater tumour radiosensitisation with less toxicity than conventional chemotherapy. The next few years are likely to see the continued development of novel therapeutic agents in combinations with radiation to improve overall survival rates.

References

1. Wylie J, Sen M, Swindell R, et al. Definitive radiotherapy for 114 cases of T3N0 glottic carcinoma: influence of dose-volume parameters on outcome. *Radiother Oncol* 1999; **53:** 15-21
2. Daly M, Lieskovsky Y, Pawlicki T, et al. Evaluation of patterns of failure and subjective salivary function in patients treated with intensity modulated radiotherapy for head and neck squamous cell carcinoma. *Head & Neck* 2007; **March:** 211-20
3. Sanguineti G, Adapala P, Endres E, et al. Dosimetric predictors of laryngeal oedema. *Int J Rad Onc Biol Phys* 2007; **68:** 741-9
4. Schoenfeld G, Amdur R, Morris C, et al. Patterns of failure ad toxicity after intensity-modulated radiotherapy for head and neck cancer. *Int J Rad Onc Biol Phys* 2008; **71:** 377-85
5. Soto D, Kessler M, Piert M, et al. Correlation between pretreatment FDG-PET biological target volume and anatmoical location of failure after radiation therapy for head and neck cancers. *Radiother Oncol* 2008; **89:** 13-8
6. Nutting C, A'Hern R, Rogers M, et al. First results of a phase III multicenter randomized controlled trial of intensity-modulated (IMRT) versus conventional radiotherapy (RT) in head and neck cancer. *J Clin Oncol* 2009; **27:** abstr LBA6006
7. Amdur R, Liu C, Li J, et al. Matching intensity-modulated radiation therapy to an anterior low-neck field. *Int J Rad Onc Biol Phys* 2007; **69:** S46-S8
8. Hall E and Wuu C. Radiation-induced second cancers: the impact of 3D-CRT and IMRT. *Int J Rad Onc Biol Phys* 2003; **56:** 83-8
9. Jefferies S, Taylor A and Reznek R. Results if a national survey of radiotherapy planning and delivery in the UK in 2007. *Clin Oncol* 2009; **21:** 204-17
10. Price P. Molecular imaging for translational research in radiation oncology. *Radiother Oncol* 2006; **78:** S22
11. Wong R. Current staus of FDG-PET for head and neck cancer. *J Surg Oncol* 2008; **97:** 649-52

12. Vanderstraeten B, Duthoy W, De Gersem W, et al. 18F-fluoro-deoxy-glucose positron emission tomography (18F-FDG-PET) voxel-intensity-based intensity-modulated radiation therapy (IMRT) for head and neck cancer. *Radiother Oncol* 2006; **79:** 249-58

13. Madani I, Duthoy W, Derie C, et al. 2006 Positron Emission Tomography-Guided, Focal-Dose Escalation Using Intensity-Modulated Radiotherapy for Head and Neck Cancer. In: *European Society for Therapeutic Radiology and Oncology (ESTRO), Leipzig, Germany, October 8–12,*

14. Daisne J, Duprez T, Weynand B, et al. Tumour volume in pharyngolaryngeal squamous cell carcinoma: comparison at CT, MR imaging, and FDG PET and validation with surgical specimen. *Radiology* 2004; **233:** 93-100

15. MacManus M, Nestle U, Rosenzweig K, et al. Use of PET and PET/CT for radiation therapy planning: IAEA expert report 2006-2007. *Radiother Oncol* 2009; **91:** 85-94

16. Emonts P, Bourgeois P, Lemort M, et al. Functional imaging of head and neck cancers. *Curr Opin Oncol* 2009; **21:** 212-7

17. Van Herk M. Errors and margins in radiotherapy. *Semin Radiat Oncol* 2004; **14:** 52-64

18. Van Herk M, Remeijer P and Lebesque J. Inclusion of geometric uncertainties in treatment plan evaluation. *Int J Rad Onc Biol Phys* 2002; **52:** 1407-22

19. Ahnesjo A. Collapsed cone convolution of radiant energy for photon dose calcualtion in heterogeneous media. *Med Phys* 1989; **16:** 577-92

20. Kim Y, Tome W, Bal M, et al. The impact of dental metal artifacts on head and neck IMRT dose distributions. *Radiother Oncol* 2006; **79:** 198-202

21. Borst G, Sonke J, Betgen A, et al. Kilo-voltage cone-beam computed tomography setup measurements for lung cancer patients; first clinical results and comparison with electronic portal-imaging device. *Int J Rad Onc Biol Phys* 2007; **68:** 555-61

22. IPEM 96: Guidance for the clinical implementation of intensity modulated radiation therapy; 2008. Institute of Physics and Engineering in Medicine (IPEM)

23. McDermott L, Wendling M, Nijkamp J, et al. 3D in vivo dose verification of entire hypo-fractionated IMRT treatments using an EPID and cone-beam CT. *Radiother Oncol* 2008; **86:** 35-42

24. PARSPORT Protocol. A multi-centre randomised study of parotid sparing intensity modulated radiotherapy versus conventional radiotherapy in patients with head and neck cancer. *ISRCTN: 48243537* 2003;

25. Eisbruch A, Foote R, O'Sullivan B, et al. Intensity-modulated radiation therapy for head and neck cancer: emphasis on the selection and delineation ofthe targets. *Semin Radiat Oncol* 2002; **12:** 238-49

26. Eisbruch A, Kim H, Terrell J, et al. Xerostomia and its predictors following parotid-sparing irradiation of head-and-neck cancer. *Int J Rad Onc Biol Phys* 2001; **50:** 695-704

27. Chao K, Deasy J, Markman J, et al. A prospective study of salivary function sparing in patients with head and neck cancers receiving intensity-modulated or three-dimensional radiation therapy: initial results. *Int J Rad Onc Biol Phys* 2001; **49:** 907-16

28. Eisbruch A. Reducing xerostomia by IMRT: what may, and may not, be acheived. *J Clin Oncol* 2007; **25:** 4863-4

29. Jellema A, Doornaert P, Slotman B, et al. Does radiation dose to the salivary glands and oral cavity predict patient-rated xerostomia and sticky saliva in head and neck cancer patients treated with curative radiotherapy? *Radiother Oncol* 2005; **77:** 164-71

30. Eisbruch A, Ten Haken R, Kim H, et al. Dose, volume and functional relationships in parotid salivary glands following conformal and intensity-modulated irradiation of head and neck cancer. *Int J Rad Onc Biol Phys* 1999; **45:** 577-87

31. Murdoch-Kinch C, Kim H, Vineberg K, et al. Dose-effect relationships for the submandibular salivary glands and implications for their sparing by intensity modulated radiotherapy. *Int J Rad Onc Biol Phys* 2008; **72:** 373-82

32. Saibishkumar E, Jha N, Scrimger R, et al. Sparing the parotid glands and surgically transferred submandibular gland with helical tomotherapy in post-operative radiation of head and neck cancer: a planning study. *Radiother Oncol* 2007; **85:** 98-104

33. Chao K, Majhail N, Huang C, et al. Intensity-modulated radiation therapy reduces late salivary toxicity without compromising tumor control in patients with oropharyngeal

carcinoma: a comparison with conventional techniques. *Radiother Oncol* 2001; **61:** 275-80

34. Cannon D and Lee N. Recurrence in region of spared parotid gland after definitive intensity-modulated radiotherapy for head and neck cancer. *Int J Rad Onc Biol Phys* 2008; **700:** 660-5
35. Boice J, Day N, Anderson A, et al. Second cancers following radiation treatment for cervical cancer. An international collaboration among cancer registries. *J Natl Cancer Inst* 1985; **74:** 955-75
36. Brenner D, Curtis R, Hall E, et al. Second malignancies in prostate carcinoma patients after radiotherapy compared with surgery. *Cancer* 2000; **88:** 398-406
37. Xu X, Bednarz B and Paganetti H. A review of dosimetry studies on external-beam radiation treatment with respect to second cancer induction. *Phys Med Biol* 2008; **53:** R193-R241
38. Ezzell G, Galvin J, Low D, et al. Guidnace document on delivery, treatment planning, and clinical implementation of IMRT: Report of the IMRT subcommittee of the AAPM radiation therapy committee. *Med Phys* 2003; **30:** 2089-115
39. Galvin J, Ezzell G, Eisbruch A, et al. Iplementing IMRT in clinical practice: a joint document of the American Society for Therapeutic Radiology and Oncology and the American Association of Physicists in Medicine. *Int J Rad Onc Biol Phys* 2004; **58:** 1616-34
40. Molineu A, Followill D, Balter P, et al. Design and implementation of an anthropomorphic quality assurance phantom for intensity-modulated radiation therapy for the Radiation Therapy Oncology Group. *Int J Rad Onc Biol Phys* 2005; **36:** 577-83
41. Webster G, Hardy M, Rowbottom C, et al. Design and implementation of a head-and-neck phantom for system audit and verification of intensity-modulated radiation therapy. *J App Clin Med Phys* 2008; **9:** 46-56
42. Georg D, Stock M, Kroupa B, et al. Patient-specific IMRT verification using independent fluence-based dose calculation software: experimental benchmarking and initial clinical experience. *Phys Med Biol* 2007; **52:** 4981-92

43. Karlsson A, Gustavsson H, Mansson S, et al. Dose integration charcteristics in normoxic polymer gel dosimetry investigated using sequential beam irradiation. *Phys Med Biol* 2007; **52:** 4697-706

44. Sakhalkar H, Adamovics J, Ibbott G, et al. A comprehensive evaluation of the PRESAGE/optical-CT 3D dosimetry system. *Med Phys* 2009; **36:** 71-82

45. Wendling M, Louwe R, McDermott L, et al. Accurate two-dimensional IMRT verification using a back-projection EPID dosimetry method. *Med Phys* 2006; **33:** 259-73

46. Cadman P, Bassalow R, Sidhu N, et al. Dosimetric considerations for validation of a sequential IMRT process with a commercial treatment planning system. *Phys Med Biol* 2002; **47:** 3001-10

47. Starkschall G, Steadham R, Popple R, et al. Beam-commissioning methodology for a three-dimensional convolution/superposition photon dose algorithm. *J App Clin Med Phys* 2000; **1:** 8-27

48. ICRU 46: Photon, electron, proton and neutron interaction data for body tissues; 1991. International Commission on Radiation Units and Measurements (ICRU)

49. Xu X, Bednarz B and Paganetti H. A review of dosimetry studies on external-beam radiation treatment with respect to second cancer induction. *Phys Med Biol* 2008; **53:** R193-241

50. Travis L, Hill D, Dores G, et al. Breast cancer following radiotherapy and chemotherapy among young women with Hodgkin disease. *JAMA* 2003; **290:** 465-75

51. Brenner D, Curtis R, Hall E, et al. Second malignancies in prostate carcinoma patients after radiotherapy compared with surgery. *Cancer* 2000; **88:** 398-406

52. Constine L, Tarbell N, Hudson M, et al. Subsequent malignancies in children treated for Hodgkin's disease: associations with gender and radiation dose. *IJROBP* 2008; **72:** 24-33

53. Stovall M, Smith S, Langholz B, et al. Dose to the contralateral breast from radiotherapy and risk of second primary breast cancer in the WECARE study. *IJROBP* 2008; **72:** 1021-30

54. Hall E 2007 *Radiobiological modelling in radiation oncology,* ed R Dale and B Jones (London: British Institute of Radiology) pp 12-34

55. Sachs R and Brenner D. Solid tumour risks after high doses of ionizing radiation. *Proc Natl Acad Sci* 2005; **102:** 13040-5

56. Lillicrap S, Morgan H and Shakeshaft J. X-ray leakage during radiotherapy. *British Journal of Radiology* 2000; **73:** 793-4

57. Williams P and Hounsell A. X-Ray leakage considerations for IMRT. *British Journal of Radiology* 2001; **74:** 98-102

58. Hall E and Wuu C. Radiation-induced second cancers: the impact of 3D-CRT and IMRT. *IJROBP* 2003; **56:** 83-8

59. Hall E. Henry S. Kaplan distinguished scientist award. *Int J Rad Biol* 2004; **80:** 327-37

60. Kry S, Salehpour M, Followill D, et al. The calculated risk of fatal secondary malignancies from intensity-modulated radiotherapy. *IJROBP* 2005; **62:** 1195-203

61. Schneider U. Calculated risk of secondary malignancies from intensity-modulated radiotherapy: In regard to Kry et al. *IJROBP* 2006; **64:** 1290

62. Verellen D and Van Havere F. Risk assessment of radiation-induced malignancies based on whole-body equivalent dose estimates for IMRT treatment in the head and neck region. *Radiotherapy & Oncology* 1999; **53:** 199-203

63. Schneider U, Zwahlen D, Ross D, et al. Estimation of radiation-induced cancer from three-dimensional dose distributions: concept of organ equivalent dose. *IJROBP* 2005; **61:** 1510-5

64. Kry S. Calculated risk of fatal secondary malignancies from intensity-modulated radiotherapy:In response to Dr Schneider. *IJROBP* 2006; **64:** 1290

65. Schneider U, Lomax A and Timmerman B. Second cancers in children treated with modern radiotherapy techniques. *Radiotherapy & Oncology* 2008; **89:** 135-40

66. Dasu A, Toma-Dasu I, Franzen L, et al. 2009 The risk fo secondary cancers in patients treated for prostate carcinoma - an analysis with the competition dose-response model. In: *IFMBE Proceedings 25/III,*

67. Dasu A, Toma-Dasu I, Olofsson J, et al. The use of risk estimation models for the induction of secondary cancers following radiotherapy. *Acta Oncologica* 2005; **44:** 339-47

68. Raggi E, Mosleh-Shiraz M and Saran F. An evaluation of conformal and intensity-modulated radiotherapy in whole ventricular radiotherapy for localised primary intracranial germinomas. *Clinical Oncology* 2008; **20:** 253-60

69. Breen S, Kehagioglou P, Usher C, et al. A comparison of conventional, conformal and intensity-modulated coplanar radiotherapy plans for posterior fossa treatment. *British Journal of Radiology* 2004; **77:** 768-74

70. Fogliata A, Nicolini G, Alber M, et al. On the performances of different IMRT treatment planning systems for selected paediatric cases. *Radiation Oncology* 2007; **2:**

71. Fogliata A, Yartsev S, Nicolini G, et al. On the performances of Intensity Modulated Protons, RapidArc and Helical Tomotherapy for selected paediatric cases. *Radiation Oncology* 2009; **4:**

72. Jones B. Correspondence: Posterior fossa treatment. *British Journal of Radiology* 2005; **78:** 285

73. Plowman P, Cooke K and Walsh N. Indications for tomotherapy/intensity-modulated radiation therapy in paediatric radiotherapy: extracranial disease. *British Journal of Radiology* 2008; **81:** 872-9

74. Brenner D. Effective dose: a flawed concept that could and should be replaced. *British Journal of Radiology* 2008; **81:** 521-3

75. ICRP PUBLICATION 89: Basic Anatomical and Physiological Data for Use in Radiological Protection: Reference Values; 2001. ICRP

76. ICRP: Reference Computational Phantoms of the Adult Male and Female (draft summary); International Commission for Radiological Protection

77. Preston D, Ron E, Tokuoka S, et al. Solid cancer incidence in atomic bomb survivors: 1958-1998. *Radiat Research* 2007; **168:** 1-64

78. Suit H, Goldberg S, Niemerko A, et al. Secondary carcinogenesis in patients treated with radiation: A review of data on radiaiton-induced cancers in human, non-human primate, canine and rodent subjects. *Radiation Research* 2007; **167:** 12-42

79. Tubiana M. Can we reduce the incidence of second primary malignancies occurring after radiotherapy? A critical review. *Radiotherapy & Oncology* 2009; **91:** 4-15

80. Jarlskog C and Paganetti H. Risk of developing second cancer from neutron dose in proton therapy as function of field characteristics, organ and patient age. *IJROBP* 2008; **72:** 228-35

81. McKenzie A, Briggs G, Buchanan R, et al.: IPEM 92 Balancing Costs and Benefits of Checking in Radiotherapy; 2006.