Radiotherapy Dose Calculation in Oesophageal Cancer:
Comparison of Analytical and Monte Carlo Methods

A thesis submitted in partial fulfilment

of the requirement for the degree of Doctor of Philosophy

Dewi M. Johns

December 2016

Cardiff University
School of Engineering

Declaration

This work has not been submitted in substance for any other degree or award at this or any other university or place of learning, nor is being submitted concurrently in candidature for any degree or other award.

Signed .... (candidate)

Date ....

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD.

Signed .... (candidate)

Date ....

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated, and the thesis has not been edited by a third party beyond what is permitted by Cardiff University's Policy on the Use of Third Party Editors by Research Degree Students. Other sources are acknowledged by explicit references. The views expressed are my own.

Signed .... (candidate)

Date ....

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed .... (candidate)

Date ....

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available online in the University's Open Access repository and for inter-library loans after expiry of a bar on access previously approved by the Academic Standards & Quality Committee.

Signed .... (candidate)

Date ....

To my parents
for their love and support.

Learning is the only thing for you. Look at what a lot of things there are to learn.

<div align="right">

T. H. WHITE

THE ONCE AND FUTURE KING

</div>

Abstract

In this work a distributed computing system (RTGrid) has been configured and deployed to provide a statistically robust comparison of Monte Carlo (MC) and analytical dose calculations. 52 clinical oesophageal radiotherapy plans were retrospectively re-calculated using the Pencil Beam Enhanced (PBE) and Collapsed Cone Enhanced (CCE) algorithm within the Oncentra v4.3 radiotherapy (RT) Treatment Planning System (TPS). Simulations were performed using the BEAMnrc and DOSXYZnrc codes. The Computing Environment for Radiotherapy Research (CERR) has been used to calculate Dose Volume Histogram (DVH) parameters such as the volume receiving 95% Dose for the Planning Target Volume (PTV) for the PBE, CCE and MC calculated dose distributions. An initial sample of 12 oesophageal radiotherapy treatment plans were simulated using the RTGrid system. The differences in the DVH parameters between the dose calculation methods, and the variance in the 12 cases, were used to calculate the sample size needed. The required sample size was determined to be 37, so a further 40 oesophageal cases were simulated, following the same method. The median difference in the PTV V95% between CCE and MC in the group of 40 cases was found to be 3%. To choose a suitable test for the statistical significance of the difference, the Shapiro-Wilk test was performed, which showed that the differences between the two sets of PTV V95% values did not follow a Gaussian. Therefore the Wilcoxon matched pairs test was indicated, which showed that the null hypothesis (i.e. that the distributions are the same) was rejected with a p-value less than 0.001, so there is very strong evidence for a difference in the two sets of values of PTV V95%. Similar statistical analyses were performed for other DVH parameters, as well as Conformance Indices used to describe the agreement between the 95% dose and the PTV, and estimates of the Tumour Control Probability (TCP). From the results, the use of MC simulations are recommended when non-soft tissue voxels make up $> 60\%$ of the PTV.

Acknowledgements

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Cancer

1.1.1 What is cancer?

Cancer is the term given to a disease in which cells of the body grow in an uncontrolled way, potentially spreading to different parts of the body, and the various illnesses caused by cancerous cells invading other tissues. There are over 200 different forms of cancer, relating to cells of different types and locations (Peedell, 2005). A neoplasm (literally, new growth) is an abnormal growth of tissue, due to rapidly proliferating cells, and may be malignant or benign (Seeley et al., 2000). A malignant (from the Latin *malignant*, 'with malice') neoplasm is a growth that can grow and spread. Benign (from the Latin *benignus*, 'kind',) neoplasms are unlikely to continue to grow or spread. The common term 'tumour' can refer to any swelling, but is most commonly used to refer to a neoplasm. Specific terms may be used depending on the type of tissue from which the tumour has grown. For example a malignant neoplasm growing from epithelial cells (type of cells lining surfaces, including the outside of the body, and the lining of cavities, digestive tract etc.) is often called a carcinoma. A malignant neoplasm growing from connective tissue cells is called a sarcoma.

Hanahan and Weinberg (2000) describe six 'hallmarks' of cancer, which distinguish most and perhaps all cancer cells from normal cells.

1. **Self sufficiency in growth signals:** Normals cells will only divide and proliferate when they receive a biochemical 'signal' to do so from the body. Cancer cells are able to signal themselves to proliferate.

2. **Insensitivity to anti-growth signals:** Within normal tissue, there are various mechanisms to prevent excessive growth by preventing proliferation. Cancer cells are able to ignore signals to stop proliferating.

3. **Evading apoptosis:** apoptosis is a term for 'programmed' cell death, a process used by the body to kill cells. Cancer cells are able to avoid apoptosis.

4. **Limitless Replicative Potential:** In order for a single cancerous cell to grow to a macroscopic tumour, it must replicate at a rate far greater than that of normal tissues.

5. **Sustained angiogenesis:** In order to proliferate, the tumour needs a blood supply, to provide oxygen and nutrients. Cancer cells have the ability to encourage new blood vessel growth, termed angiogenesis.

6. **Tissue invasion and metastasis:** Most tumours are able to produce 'pioneer' cells that spread, typically to nearby tissues and lymph nodes, and then to further sites.

These differences are due to mutations in genes that regulate cell proliferation and division. These mutations are predominantly acquired, rather than inherited (Peedell, 2005), and may be caused by chemical, radiological or biological factors.

1.1.2 Cancer statistics

The number of people being diagnosed with cancer is increasing. The lifetime risk of cancer for people born in the UK since 1960 is now greater than 1 in 2 overall (53.3% for men, 47.5% for women) (Ahmad et al., 2015). This is an increase from the figure of 1 in 3 for those born since 1930. The lifetime risk of cancer, as defined by Ahmad et al. (2015), means that as average lifespans increase there will be a corresponding increase in the lifetime risk of cancer. The proportion of the population living to at least 65 in the UK is expected to reach 24% in 2040, compared with 16% in 2009 (Maddams et al., 2012). There are other factors that lead to the observed increase in lifetime risk of cancer diagnosis, such as increasing obesity, women having fewer children and having children later in life, increases in red meat consumption, and better screening for certain cancers, e.g. Prostate Specific Antigen (PSA) screening for prostate cancer, leading to more cancers being diagnosed (from Ahmad et al. (2015) and references therein).

There are two related, but different ways of looking at death from cancer: mortality rate, which is the proportion of people dying from cancer per 100,000; and cancer deaths, which is the total number of deaths from cancer, usually expressed as deaths within a calendar year. Mortality rates from cancer have decreased since 1990, and are expected to continue to decline (Olsen et al., 2008). However, the absolute increase in cancer diagnoses outweighs the decrease in mortality, so that the number of cancer deaths is still increasing (Olsen et al., 2008).

The chances of surviving cancer are steadily increasing. Quaresma et al. (2015) have analysed cancer survival rates in terms of a survival index, which is independent of changes in population age and prevalence of cancer in groups of different age and gender. This survival index shows an increase in 10 year post-diagnosis survival for both genders combined from 1 in 3 (34.4%) for 1990-1991 to 1 in 2 (49.8%) in 2010-2011. Quaresma et al. (2015) caution that this does

not mean that half of patients will have their cancers cured, rather that half will not die from their cancer. The survival at 10 years for different cancers varies dramatically, from 1% for pancreatic cancer to 94% for testicular cancer.

Furthermore, the absolute number of people surviving cancer is increasing. Maddams et al. (2012) expect that the number of cancer survivors in the UK will increase by one million per decade from 2010 to 2040, so that by 2040 1 in 4 people in the UK will be a cancer survivor.

There are currently three main treatments for cancer, surgery, radiotherapy (RT) and chemotherapy. These may be used on their own, or in a combination. It is estimated that radiotherapy contributes to the cure of 40% of patients diagnosed with cancer who survive five years or more, second only to surgery in its curative effectiveness (Bentzen et al., 2005). RT is the use of ionising radiation to treat cancer. The majority of RT is External Beam Radiotherapy (EBRT), in which a beam of radiation from a treatment machine is directed at the cancer from outside the body. However, the term includes brachytherapy, in which radioactive sources are placed into the body, and unsealed source therapy, in which radioactive chemicals in liquid form are introduced into the body. There is increasing interest in immunotherapy, which aims to either enhance the body's natural immune response to cancer, or to help the immune system identify cancer cells, so that they will be attacked[1].

1.2 Oesophageal cancer

Oesophageal cancer was the eighth most commonly occurring cancer worldwide in 2012 (Stewart and Wild, 2014), and the thirteenth most commonly occurring

---

[1]Immunotherapy. Available at: http://www.cancer.net/blog/2016-02/immunotherapy-2016-clinical-cancer-advance-year (Accessed: 16 December 2016).

cancer in the UK in 2013[2]. The 8,784 new cases of Oesophageal cancer diagnosed in the UK in 2013 accounted for 2% of the total cancer cases in the UK that year. However, oesophageal cancer accounted for 5% of the deaths from cancer in the UK in 2014[3]. These trends are similar for Wales, where oesophageal cancer accounted for 2,181 cancer diagnoses between 2010 and 2014, out of a total of 93,976 (2.3%). However, oesophageal cancer accounted for 2,091 deaths from cancer between 2010 and 2014 in Wales, out of a total of 43,323 deaths (4.8%) in the same period[4].

Oesophageal cancer commonly presents with locally advanced or metastatic disease (Mukherjee et al., 2015). Survival rates for oesophageal cancer in the UK for 2010-2011 were 42% at one year after diagnosis, 15% at five years after diagnosis and 12% at ten years after diagnosis[5].

Stewart and Wild (2014) state that the two most common types of oesophageal cancer are squamous cell carcinoma and adeno-carcinoma. These are cancers of squamous cells (which are found in the skin, and linings of certain organs) and cancers of the epithelial tissue (which lines the oesophagus, as well as other organs), respectively. According to Peedell (2005), 65%-70% of oesophageal cancer cases are adeno-carcinomas, although 20-30 years ago the majority of cases were squamous cell carcinoma. Risk factors for squamous cell carcinoma of the oeso-

---

[2]Cancer Research UK. Oesophageal Cancer Incidence Statistics. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer/incidence#heading-Zero (Accessed: 18 November 2016).

[3]Cancer Research UK Oesophageal Cancer Mortality Statistics. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer/mortality (Accessed: 18 November 2016).

[4]Welsh Cancer Intelligence and Surveillance Unit, Welsh Cancer Statistics. Available at: http://www.wcisu.wales.nhs.uk/offical-statistics-exel-files-of-trend (Accessed: 13 December 2016).

[5]Cancer Research UK. Oesophageal Cancer Incidence Statistics. Available at: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/oesophageal-cancer/incidence#heading-Zero (Accessed: 18 November 2016).

phagus include alcohol consumption, smoking or chewing tobacco, and drinking hot beverages. Squamous cell carcinoma of the oesophagus tends to arise in low resource regions of the world.

The most significant risk factors for adeno-carcinoma of the oesophagus are obesity, gastro-oesophageal reflux disease, and smoking or chewing tobacco. Another risk factor is Barratt's Oesophagus, a complication of gastro-oesophageal reflux disease in which the lining of the oesophagus begins to resemble that of the intestine. However, only 5% to 10% of patients with Barratt's Oesophagus go on to develop adeno-carcinoma of the oesophagus. The UK has the highest incidence of adeno-carcinoma of the oesophagus in the world (Stewart and Wild, 2014), and the highest age-standardised incidence of oesophageal cancer in Europe (Gatenby et al., 2011). The incidence of oesophageal cancer has increased in men over the last 35 years (Figure 1.1), whereas in women, after a peak incidence in 1990s, rates have slightly declined. According to Parkin et al. (2011), 89% of cases of oesophageal cancer in the UK are preventable, i.e. due to lifestyle factors.

Quaresma et al. (2015) grouped the 21 most common cancers into three clusters, based on survival. The first cluster has the best prognosis, and includes breast, prostate and testicular cancer. The second group has poorer prognosis, and includes bladder cancer and leukaemia. Oesophageal cancer falls into the cluster with the poorest prognosis, with pancreatic and lung cancer.

The most common symptom of oesophageal cancer is difficulty swallowing, initially solids, and later liquids, which is also known as progressive dysphagia. Patients may also suffer from weight-loss. If the oesophagus is blocked, this may cause coughing, regurgitation of food, choking on food, and pneumonia (Peedell, 2005).

Screening using endoscopy has proved to be expensive and error-prone (Stewart and Wild, 2014). Non-endoscopic methods involve threading a deflated balloon through the oesophagus, inflating the balloon, then pulling the balloon back
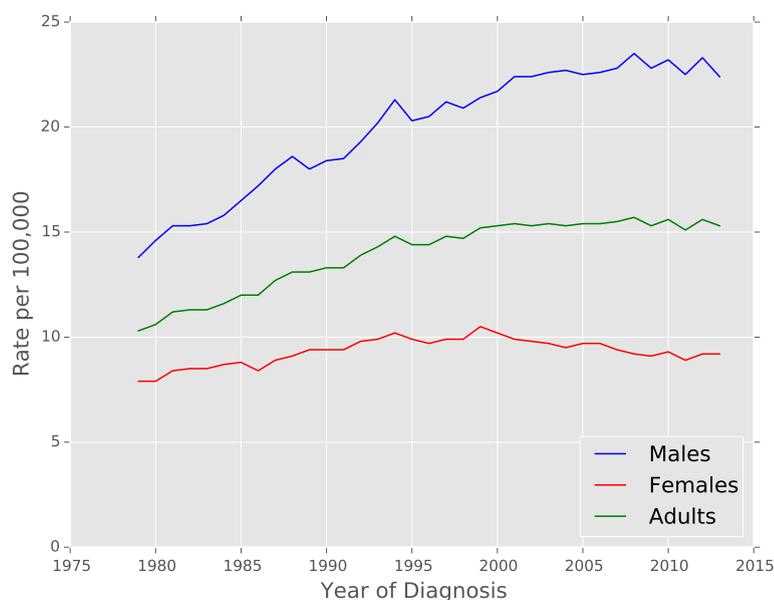
Figure 1.1: Age-standardised incidence rates of oesophageal cancer per 100,000 population, by sex, for Great Britain.

through the oesophagus. Cells from the lining of the oesophagus adhere to the surface of this balloon, which can be tested for cancerous cells. Developing a blood test for oesophageal cancer is the subject of significant current research interest in Wales[6].

Surgery is the most common treatment for early stage oesophageal cancer. Medical Research Council (MRC) trial OE02 (Allum et al., 2009) showed an advantage to neo-adjuvant chemotherapy (N.B. no radiotherapy) and surgery compared to surgery alone. Chemo-radiotherapy (CRT) is commonly used if the cancer has spread, or the patient is too unwell to tolerate surgery (Gwynne et al., 2011). CRT may be used in two main ways, as the sole treatment, known as definitive CRT, or prior to surgery, called neo-adjuvant CRT. Survival at two years from definitive

---

[6]Cancer Research Wales. New Smoke Detector Test For Oesophageal Cancer. Available at: https://www.cancerresearchwales.co.uk/blog/new-smoke-detector-test-for-cancer/ (Accessed: 28 November 2016).

CRT was reported by Gwynne et al. (2011) as over 40%, which is comparable to surgery alone. Gwynne et al. (2013) discussed that most relapses following definitive CRT are due to local progression or local recurrences, suggesting that the standard doses of CRT are not high enough.

1.2.1 SCOPE1 Trial for oesophageal cancer

SCOPE1 (Hurt et al., 2011) was a multi-centre randomised controlled trial to compare CRT with and without a chemotherapy agent called Cetuximab. The trial closed early due to the research arm (CRT plus Cetuximab) showing higher than expected toxicity, poorer completion (completing the whole course of therapy), and poorer survival. However, the control arm (CRT only) showed better overall survival (56%) compared with the prediction of 35% (Crosby et al., 2013). This was attributed to the radiotherapy protocol design and quality assurance used for SCOPE1. However, the authors also cite other positive contributory factors such as patient selection and improvements in the organisation of cancer services, access to Positron Emission Tomography (PET) and the role of Multi-Disciplinary Teams (MDTs) in the management of oesophageal cancer. The trial is considered a watershed in the design of radiotherapy trials, and the importance of quality assurance within such trials - this despite the main trial aim being related to the effect of an additional chemotherapy agent.

The details of the protocol, as they relate to this project, will be discussed in section 3.2.1.

As part of the SCOPE1 trial, test cases were sent to the prospective centres, which required all clinicians to delineate the diseased tissue on an anonymised Computed Tomography (CT) scan. A RT treatment plan had to be prepared on this test case, and the outlines, plan and doses returned to the Radiotherapy Trials Quality Assurance (RTTQA) centre for assessment. RTTQA are a group

funded by the National Cancer Research Institute (NCRI), which ensures that patients in all NCRI radiotherapy trials are treated in accordance with the trial protocol, and are treated according to nationally accepted standards. RTTQA are responsible for the quality assurance of radiotherapy in clinical trials. A centre was only able to enrol patients in the trial if the outlining, planning and dose calculation were deemed to have passed the quality checks by the RTTQA team.

1.3 Radiobiological rationale for radiotherapy

It is common to deliver radiotherapy in a series of 'fractions', with each fraction typically delivered on separate weekdays, over a period of weeks. However, some fractionation schemes may involve multiple fractions per day, or no breaks for weekends, or a single fraction per week. Radiotherapy kills cells by damaging the deoxyribonucleic acid (DNA) in cells. Due to the famous double helix structure of DNA, the damage to it can be classified as a Single Strand Break (SSB) or a Double Strand Break (DSB). The aim of the fractionation schedule is to make the radiation more damaging to the cancerous cells than the surrounding tissue. The five factors, known at the '5 Rs', which affect the choice of fractionation schedule identified by Steel (1997) are Radio-sensitivity, Re-oxygenation, Redistribution, Re-population and Recovery.

1.3.1 Radio-sensitivity

Radio-sensitivity is the amount of damage done by a standard dose of radiation. The ability of the cell to rejoin DSBs in DNA, and the fidelity of such repairs determine the sensitivity of a cell to radiation. (SSBs are easy for the body to repair, as the double helical structure will not alter; a DSB, or many breaks in a small area can cause sections of DNA to become detached, which is hard to repair

correctly. Mis-repair can lead to cell death or mutation; mutation can then lead to carcinogenesis) .

1.3.2 Re-oxygenation

Since 70% of the body is comprised of water, the most common molecule for radiation to interact with is water, or H_2O, commonly producing a positively charged H_2O^+ molecule, and a negatively charged electron, e^-. There follows a series of interactions, leading to various highly reactive ions with unpaired electrons, known as free radicals. Oxygen reacts with these free radicals to produce hydro-peroxy radicals, such as HO_2^-, which are highly toxic to biological structures such as DNA. Therefore damage due to RT is affected by the availability of oxygen. The Oxygen Enhancement Ratio (OER) is the ratio of doses needed to produce a given effect in the presence and absence of oxygen. The OER is 2.5-3 for low Linear Energy Transfer (LET) radiation, such as x-rays or gamma rays (for the higher LET values found in particle therapies, OER values are lower, approaching unity for carbon ion beams). Some regions of tumours have a very low oxygenation level, which makes them resistant to radiation compared to the rest of the tumour and normal tissue.

Re-oxygenation is important as cells that are Hypoxic during the first treatment may become oxygenated for subsequent fractions, increasing their sensitivity to radiation-induced damage.

1.3.3 Redistribution

The sensitivity to radiation damage varies over the cell cycle, the sequence of changes in a cell in which it splits into two cells (Figure 1.2). The cell cycle can broadly be split into:-

- **Gap 0** or **G$_0$:** a resting phase, after the previous division has completed and before a new division starts;

- **Gap 1** or **G$_1$:** cell increases in size;

- **Synthesis** or **S:** DNA is replicated, resulting in two copies of the DNA molecule;

- **Gap 2** or **G$_2$:** further increase in cell size;

- **Mitosis** or **M:** Cell splits into two, each with one of the DNA molecules.
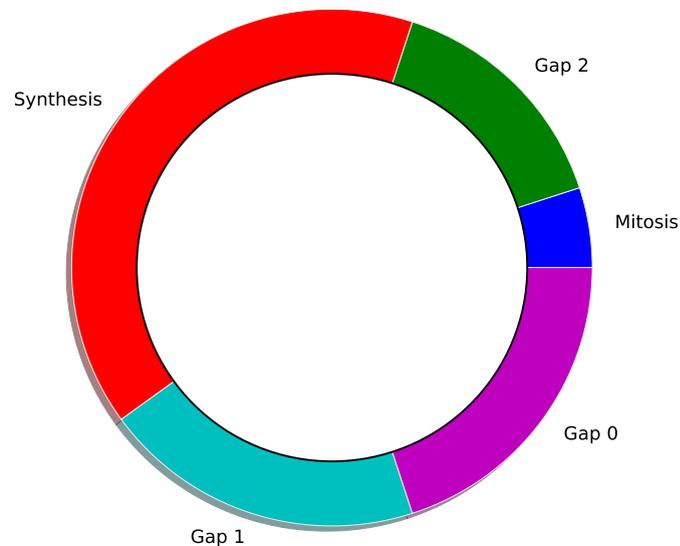
Figure 1.2: Graphic of the phases of the cell cycle, showing Gap 0, Gap 1, Synethsis, Gap 2 and Mitosis phases.

A given cell is relatively more resistant in the late S phase, when the DNA molecule is being replicated, compared to the G$_2$ or M phases. At irradiation, tumour cells will be in different stages of the cell cycle. A cell in the S phase during the first fraction may be in the G$_2$ or M phase for subsequent fractions, so may be killed by the radiation - this effect is known as **Redistribution**. This

re-assortment means that fractionating the treatment will kill more cells. This effect is complicated by checkpoints in the cell cycle, which slow down the cycle in cells with damaged DNA.

1.3.4 Re-population

Tumours undergo **Re-population** between fractions by cell proliferation. The rate of cell division in treated tumours can be accelerated, possibly due to the increased abundance of nutrients. This limits the total time a treatment can last.

1.3.5 Recovery

Cells can also undergo **Recovery** from radiation damage, which reduces the number of cells killed during prolonged treatments, or following missed fractions.

1.4 Therapeutic window

The aim of radiotherapy is to give enough radiation to the cancerous cells to produce a desired effect , while keeping the effect on healthy tissue to acceptable levels. The desired effect is to kill enough cells so that the body can deal with the remainder, or reducing tumour size to alleviate symptoms. The proportion of cells in a sample that is killed by a dose of radiation can be approximated by a sigmoid ('S'-shaped) curve. Up to a certain dose of radiation, virtually no cells are killed; as the dose is increased, the proportion of cells rises rapidly at first, as the more radio-sensitive cells are killed. As the dose is further increased, the proportion of cells killed begins to slow, as only the more radio-resistant cells are left. Assuming healthy and cancerous cells exhibit this behaviour leads to the concept of the 'therapeutic window', illustrated in Figure 1.3. If the curves are close together, with the curve for the cancerous cells to the left of the healthy
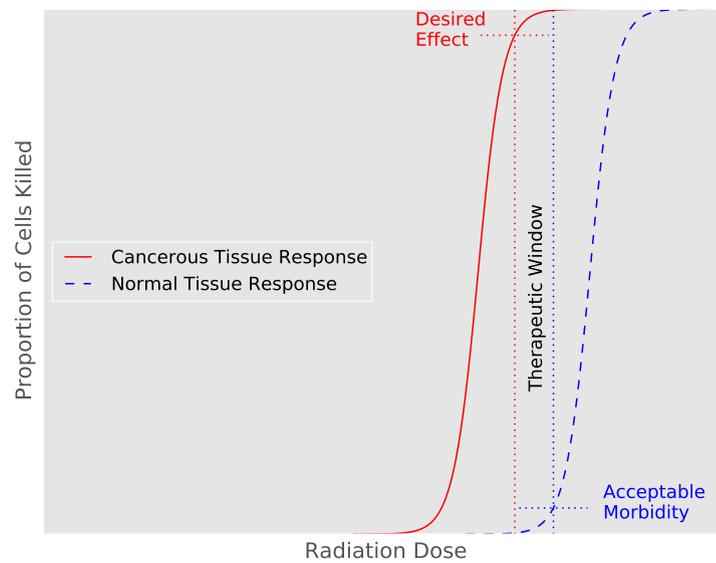
Figure 1.3: Illustration of the concept of the therapeutic window in radiotherapy.

cells, there will be a small range of suitable radiotherapy doses that can satisfy both the requirement to treat the tumour, and to spare healthy tissue. If the curve for the cancerous cells is to the right of the the curve for the healthy cells, there will be no therapeutic window, as is the case for stomach cancer. Different fractionation schedules (or the use of more advanced radiotherapy techniques, to shield healthy tissue better) may widen the therapeutic window. It should be noted that the therapeutic window is essentially an illustrative concept - the curves in Figure 1.3 cannot be determined for an individual patient.

1.5 Reactions to radiotherapy

The reactions to RT can be divided into three categories:

- **acute or early reactions:** Occur within a few weeks of treatment, and

are associated with damage to rapidly dividing cells, and include reactions such as hair loss and reddening of the skin, known as erythema;

- **sub-acute reactions:** Occur between one and six months after radiotherapy, and include inflammation of lung tissue, known as radiation pneumotitis, and fatigue;

- **late reactions:** are usually the reactions that prevent higher doses of radiation being used, and vary between irradiation site, but include tissue death (necrosis), ulceration, nerve damage (neuropathy), inflammation of the sac surrounding the heart (pericarditis) and dry mouth due to damage to the salivary glands (xerostomia).

1.6   Radiotherapy planning and delivery

1.6.1   Treatment Planning Systems

Modern-day RT may be radical (aiming to cure the disease) or palliative (aiming to provide relief from the symptoms of the disease). The RT may involve simple irradiation conditions, or more complex radiation dose distributions, tailored to the individual patient's anatomy. This more complex form of RT is called Conformal Radiotherapy (CFRT). CFRT is recommended by the National Institute for Clinical Excellence for RT of many sites, including Urological, Breast and Head and Neck cancers[7].

In CFRT, a CT scan is taken of the patient and images loaded into specialised software, called a 3D Treatment Planning System (TPS). Oncologists will outline the cancerous tissue they wish to treat, and Organs At Risk (OARs) to which they wish to limit the dose, and state the total radiation dose they want to deliver.

---

[7]NICE website. Available at: https://www.nice.org.uk/ (Accessed: 21 November 2016).

Treatment planning staff then use the 3D TPS to model the radiation treatment to be delivered by a medical Linear Accelerator (linac). Treatment planning staff choose suitable beam entry angles, sizes and shapes to deliver the required dose of radiation to the target volume, while avoiding healthy tissue as far as possible.

The TPS calculates the dose that will be absorbed by the patient anatomy from the chosen beam arrangement, and displays the dose distributions overlaid on the CT scans, using visualisation tools such as Isodoses and Dose Volume Histograms (DVHs). Isodoses are lines (often coloured) that indicate regions getting the same dose (c.f. isobars, lines of equal pressure on a weather map). The isodoses can be used to see regions of the target that may not be getting enough dose, or regions of healthy tissue getting too much dose. DVHs are a type of histogram used to show the volume of a target or OAR that is received a given dose. Commonly, cumulative histograms are used, so the histogram shows the volume getting up to and including a specified dose. Commonly, for a given treatment site there will be objectives given in terms of DVH parameters, i.e. at least 99% of the target volume must receive at least 95% of the prescription dose. The planning staff then adjust the beam arrangement and parameters to maximise the tumour dose, and minimise the sensitive tissue dose, based on the isodoses and DVH.

The decisions made to optimise the dose distribution assume that the dose calculated by the TPS is accurate. However, the two current dose calculation methods used in the TPS used at Velindre Cancer Centre, Cardiff, Elekta Oncentra[8], known as the Pencil Beam (PB) and Collapsed Cone Algorithm (CC) algorithms (Ahnesjö and Aspradakis, 1999), each involve assumptions and approximations to allow dose calculation in a reasonable time. These assumptions may cause the calculated dose to deviate from the true dose at interfaces between tissues of different densities (IAEA, 2013). It is generally accepted that CC is the more accurate algorithm, but it is not always used due to the longer calculation times,

---

[8]Oncentra. Available at: https://www.elekta.com/software-solutions/#treatment-management (Accessed: 21 November 2016).

or poorer resolution for comparable calculation times. There is also the complicating factor that CC reports dose to material, not dose to water as PB does (Ahnesjö, 1989).

There are several advantages of using an algorithm that reports dose to water. Clinical experience is based on calculations that yield dose to water. There may be uncertainties in deducing the material involved in the dose to material calculation. In cases where bone is contained within the Planning Target Volume (PTV), it may not be possible to deliver sufficient radiation to treat the bone to the prescription dose if the dose is reported as dose to medium. The advantages of reporting dose to material are that it is more closely related to the energy absorbed by the physical matter in the patient. There is some discussion on whether the dose to water is a better surrogate for the radiobiological effects on bony tissue (Liu et al., 2002).

The gold standard of dose calculation in radiotherapy is the Monte Carlo (MC) method, which involves simulating radiation transport through the linac and the patient on a particle by particle basis (Verhaegen and Seuntjens, 2003). It is therefore intrinsically very accurate, but very computationally intensive; the dose deposited in a patient from a CFRT treatment may take days of CPU time to calculate by MC, using current computational methods and technology.

1.6.2 Linear accelerators

The majority of EBRT is delivered by linacs (Figure 1.4). A linac is a medical device that accelerates electrons through an evacuated wave-guide to a potential of several million volts. Generally, this wave-guide will not be directed at the patient. Due to the length of the wave-guide, if it were directed at the patient, the assembly would not be able to be positioned under the patient without the patient being suspended several meters above the floor. Therefore, the electron

Figure 1.4: Photograph of a Varian 2100CD linac at Velindre Cancer Centre, aka 'LA6'.

beam must be deflected by beam steering magnets towards the patient. As the electron beam is poly-energetic, different energies would be deflected by different angles in a simple magnetic field, so different manufacturers have devised different methods to perform this bending in a way that does not lead to dispersion, known as achromatic bending. The electron beam is then incident on a heavy metal target, typically tungsten or tungsten-alloy, whereby photons are produced by a process called Bremsstrahlung. As this is the origin of the photons this is termed the 'source'. Distances are defined relative to the plane containing this source, for example the 'source to skin' distance is the distance the patient's skin surface is from the source plane. The generation of photons in this way produces a photon beam that is more intense in the centre of the beam than at the edges, which is not generally suitable for treating patients. Therefore a shaped piece of metal, known

as a flattening filter, is placed in the beam, which absorbs more radiation on the beam axis than at the edges to produce a relatively flat beam profile at a specific depth. This flattening filter causes differential beam hardening across the beam, resulting in a more energetic beam on the central axis than off axis. Currently there are two main linac manufacturers, Varian (Varian Medical Systems, Palo Alto, CA, USA)[9] and Elekta (Elekta Limited, West Sussex, UK)[10]. A specific model of Varian linac, known as a Varian 2100CD has been used throughout this work. Where the design details of a Varian 2100CD differ from a generic linac, this will be highlighted.

A lead block with a cone cut from it, known as the primary collimator, confines the beam to a cone. If the secondary collimators are opened fully the corners of the field will be missing, due to the shape of the primary collimator.

Secondary collimators or jaws are blocks that move in and out of the beam to define the size of the radiation beam (Figure 1.5). There are upper collimators that define the field size in one direction, and lower collimators mounted below them to define the field in the perpendicular direction. Due to the finite radiation source size, there will not be a drop in dose from 100% to 0% at a single point at the edge of the field. There will be a penumbral region in which the collimator blocks part of the source (the effects of scatter blur the beam edge further). The size of the penumbra region depends on the source to skin distance, and the distance between the collimator and the point halfway between the source and the patient surface. Therefore, the upper collimators have a larger penumbra than the lower collimators. To reduce the penumbra, in some linacs such as the Varian Clinac 2100CD, the edge of the collimator can be aligned with the beam divergence (Figure 1.5). For this the collimator must move along an arc,

---

[9]Varian Linacs. Available at: https://www.varian.com/oncology/products/treatment-delivery (Accessed: 3 December 2016)

[10]Elekta Linacs. Available at: https://www.elekta.com/radiotherapy/treatment-delivery-systems.html (Accessed: 3 December 2016)

which increases the height of the treatment head. When radiation is incident on the face of the collimator, rather than the upper surface, the radiation will not be attenuated by the full thickness of the collimator. Radiation entering the collimator near the lower edge will have a shorter path length, and be attenuated
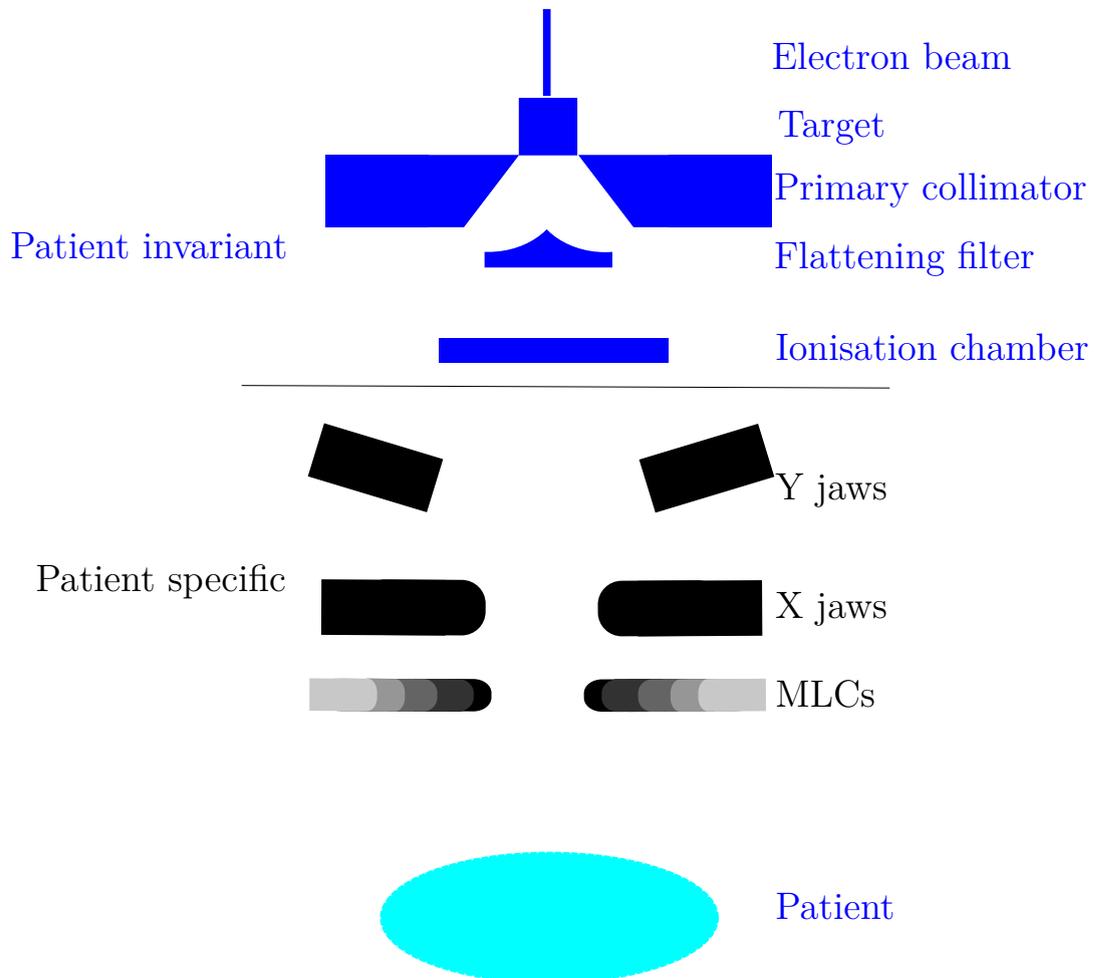
Figure 1.5: Block diagram of Varian 2100CD linac. Y jaws, X jaws and Multi-Leaf Collimators (MLC) shown as moving in same plane for illustrative purposes only. The target, primary collimator and flattening filter are fixed for a given energy, and are referred to as the patient invariant portion of the linac head. The jaws and MLCs are positioned depending on the patient, so are referred to as the patient specific portion of the linac head.

less, than radiation entering near the upper edge. This contributions to the size of the penumbra; Figure 1.6 illustrates that angling the face of the collimator to align with the beam divergence reduces the size of the penumbra. On Varian machines, only the upper collimators move in an arc, to save on isocentric height.
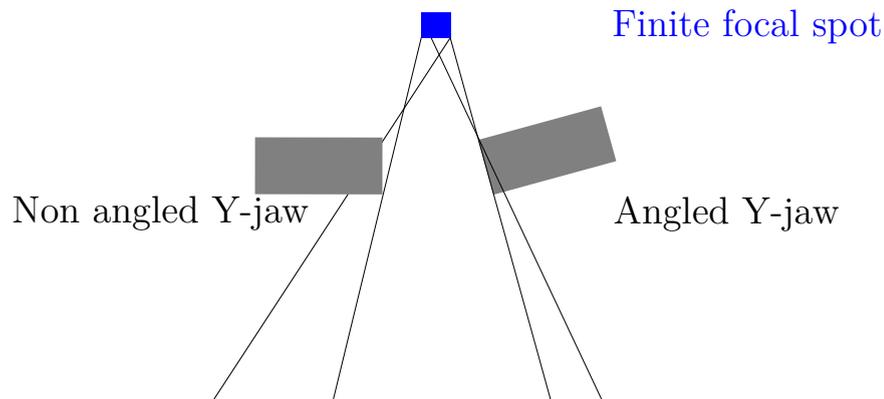
Figure 1.6: Diagram illustrating how aligning the face of the collimator with the beam divergence reduces the beam penumbra.

A hard wedge is a physical piece of metal, shaped to give an isodose angled at 15°, 30°, 45° or 60° to a plane at right angles to the central axis at a specified depth. For Varian hard wedges, this depth is 5 cm along the the central axis, following the approach of Hughes et al. (1972). The varying thickness of metal produces differential beam hardening and scatter across the field.

A flying wedge is a motorised 60° physical wedge in the linac head, which can be driven into and out of the beam. If for example the flying wedge is in the beam for a given fraction of the treatment time, a similar profile will be produced as if a 30° wedge had been used. These wedges still have differential beam hardening and scattering across the beam. Varian 2100 linacs do not use flying wedges.

For machines in which the collimators move asymmetrically, one jaw can be driven across the beam during irradiation, such that a wedged isodose will be produced. On a Varian linac, the wedge angle is defined at 10 cm deep, following the approach of ICRU (1976). This means there is no differential beam hardening or

scatter across the beam. Any wedge angle up to 60° could be produced in theory, but profiles would have to be measured, and quality assurance performed, on each. Only 15°, 30°, 45° and 60° Enhanced Dynamic Wedges (EDWs) are used at Velindre Cancer Centre.

An MLC consists of scores of pairs of thin leaves (40, 60 and 80 are common configurations), which can be moved to produce an irregular field shape, to improve conformity $\big($Figure 1.7(a)$\big)$. Due to restrictions on isocentric height, the leaves cannot move in an arc to keep their edges at the divergence angle of the field, so are rounded, to produce as low a penumbra as possible at all positions $\big($Figure 1.7(b)$\big)$.

One of the most important components of the linac is the monitor chamber. This is a thin, transmission ionization chamber, positioned below the flattening filter, and above any beam shaping collimators. While most radiation will pass through the chamber (it is constructed from low attenuation materials) some will cause ionisation in the gas within the ionization chamber. The chamber has parallel plate electrodes, which collect the charge. The resultant signal will be proportional to the amount of radiation passing through the chamber, so can be used to monitor (hence the name) the radiation fluence above the beam shaping collimators.

The chambers are divided into segments, and analysing the relative signal from different segments can be used in feedback loops to correct problems with the beam generation. For example, changes in the electron beam energy incident on the target will affect the relative intensity in the centre of the beam compared to the edge of the beam. If the electron beam hits the tungsten target off centre, the beam will be more intense on one side of the beam than the other. It is also common to have two chambers, one 'downstream' of the other, where the second acts as a backup to the first. The linac will be calibrated in terms of Monitor Units (MU), a way of linking the signal from the monitor chamber to a radiation

dose. The MU will be calibrated in such a way that 100 MUs will deliver 1 Gray (J/Kg) of dose, under reference conditions. Measuring under reference conditions involves measuring at 10 cm deep in a water phantom at standard temperature and pressure, placed with its surface 90 cm from the source of the radiation (the tungsten target), when the collimators define a 10 cm by 10 cm field. When delivering 100 MU, the dose given to a patient, or test phantom, can be altered by altering the collimator settings or MLC shape, or moving the patient closer or further from the linac. The MU is designed to take account of fluctuations in the intensity of the radiation beam, but to be independent of how the patient is set up, or how the collimators are configured. However, some linacs are affected by backscatter from collimators into the monitor unit chamber (section 3.2.4.6). Calculating the number of MUs to deliver the required dose in a patient is usually performed by a TPS.

1.6.3 Radiotherapy computer systems

The planning, delivery and monitoring of RT involves the use of many computer systems, which must communicate with one another. Medical images and certain other data are often transferred between computer systems in Digital Imaging and COmmunications in Medicine (DICOM) file format, which is an international standard for the transfer of medical images. DICOM is used to transfer CT, PET and many other medical images, each known as a modality in DICOM terminology. Digital Imaging and COmmunications in Medicine, Radiotherapy Supplement (DICOM-RT) is an extension to the DICOM standard for data used in radiotherapy, and defines seven additional modalities: -

- RTSTRUCT, which describes the details of the OARs, target volumes and the external contour;

- RTPLAN, which describes details about the treatment plan;

- RTIMAGE, for a type of image peculiar to radiotherapy, called a Digitally Reconstructed Radiograph (a prediction of a 2D image based on a 3D dataset);

- RTDOSE, which describes the dose calculated for a plan, in a matrix of voxels;

- RTRECORD, to record the parameters used during external beam treatment delivery;

- RTBRACHYRECORD, to record the parameters used during delivery of Brachytherapy;

- RTSUMMARYRECORD, to record dose delivered during a course of treatment

Due to the extensive use of DICOM and DICOM-RT as a transfer standard, it is very common to store files that follow the formatting defined by the DICOM and DICOM-RT standards, so the transfer standards have become *de facto* file format standards too. There are therefore a number of tools available to read, manipulate and save DICOM and DICOM-RT files.

1.7   Summary

Cancer is one of the leading causes of death in the world. While oesophageal cancer is not the most prevalent form of cancer, it does have significant mortality rates in Wales, UK and across the world. Many forms of cancer are treated with radiotherapy, either as the sole treatment or in conjunction with other treatments, such as chemotherapy or surgery. The SCOPE1 trial demonstrated the importance of RTTQA for oesophageal radiotherapy, and established a protocol for delineation and treatment planning that is still the standard for oesophageal

treatments in Velindre Cancer Centre. These treatments are designed on TPS, and delivered by linacs. If the TPS employs an algorithm with too many approximations or assumptions, the treatment delivery parameters produced may not result in the delivery of the required dose, which may compromise the efficacy of the treatment.

1.8   Thesis objectives

The aim of this study is to establish whether MC methods provide significantly different results to conventional analytical dose calculations used in conformal photon radiotherapy treatment planning for oesophageal cancer. MC calculations are much slower than the algorithms used in TPS. Using MC clinically will require additional resources, and determining whether MC is important will allow appropriate and effective deployment of resources in the future.

Over the last 10 years, work has been undertaken at Velindre Cancer Centre to develop a system known as RTGrid (Downes et al., 2009) to perform MC dose calculations for CFRT, on a variety of computational resources. These resources include those at the Welsh e-Science Centre (WeSC)[11] which can be controlled by RT physicists at Velindre Cancer Centre, or computational resources at Velindre Cancer Centre itself. The development of RTGrid was originally funded by a £400,000 grant from the Engineering and Physical Sciences Research Council[12]. Expected plan turnaround is a few hours, rather than the several days required for the same task on a single PC. These resources have been utilised for this research. This study will compare PB, CC and MC, for a large number of clinical cases to determine in which cases each mode of dose calculation is appropriate.

---

[11]WeSC. Available at: http://www.cs.cf.ac.uk/wesc/ (Accessed: 21 November 2016).

[12]EPSRC grant. Available at: http://gow.epsrc.ac.uk/NGBOViewGrant.aspx? GrantRef=EP/C538641/1 (Accessed: 21 November 2016).

This will allow effective deployment of clinical resources to the situations where they are most needed.

The study cohort will be conventional clinical CFRT plans calculated using PB, on the clinical TPS, Elekta Oncentra, at Velindre Cancer Centre. These can be re-calculated using CC, also using the clinical TPS, and using MC on the RTGrid. Differences between MC, CC and PB calculations can then be determined in the required regions of interest. These will inevitably include the PTV, which includes the tumour plus a margin for clinical spread, and a further margin for set-up uncertainties (ICRU, 1993, 1999) together with any clinically important radiation-sensitive normal structures, or OARs. Over 2500 CFRT plans are prepared per annum in Velindre Cancer Centre, including approximately 50 oesophageal plans per annum, which can be used for the study. The RTGrid system will be used to export the necessary parameters from the clinical plan and return the MC dose distributions to the PTV and OARs for comparison. Data analysis and reduction will be performed using a variety of systems including analysis software previously developed in Velindre Cancer Centre for comparing MC calculated doses and doses from commercial treatment planning systems (Spezi et al., 2002).

1.9 Thesis outline

The structure of the thesis is as follows: -

- Chapter 1: Introduction

- Chapter 2: Overview of dose calculation methods in external beam Radiotherapy. Distributed computing for MC calculations. Clinical relevance of dose calculations.

- Chapter 3: The use of the RTGrid system, including changes required to

allow simulation of oesophageal plans. Pilot study and sample size calculation.

- Chapter 4: Main study, analysis of differences between calculation methods.

- Chapter 5: Conclusions.

- Appendix: Dissemination
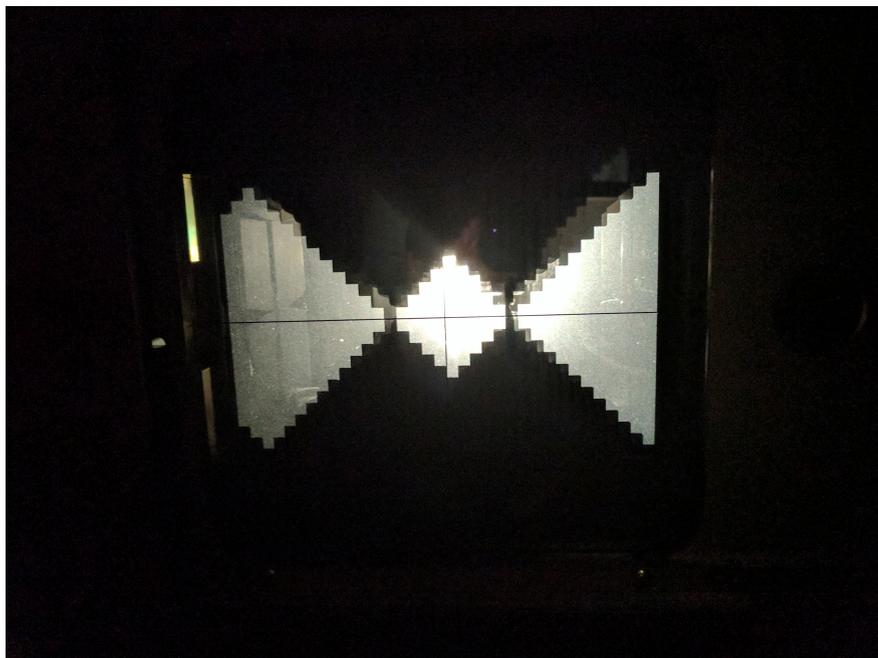
(a) Millenium 120-leaf MLC, defining the field according to a standard test shape.
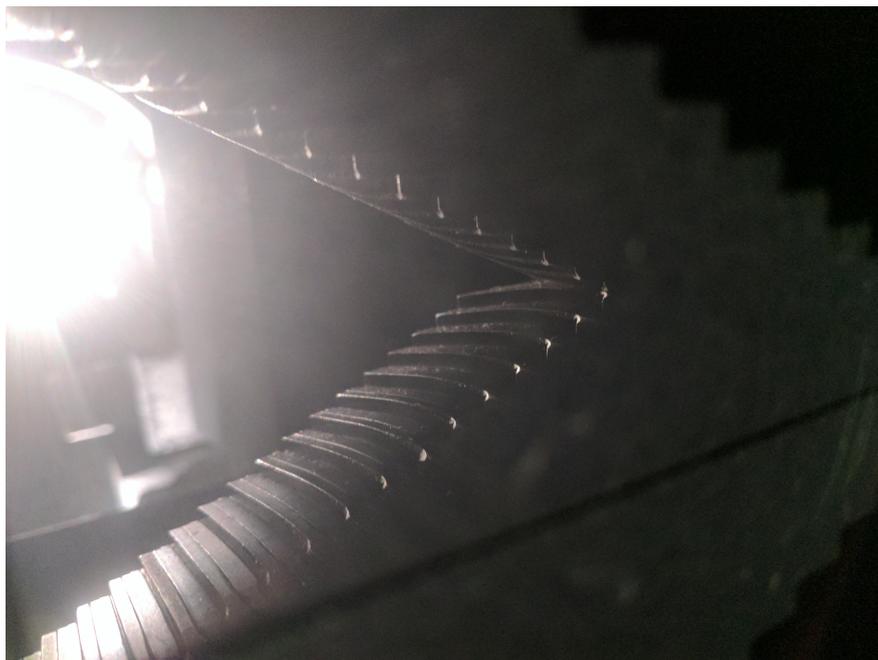
(b) Millenium 120-leaf MLC, in which the curved leaf ends can be seen.

Figure 1.7: Photographs of Millennium 120-leaf MLC on a Varian 2100CD linac.

Chapter 2

Dose calculation for external beam radiotherapy

2.1 Introduction

This chapter will provide an overview of dose calculation methods for EBRT, starting with a discussion of the interaction of radiation with matter. Superposition and convolution dose calculation methods will be discussed, and the PB and CC algorithms used in Oncentra v4.3 will be described. An overview of the MC method will be presented, with a focus on the BEAMnrc package. To conclude the chapter, a review of MC simulations on distributed computing platforms will be presented.

2.2 Photon interactions with matter

The attenuation of photon beams passing through an absorber is due to five major types of interactions of photons with matter (Khan, 2003):

1. Coherent scattering;

2. Photo-electric effect;

3. Compton scattering;

4. Pair production;

5. Photo-disintegration.

Photo-disintegration only occurs above a threshold, which is over 10 mega-electron volts (MeV) for most nuclei (but is approximately 2 MeV for beryllium and deuterium). Even above the threshold, it has a low probability of occurring, and will not be discussed further.

Neglecting photo-disintegration, the total mass energy attenuation co-efficient for a photon beam, μ/ρ, is given by Equation 2.1:

$$\mu/\rho = \sigma_{coh}/\rho + \tau/\rho + \sigma_c/\rho + \kappa/\rho \tag{2.1}$$

where σ_{coh}, τ, σ_c and κ are the attenuation coefficients for coherent scattering, photoelectric effect, Compton scattering and pair production respectively. These four processes will be discussed below.

In EBRT, the range of energies encountered is $\leqslant$25 MeV (Velindre Cancer Centre uses energies of 6 MeV and 10 MeV). The effective atomic number (Z) is 7.4 for muscle and water and is 13.8 for bone, although higher Z values may be encountered in metal implants - hip prostheses commonly contain titanium ($Z = 22$). Figure 2.1 shows the most probable interaction for a range of Z values and energies. It follows that Compton scattering is the most probable interaction for EBRT, while Photo-electric effect (PE) is most probable for the lower energies used in medical imaging, and pair production is important in EBRT for materials with high Z.
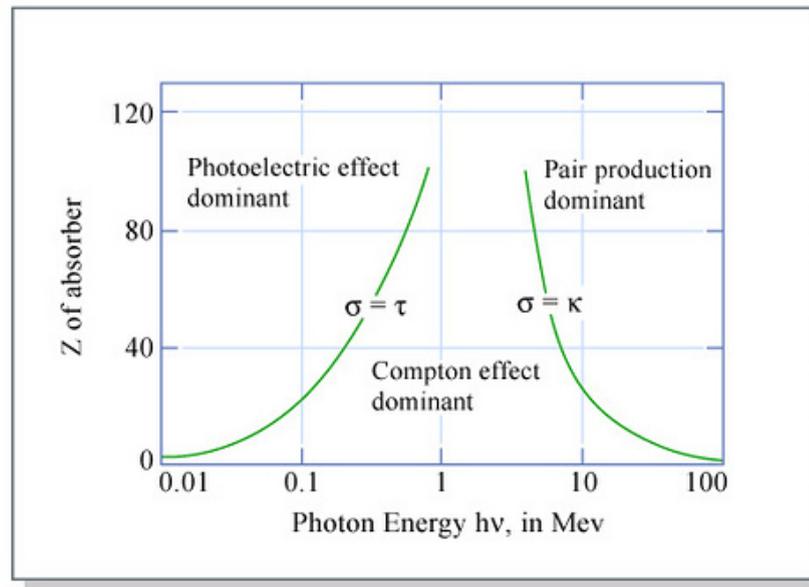
Figure 2.1: Illustration of the relative importance of the three principal types of photon interaction in an absorber of atomic number Z, for a photon with energy $h\nu$ in the range 0.01-100 MeV, where τ, σ and κ are the probability of photoelectric, Compton and pair interaction occurrence, respectively.

Short M. (2015) *Photoelectric Effect*[Online]. Available at:
https://www.flickr.com/photos/mitopencourseware/3706466731 (Accessed: 25 July 2017).

2.2.1 Coherent scattering

In coherent scattering, also known as classical scattering or Rayleigh scattering, no energy is converted to kinetic energy, but photons are deflected (Figure 2.2). Considering the radiation beam as an electro-magnetic wave (or a superposition of waves), the wave induces oscillations in the electrons bound to atoms (c.f. induced oscillations of electrons in an antenna). Coherent scattering leads to a broadening of radiation beams below 100 kilo-electron volts (keV), such as those used in diagnostic radiology (Johns and Cunningham, 1983). It is not shown in Figure 2.1 as it is not the most probable interaction for any energy range or value of Z encountered in radiotherapy.
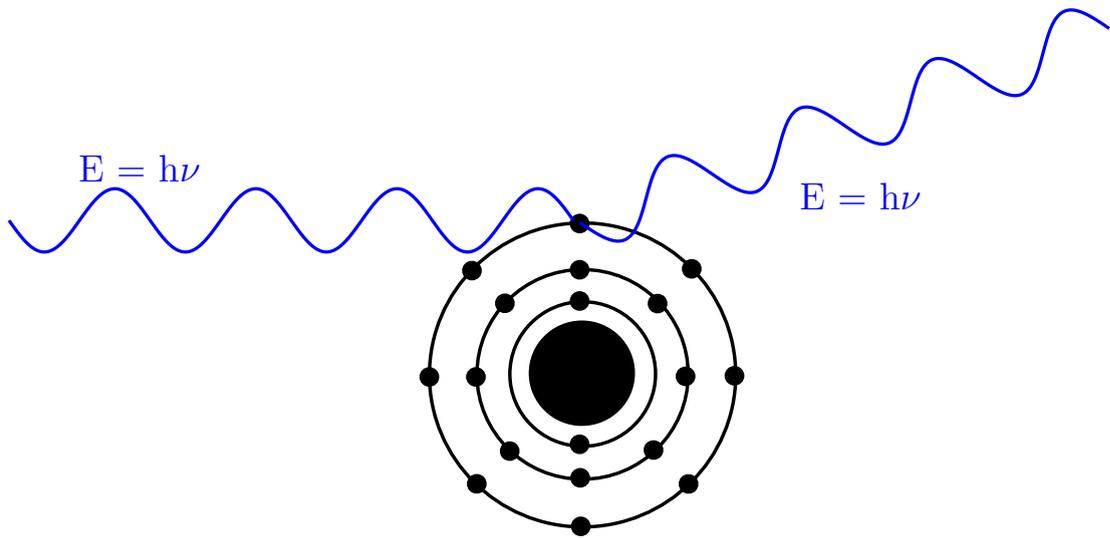
Figure 2.2: Illustration of a photon undergoing coherent scattering. The scattered photon has the same energy, $h\nu$, as the incident photon, and no energy is transferred.

2.2.2 Photo-electric effect

PE is the process in which a photon interacts with an atom and causes a bound electron to be ejected (Figure 2.3). The energy of the electron is entirely absorbed by the atom, and is then transferred to the ejected electron. If the photon energy is given by $E = h\nu$, where h is Planck's Constant and ν is the frequency of the photon, the ejected electron will have an initial energy of $h\nu - E_B$, where E_B is the binding energy of the electron. If $h\nu$ is less than E_B for the lowest energy electron state, PE is not possible. If the electron is ejected from an inner shell, an electron from an outer shell can drop down to fill the hole. A photon will then be emitted with a characteristic energy, E_{Char}, equal to the difference between the energy levels of the two shells (Figure 2.4).
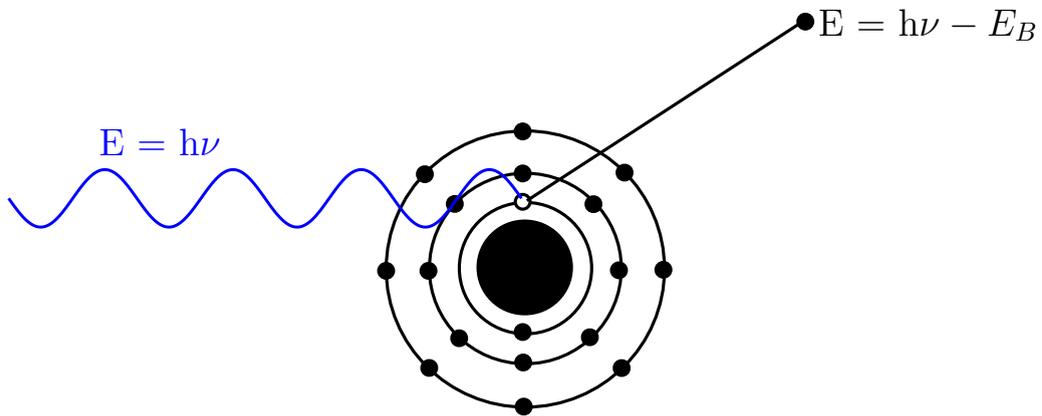
Figure 2.3: Illustration of the photo-electric effect, in which an electron is emitted with E $= h\nu - E_B$.

Figure 2.1 shows that the photo-electric effect is the most probable interaction photons of energy $\ll 0.1$ MeV for the ranges of Z naturally occurring in the human body, so is most relevant in imaging rather than EBRT.

2.2.2.1 Auger electrons

Instead of a characteristic photon being emitted, the photon can be reabsorbed by the atom, and an electron from a higher shell is emitted to carry away the excess energy from the atom. This electron is known as an Auger electron. It the electron emitted is due to transitions within sub-shells of the same shell, the electron is known as a Coster-Kronig electron.
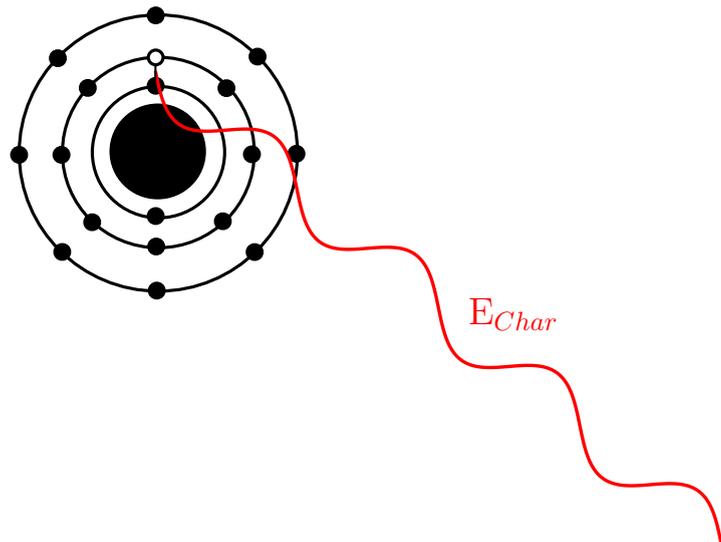
Figure 2.4: Characteristic photon emitted following photo-electric event.

2.2.3 Compton scattering

Compton (or incoherent) scattering is the process in which a photon, of energy $E = h\nu$, collides with an electron, and is deflected at an angle θ, leaving the photon with a lower energy, E'. The energy lost by the photon is converted to the kinetic energy of the electron, which is deflected at an angle ϕ (Figure 2.5).
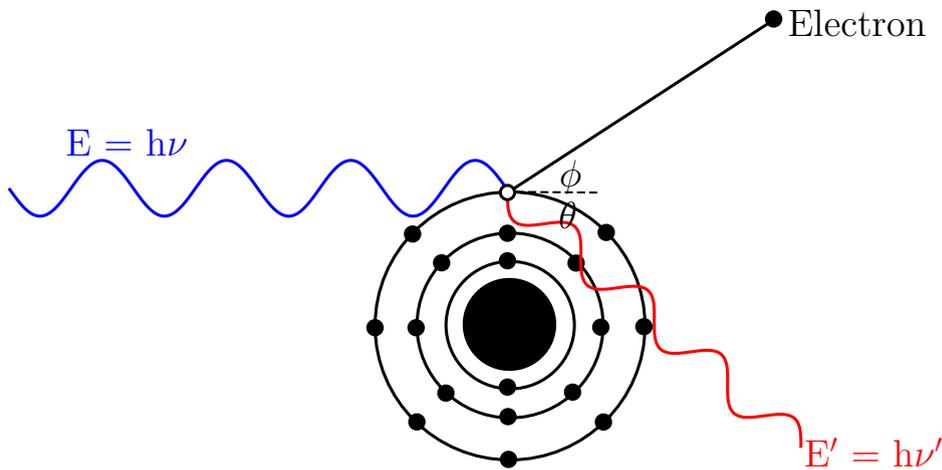
Figure 2.5: Illustration of Compton scattering.

The greatest energy transfer from the photon to the electron occurs when the photon is scattered backwards, i.e. $\theta = 180^o$, and the electron is scattered in the forwards direction, i.e. $\phi = 0^o$. Zero energy is transferred from the photon to the electron where $\theta = 0^o$ and $\phi = 90^o$, so the photon is not deflected, and the electron is scattered at right angles to the photon. Figure 2.1 shows that Compton scattering is the most probable interaction the energies encountered in EBRT, given the range of Z naturally occurring in the human body.

2.2.4 Pair production

If a photon of energy $\geqslant 1.022$ MeV, i.e. twice the rest mass of an electron, interacts with the strong Coulomb field of a nucleus, the photon may be converted into an electron-positron pair, following $E = mc^2$ (Figure 2.6). While the threshold for pair production is 1.022 MeV, Figure 2.1 shows that it is only the most probable interaction for much higher energies, particularly in higher Z materials such as titanium, so pair production is significant in hip prostheses.

The energy of the photon above 1.022 MeV is transferred into the kinetic energy

of the electron and positron, although the kinetic energy may not be distributed equally between the particles (Khan, 2003). The nucleus will recoil, and carry away some energy too, so momentum is not conserved between the photon, electron and positron. The positron will annihilate when it collides with an electron, releasing two 511 keV photons (Figure 2.7).
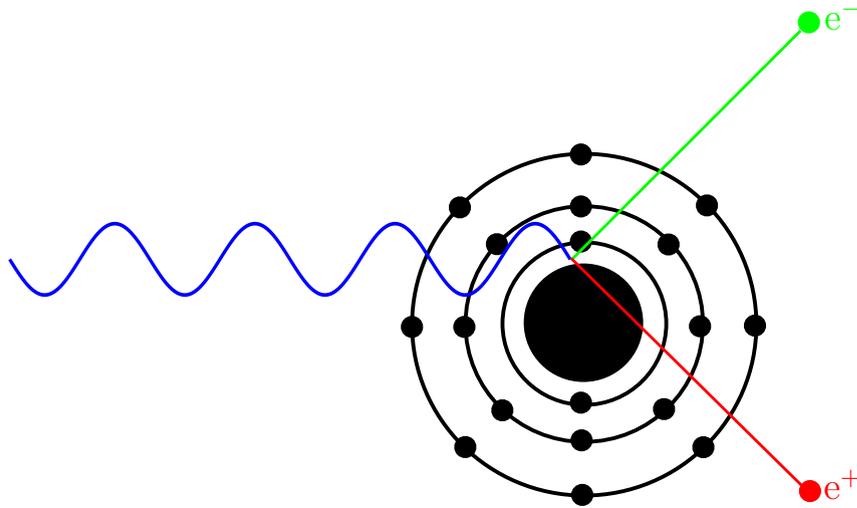
Figure 2.6: Illustration of pair production.

2.2.4.1 Triplet production

A related process to pair production is triplet production, which occurs in the Coulomb field of an electron rather than a nucleus. In triplet production, an electron-positron pair is produced, and energy is imparted to the original electron. The threshold for triplet production is twice that for pair production.
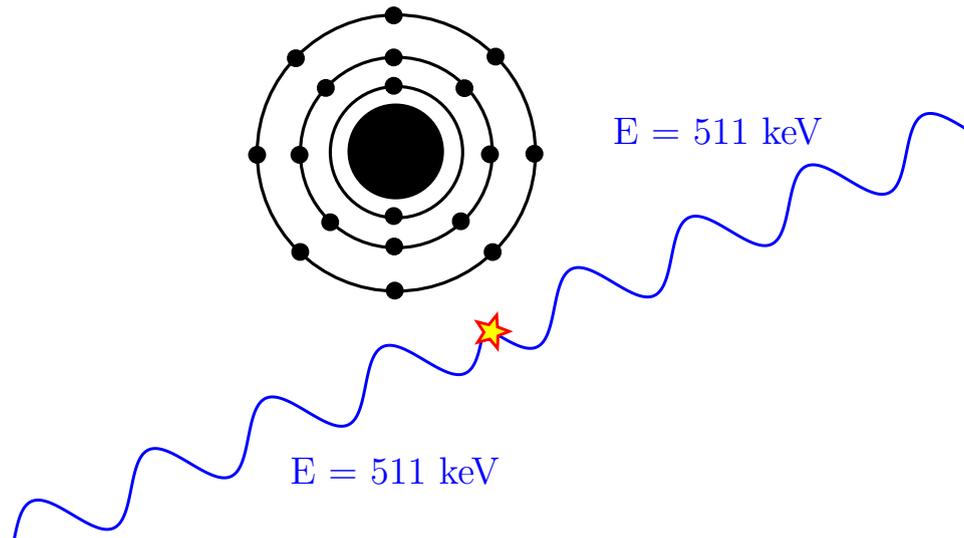
Figure 2.7: Illustration of production of annihilation photons, arising from the positron created during a pair production event.

2.2.5 Charged particle interactions with matter

Charged particles such as electrons interact with matter via ionisation and excitation, which are mediated by the Coulomb force between the electrons and the bound electrons or nuclei of atoms making up the material (Khan, 2003). For electrons in water or soft tissue, the collisions may be:

- inelastic collisions with bound electrons (ionisation and excitation);

- inelastic collisions with nuclei (which leads to Bremsstrahlung);

- elastic collisions with bound electrons;

- elastic collisions with nuclei.

For elastic collisions, no kinetic energy is lost (but may be re-distributed between particles). As electrons move through a material, they continually lose energy (which is deposited in the medium) until they reach thermal energies, and are captured by the atoms of the material. In higher Z materials, such as lead, electrons may interact via the Bremsstrahlung process, giving rise to new photons.

2.3 Characteristics of radiotherapy beams
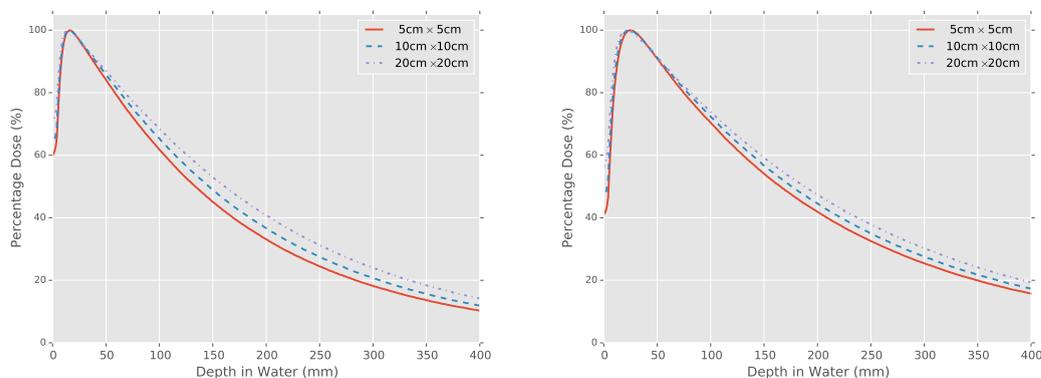
(a) Percentage depth dose curves at 6 MV. **(b)** Percentage depth dose curves at 10 MV.

Figure 2.8: Percentage Depth Dose curves for Varian 2100CD at Velindre Cancer Centre.

Figures 2.8(a) and (b) show the variation of absorbed dose, normalised to the maximum value and expressed as a percentage, as a function of depth along the central axis of a photon beam from a Varian 2100CD linac at Velindre Cancer Centre. These Percentage Depth Dose (PDD) curves were acquired during linac commissioning, a process that involves taking measurements of the photon beams generated by the linacs, to allow the dose to be calculated by TPSs. Absorbed dose is defined as the energy absorbed per unit mass in a material, in this case water. The interactions between photons and materials outlined in sections 2.2.2,

2.2.3 and 2.2.4 describe the photon imparting energy to the medium (often the electrons of the medium). Considering PE, energy is removed from the photon beam when the electron is ejected from the atom, but the energy is only absorbed by the media when the electron interacts with other atoms, depositing energy. This leads to two related but distinct concepts:

1. kerma is the kinetic energy *released* per unit mass, and relates to energy leaving the photon beam due to interactions that set charged particles in motion;

2. absorbed dose is the energy *absorbed* per unit mass, and related to energy being deposited into tissue by electrons

Kerma may be separated into two parts, K^{Col} due to inelastic collisions (i.e. ionisation and excitations); and K^{Rad} due to radiative transfers.

Figures 2.8 (a) and (b) show that the absorbed dose near the surface is relatively low (60% for 6 MV and approximately 40% for 10 MV), before climbing to a maximum dose at a depth denoted D_{Max}. This is an effect called build-up, and is due to the electrons that are set in motion near the surface not being absorbed until they have travelled a finite distance in the water. At D_{Max}, the energy imparted to electrons by the photon beam is balanced by the energy absorbed by the medium from the electrons, a condition known as Charged Particle Equilibrium (CPE). Once CPE has been established, absorbed dose and K^{Col} are equal to each other. Strictly, this is not possible due to the attenuation of the beam, so the absorbed dose is greater than K^{Col}, this is termed transient CPE. In situations where (transient) CPE is not present, i.e. at surfaces, or interfaces between two materials of different densities, K^{Col} and absorbed dose are not equal; it is often harder to calculate absorbed dose in such regions.

After D_{Max}, the PDD curves shown in Figures 2.8 (a) and (b) fall, due to attenuation by the water phantom, and the effect of the inverse square law. As

the radiotherapy beams are poly-energetic, the lower energy components will be preferentially attenuated, and the mean energy of the beam will shift towards higher energy with increasing depth. It can also be seen that 10 MV beams are more penetrating than 6 MV beams are, due to their higher energy (Figure 2.8 (b) shows a higher value at 400 mm deep than Figure 2.8 (a) does). Also, the larger field sizes show higher PDD values at depth than the smaller field sizes, due to radiation scattering from regions off the central axis onto the central axis at depth, and increases in head scatter and backscatter with increasing field size.

Figures 2.9 (a-d) show 'profiles', plots of the variation in absorbed dose across the width of a 20 cm by 20 cm field, for 6 MV and 10 MV, either normalised at each depth separately, or normalised to a common dose. It can be seen that at 5 cm deep, the beams have pronounced 'horns', i.e. the profile has a local minimum in the centre of the beam, and the profile has higher readings off axis. For the 10 cm deep profiles, this effect is much smaller, and the beam is almost flat in the central region. At 20 cm deep, the profile has pronounced 'shoulders', with the measurement falling off away from the central axis. This effect is due to the production of photons via Bremsstrahlung and the flattening filter. The beam leaving the target is much more intense on the central axis than off axis. The flattening filter attempts to remove this effect by preferentially absorbing photons on the central axis. As the photon beam is poly-energetic, lower energy photons are absorbed preferentially in the flattening filter, and this effect is more pronounced on the central axis where the filter is thickest. Therefore the condition that the beam is flat can only be satisfied at one depth, usually chosen to be 10 cm.

At approximately ± 10 cm laterally, there is a sharp fall off in all profiles, i.e. the beam edge, due to the position of the collimators. It can be seen that the beam falls to approximately 5% over a few millimetres. This is partly due to the finite source size, so that the edge of the collimators is blurred out, and partly due to lateral scatter blurring the edge of the beam. This region is known as the

penumbra of the beam. The region outside the beam is called the umbra, and dose here is due to the small portion of radiation passing through collimators, scatter radiation from the directly irradiated portion of the phantom, and indirect head scatter.
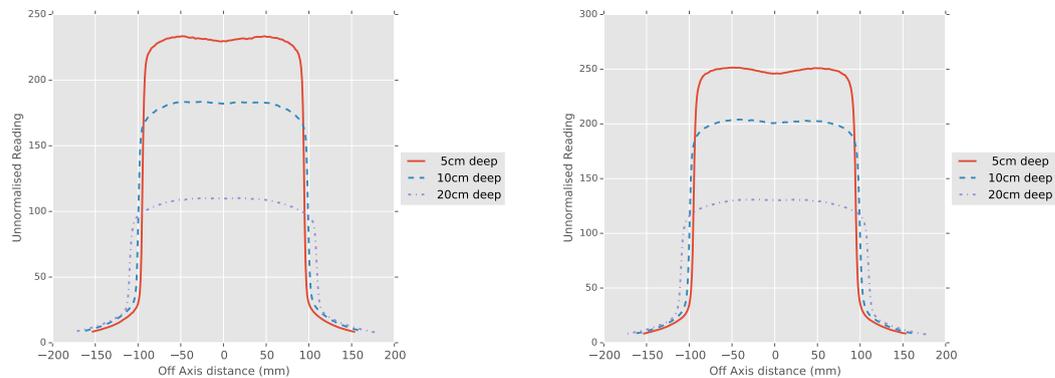
(a) Un-normalised beam profiles at 6 MV.

(b) Un-normalised beam profiles at 10 MV.
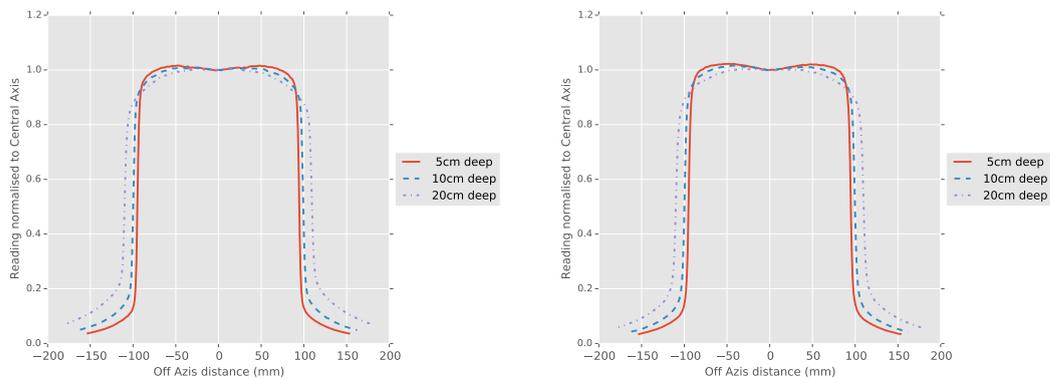
(c) Normalised beam profiles at 6 MV.

(d) Normalised beam profiles at 10 MV.

Figure 2.9: Beam profiles for Varian 2100CD at Velindre Cancer Centre.

2.4 Analytical dose calculation methods

2.4.1 Factor based dose calculations

The dose at a point on the central axis of radiotherapy beam can be calculated by Equation 2.2, where PDD is the PDD at depth d, MU is the number of MU delivered, K is the calibration factor, typically 1 Gy per 100 MU, $S_C(r_C)$ is the collimator scatter factor of a field size r_C defined at the source to axis distance, $S_P(r)$ is the phantom scatter factor of a field size for r defined at the patient surface, SCD is the source to calibration distance, SSD is the source to surface distance, and t_0 is the reference depth (Khan, 2003).

$$Dose = PDD(d) \times MU \times K \times S_C(r_C) \times S_P(r) \times \left(\frac{SCD}{SSD + t_0} \right)^2 \qquad (2.2)$$

Equation 2.2 does not account for the presence of wedges, heterogeneities, or allow for the calculation of dose off axis. The dose on the central axis in the presence of a wedge (hard or dynamic) may be accounted for using a Wedge Factor (WF). The WF is defined as the ratio of doses at a reference depth with and without a wedge in the beam; the reference depth is often taken to 10 cm deep. The WF depends on energy of the beam, angle of wedge, and for dynamic beams the field size. Calculating the dose off-axis may be achieved using an Off Axis Factor (OAF), which is the ratio of absorbed dose at a point away from the beam axis, to the absorbed dose at a point on the beam axis at the same depth. The OAF depends on energy of the beam, depth, and the field size. Accounting for heterogeneities is more complex, but may be achieved using a Power Law approach, as outlined in section 2.4.2.

In modern radiotherapy, factor based dose calculations are primarily used to check more complicated dose calculation methods, but illustrate that:-

- absolute dose changes with jaw size and source to surface distance;

- variation of dose with depth depends on depth, collimator sizes, and source to surface distance;

- energy and dose change with depth and off axis position;

- wedges change absolute dose, as well as the shapes of isodoses;

- scatter from the head of the linac and scatter from the phantom behave differently, and should be considered differently.

2.4.2 Power Law corrections

Consider calculating the dose in a water-like medium, in which there is an irregularly shaped region of low density material (e.g. a lung). Clearly, the dose within the low density material will be different to the dose at the same point in a homogeneous water phantom. The dose in the region 'downstream' of the radiotherapy beam will also be different to the dose in a similar homogeneous water phantom. This situation is too complex to have a series of pre-defined factors to correct for the presence of the heterogeneity. The dose will be affected by the size and shape of the heterogeneity, as well as the proximity of the dose calculation point to the heterogeneity. Johns and Cunningham (1983) describe a power law correction, based on the method of Batho, later extended by Sontag, for correcting the dose at a point in or 'downstream' of a heterogeneity. This method is encapsulated in Equation 2.3, in which:

- TAR is the Tissue Air Ratio (TAR), where the TAR at a point in a water phantom is the ratio of the total absorbed dose at that point, divided by the total absorbed dose at a point with the same distance from the source, but measured with the phantom shifted so that this point is at the depth of maximum dose (British Institute of Radiology, 1996);
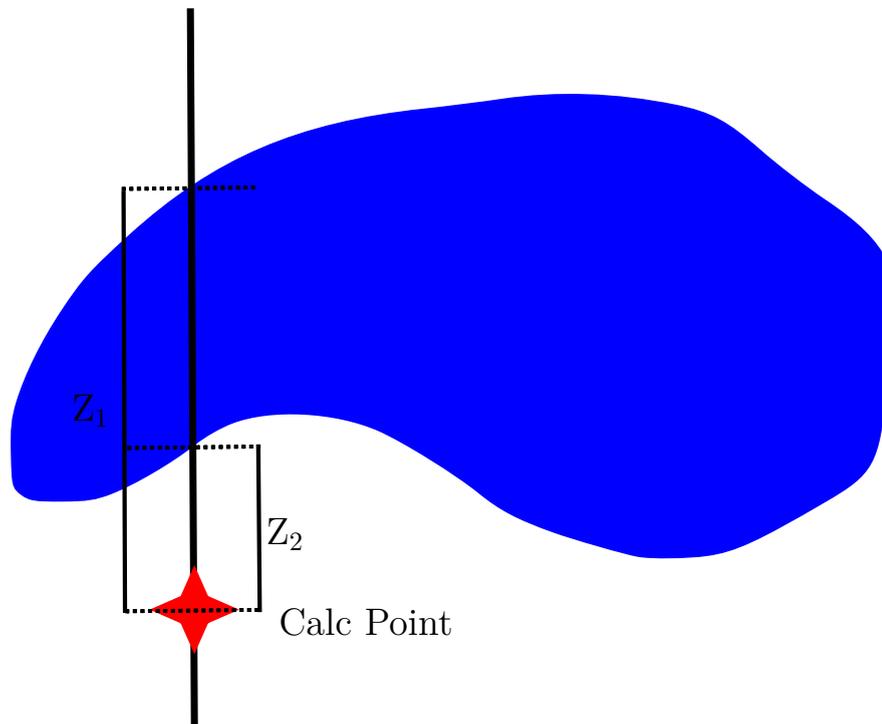
Figure 2.10: Diagram depicting an inhomogeneous region, in which a Batho type calculation may be employed. Z_1 is the distance from the point at which the beam enters the inhomogeneity to the calculation point, and Z_2 is the distance from the point at which the beam exits the inhomogeneity to the calculation point.

- z_1 is the distance along the ray line from the calculation point to the point where the beam enters the heterogeneity;

- z_2 is the distance along the ray line from the calculation point to the point where the beam exists the inhomogeneous region (Figure 2.10);

- f_d is the field size defined at a depth d;

- ρ_1 is the density of the heterogeneity;

- ρ_2 is the density of the outer material (i.e. the water phantom).

This correction still does not account for lack of CPE close to boundaries between tissues, or the shape of the heterogeneities.

$$C = \frac{TAR(z_1, f_d)^{\rho_1 - \rho_2}}{TAR(z_2, f_d)^{1 - \rho_2}} \tag{2.3}$$

2.4.3 Type A and type B algorithms

Dose calculation methods can be divided into two types, those that cannot model lateral electron scatter, termed *type A* algorithms, and those that can, termed *type B* algorithms. In those algorithms that cannot model lateral electron scatter, there will be errors in regions without Charged Particle Equilibrium, i.e. at the surface of the patient, or near interfaces between tissues of different densities. Also, the width of the penumbra in low density media, such as lungs, will be incorrectly modelled in *type A* algorithms.

Examples of *type A* include PB in Oncentra, the Pencil Beam algorithm in Eclipse and any factor based method that does not account for lateral scatter. Examples of *type B* include CC in Oncentra, and the Anisotropic Analytical Algorithm in Eclipse, which accounts for lateral scatter within a plane only, so is sometimes referred to as '2.5D', as it does not account for heterogeneities in 3D.

2.4.4 Kernel based methods

Modern TPS dose calculation algorithms use dose deposition kernels in the calculation, which are matrices or analytical functions describing the dose deposited around the point at which a photon interacts with matter. These kernels may be combined with the photon fluence via convolution or superposition methods. Convolution is a mathematical operation to describe the overlap of two functions

as one function is moved over the other. The modifying function must be invariant over the space in which the convolution takes place. In convolution-based dose calculations, the dose $D(\vec{r})$ at a point $\vec{r}$ may be given by Equation 2.4, where $\frac{\mu}{\rho}$ is the mass attenuation co-efficient, $\Psi_p(\vec{r'})$ is the primary photon energy fluence at a point $\vec{r'}$, $A(\vec{r}-\vec{r'})$ is the convolution kernel. This convolution kernel is a matrix or function describing the dose distribution deposited by scatter photons and electrons set in motion at the primary interaction site. This may be expressed as a function of total energy released per unit mass (terma), $T_p(\vec{r'})$, a concept analogous to kerma, but which includes all energy released, not just kinetic energy, i.e. radiative energy is included. In practice, Fast Fourier transforms may be used to perform convolution calculations quickly.

$$D(\vec{r}) = \int \frac{\mu}{\rho} \Psi_p(\vec{r'}) A(\vec{r} - \vec{r'}) d^3\vec{r'} \qquad (2.4)$$

$$D(\vec{r}) = \int T_p(\vec{r'}) A(\vec{r} - \vec{r'}) d^3\vec{r'} \qquad (2.5)$$

Convolution is only valid if the modifying function (the kernel, in dose calculations) is spatially invariant, so convolution is not always applicable. If the kernel varies with depth, or off axis distance, superposition must be used in place of convolution. In superposition, the function describing the incoming radiation and the function describing the spread of dose must be evaluated at each point, then the resulting doses summed together, or super-imposed - hence the name. Superposition is mathematically much less efficient than convolution, but convolution-based methods cannot take account of variations in dose deposition behaviour that superposition methods can.

2.4.5 Oncentra MasterPlan algorithms

Oncentra has two algorithms for photon dose calculation, the PB algorithm, which uses convolution of pencil, point and planar beam kernels, and the CC algorithm that uses superposition of point kernels. The PB algorithm does not account for the lateral transport of secondary particles in heterogeneous materials, and is therefore a *type A* algorithm. CC does model this lateral transport, so is a *type B* algorithm.

The algorithms are designed to be general enough to handle all linac beam geometries, i.e. regular fields shaped by collimators, irregular fields defined by MLCs, the presence or absence of wedges (either dynamic or physical wedges) and modulated or dynamic fields such as Intensity Modulated Radiotherapy (IMRT) and Volumetric Modulated Arc Radiotherapy (VMAT).

Both algorithms calculate the absolute dose per MU, using the dose per MU under reference conditions. The calculation of dose per MU takes account of the signal scattered backwards from the collimators into the MU chamber, which reduces the dose per MU for small field sizes, above that which would be expected due to the reduction in phantom scatter. Each algorithm can be divided into two steps: the initial step involves calculating the energy fluence exiting the head of the linac; and the 'dose engine', which calculates the dose to the patient. Both PB and CC use the same energy fluence modelling. What distinguishes the PB and CC algorithms is the dose calculation engine they use.

2.4.5.1 Energy fluence modelling

The energy fluence is the amount of radiant energy per unit area traversing a plane perpendicular to the beam, at a given position, and is a quantity that is independent of the patient or phantom. Energy fluence is determined wholly by the linac configuration, including beam energy, collimator and MLC positions and

presence of wedges. The energy fluence model is produced from PDD, profiles and Field Size Factor (FSF) measurements taken for different energies and collimator sizes, and knowledge of the linac head design for different manufacturers. Parameters in the general model are adjusted until the PDD, profiles and FSF calculated by the model agree with the measured values. The energy fluence from the linac includes contributions from:

- **direct photons**, which do not undergo scattering in the linac head;

- **photons scattered from the flattening filter and primary collimator**, which are generally the main sources of scattered energy fluence;

- **photons scattered from the collimators and MLCs**, which varies with collimator and MLC positions;

- **photons scattered from physical wedges** and any auxiliary modulators;

- **contaminant photons and electrons**, which are low energy photons and electrons that affect the dose near the surface of the patient or phantom.

Each of these contributions follows a different form, and are calculated in different ways. The direct fluence is obtained by modulating the fluence for an 'open' (un-collimated) beam by the attenuation from the collimators and MLCs. The calculation of the photons scattered from the flattening filter and primary collimator involves a two step process. Firstly, these components are divided into a grid, to determine which elements of this grid are 'visible' to each calculation point, using a 'virtual light source' approach (visible meaning not hidden by collimators or MLCs). Secondly, an integration is performed of the energy fluence over these visible elements. The collimator scatter is not calculated from scratch for each collimator and MLC position; rather the scatter kernels are calculated for a standard geometry, and then shifted and scaled based on the actual beam defining geometry.

2.4.5.2   Dose calculation engines

Oncentra represents the patient (or phantom) geometry as a 3D density matrix. The density of each voxel in the matrix is determined either from the pixel values of the CT scan, or from user specified values of mass or electron density. The resolution of this density matrix can be varied; the smaller the voxel size, the more accurate the calculation, but the calculation time increases linearly with the number of voxels. (In practice, the voxel size in one direction is set by the CT slice thickness; within a plane, it is common to use the same resolution in both directions, so the calculation time varies as the inverse of the square of the voxel side).

A dose engine is an algorithm that calculates the absorbed dose in this density matrix from the incident energy fluence, calculated in the first step. The Oncentra photon dose engines are based on energy deposition kernels, which describe the deposition of dose around a elementary radiation beam. Three different types of kernels are used. The first kind are point kernels, relating to the energy deposition around a point interaction. The second kind are pencil kernels, relating to the energy deposition from a mono-directional 'beamlet', known as a pencil beam (hence the name of the algorithm). The third kind are planar kernels, relating to the energy deposition from a broad parallel beam.

Primary dose is defined as the energy imparted per unit mass due to photons interacting with the medium for the first time. Head scatter dose is the energy imparted per unit mass due to photons that have been scattered in the head of the linac (i.e. in the flattening filter, collimators, wedges etc.). Phantom scatter dose is defined as the energy imparted per unit mass due to photons that have interacted in the medium more than once, as well as Bremsstrahlung and annihilation photons from pair production.

There are two types of kernels used in this calculation. Firstly, there will be

kernels to describe the deposition of energy from the energy fluence from the linac head - this leads to primary and head scatter dose. There will also be kernels to describe the deposition of energy from energy fluence that has been scattered in the patient or phantom - this leads to phantom scattered dose.

2.4.5.3 Pencil Beam dose engine

The PB (Ahnesjö et al., 1992) dose engine performs a 'pre-convolution' step, in which point kernels are integrated over the depth co-ordinates, to produce pencil beam kernels. This saves computation time, but precludes any way to include density scaling for lateral heterogeneities, as the pencil kernels are then fixed, and cannot be adapted for the presence of lateral heterogeneities.

The variation of energy with depth is accounted for by using two mono-energetic pencil kernels, which are combined with a different weighting at different depths. This produces a poly-energetic pencil kernel, which is integrated with the incident energy fluence. This creates a dose distribution that reproduces the variation of beam profile with depth described in section 2.3, without the need to vary the point kernels with depth, which would lead to a loss in calculation efficiency.

Contaminant photons and charged particles are produced as the direct photons pass through the air between the linac and the patient, and any auxiliary beam modulators near the patient. This is modelled using a Gaussian pencil beam kernel. The contamination dose falls rapidly with depth.

The effect of inhomogeneities along a pencil beam for primary dose and contaminant dose is accounted for by scaling the pencil beam kernels by the radiological depth. The radiological depth is defined as the depth of water that would give the same attenuation as the path through the inhomogeneous region.

Ahnesjö et al. (1992) describe a correction factor that is applied to the phantom scatter to account for heterogeneities, which is a function of an 'effective attenu-

ation coefficient'. From the definition of this correction factor, it follows that the final dose is relatively insensitive to the actual value of this effective attenuation coefficient. A value of 0.8 is used in the Pencil Beam Enhanced (PBE) algorithm. Using a constant value enables the scatter dose to be corrected for heterogeneities using a 1D convolution with a planar kernel performed along the beam path. A planar kernel is used as the kernel does not vary across the width of the beam. It is noteworthy that the PBE uses point, pencil-beam and planar kernels at different stages of its calculation of dose. It is not a simplistic algorithm as it is often described; rather it is an algorithm that applies approximations in order to allow efficient calculation methods to be exploited.

Another consequence of the production of pencil beam kernels via the 'pre-convolution' step is that the absorbed dose in a voxel is calculated as if the voxel were water (even if it is not). This is because the pencil beam kernels describe dose deposition in water, and this cannot be altered on a voxel by voxel basis if convolution is to be used.

2.4.5.4 Collapsed Cone dose engine

The CC algorithm (Ahnesjö, 1989) uses the same inputs to the beam model as PB. Older versions of Oncentra limited the resolution that CC was performed at to decrease execution time. However, advances in computing, particularly implementation of the CC algorithm on the Graphics Processing Unit (GPU), allows calculations to be run faster than the PB algorithm on the Central Processing Unit (CPU) for all but the simplest calculations. For very simple cases the losses in the overhead involved in transferring data to and from the GPU outweigh the speed advantages of the GPU itself.

The CC algorithm uses pre-calculated point kernels, which describe the deposition of energy from a photon interaction site as a function of direction and distance.

As with the PB algorithm, a 3D density matrix is used to describe the patient or phantom geometry. Raytracing is used to calculate the amount of radiant energy released in each voxel from the incident beam; considering the direct radiation and head scattered radiation separately. The variations in beam energy along the beam axis, and across the beam profile, are accounted for during the raytracing. There are two sets of point kernels used, one to calculate primary dose from direct and head scattered energy fluence; and one to calculate phantom scattered dose from the phantom scattered energy fluence. The point kernels for the primary dose are tilted such that the axis of the kernel is aligned to the beam divergence angle.

In a homogeneous, water equivalent medium the dose deposition will be the radiant energy released at each point, convolved with the point kernel at each point. However, in order to account for inhomogeneities, the dose deposition kernel is divided into a number of cones, with the cones being closely packed in the direction of higher energy deposition, i.e. along the beam axis. In the initial paper describing the CC algorithm (Ahnesjö, 1989) between 48 and 384 cones were used for different cases, but 27 was the preferred number; the Collapsed Cone Enhanced (CCE) as implemented in Oncentra uses 106. The energy deposited within a cone is assumed to be deposited along the axis of the cone, and the energy deposition is scaled according to the density of the media along this axis. This allows for heterogeneities lateral to the beam axis to be accounted for. Since the energy deposition is performed for the voxel medium, the dose calculated by the CC algorithm is the dose to the medium, rather than the dose to water as reported by the PB algorithm. The name of the algorithm comes from the assumption that the dose deposited in a cone can be 'collapsed' onto the axis of the cone, as an approximate representation of actual dose deposition within that cone.

The dose from contaminant photons is calculated separately, using the same Gaussian pencil beam based method as in the PB algorithm.

2.4.6 Enhanced algorithms in Oncentra

In Oncentra v3.1, 'enhanced' versions of the PB and CC algorithms were introduced. These algorithms include changes to the energy fluence modelling. The fluence grid is aligned with the beam collimation system, and a higher resolution grid is used to describe the fluence than in the non-enhanced versions of the algorithms.

For multi-segment beams, such as IMRT beams, the energy fluence for all segments is calculated, and the dose engine performs the dose calculation once for that beam, not once per segment as for the non-enhanced versions of the algorithms.

The enhanced PB algorithm includes further increases in speed due to use of an interpolation scheme. The energy fluence is convolved with primary and scatter kernels; this is done at two depths for the primary dose, and three depths for the scattered dose. The depths are determined during the calculation, as a function of beam spectrum. The calculation point in the patient is transformed onto a water phantom, where the point has the same radiological depth and lateral position. Lateral interpolation is performed in kernel space to calculate the dose deposition throughout the patient; i.e. the kernels required at other depths are interpolated, rather than the dose itself being interpolated.

The enhanced versions of the PB and CC offer better calculation accuracy that the standard versions, due to the higher resolution grid used for the energy fluence, and its alignment with the beam collimation system. Efficiencies in the way the fluence matrix is used allow for faster calculations (for same beam and patient geometry, on the same computer), particularly for IMRT beams.

Throughout this work the enhanced versions of the algorithms from Oncentra v4.3 have been used, and will be termed PBE and CCE.

2.5   Monte Carlo methods

The MC method solves equations or calculates integrals using random number sampling (Seco and Verhaegen, 2013). Early examples include solutions to numerical problems such as a method to calculate the value of π. However, MC is perhaps best suited to problems such as those in particle and nuclear physics in which the random nature of fundamental interactions can be accurately represented by MC methods. Point kernels describing the spread of energy around the point of a primary photon interaction, as used in the CCE algorithm were calculated from MC simulations (Mackie et al., 1988). Simulating radiation transport using MC methods is, in principle, simple; the microscopic interactions are simulated, using the interaction cross-sections for the radiation energy and type, and the material under investigation. Consider pair production, discussed in section 2.2.4: a photon passing near a nucleus may or may not lead to a pair production process, with a probability that is dependent on the Z of the nucleus. If pair production does occur, the angles at which the electron and positron travel at will be random, as will be the share of the energy carried away by each particle, and the nucleus. MC and a knowledge of the relevant interaction cross-sections should allow radiation transport to be simulated. MC is conceptually simpler than, for example, the PB algorithm with its use of point, pencil beam and planar beam kernels at different stages of the calculation.

However, full simulation of all interactions is prohibitively time consuming. For the energies associated with radiotherapy beams, electrons undergo a large number of interactions, making the simulation of one electron track much more computationally expensive that simulating a photon track, to the extent that simulating electron tracks is not feasible. While MC for particle transport was proposed by Stanisław Ulam (Metropolis and Ulam, 1949), it was the work of Martin Berger in Alder et al. (1963) that led to MC for photon and electron transport. Due

to the small amount of energy transferred from an electron to the tissue and the small scattering angles that the electron is deflected through, an approach known as the Condensed History (CH) technique was proposed by Berger. Interactions are divided into 'hard' and 'soft' collisions, by use of a parameter, E_C. Collision events with an energy transfer greater than E_C are hard collisions, and are simulated explicitly, as for photons; collisions with an energy transfer less than E_C are termed soft collisions. Soft collisions are modelled by assuming a continuous energy transfer into the surrounding medium, and a single large scattering angle rather than a series of small scattering angles. If this large angle deflection is applied only at the start, the distance travelled by the electron may be overestimated. Using a second scattering angle, known as a 'random hinge' can be used to correct for this effect (Figure 2.11).

The parameter E_C can be selected by the user, to decrease the time required for a MC simulation; but choosing too high a value may lead to inaccurate results.
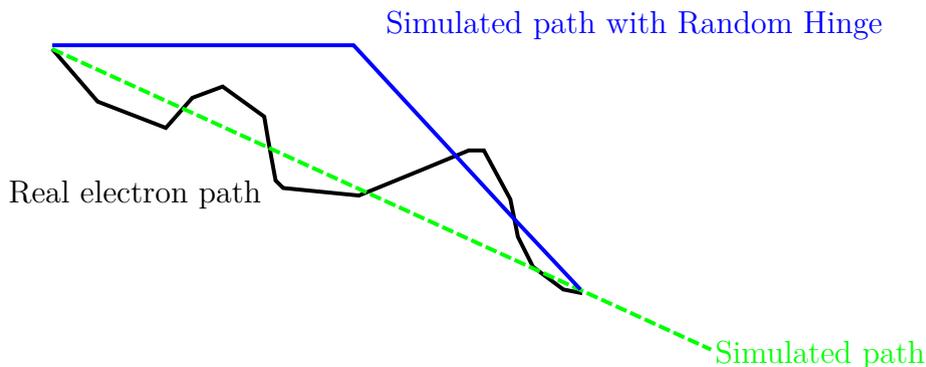
Figure 2.11: Use of the random hinge method, to avoid over-estimation of particle range.

The continuous energy loss and random hinge approach for soft interactions has limitations. For example, all electrons with an initial energy E modelled using a step length of s will reach the end of the step with the same energy, whereas in reality the electrons should have a spread of energies. Many implementations of the random hinge angle use Fermi-Eyges multiple scattering theory (Eyges, 1948)

to compute a distribution of scattering angles to sample from. However, this theory uses a 2D Gaussian distribution that underestimates the probability of large scattering angles. Another issue is that the random hinge method assumes that all soft scattering occurs within the same medium; therefore the CH technique can only be applied if a single step will not lead to the particle crossing a boundary into another medium. Commonly, the step size is decreased near boundaries such that it is much less than the distance to a boundary, but this will increase calculation times.

Variance Reduction techniques are techniques to reduce the calculation time required to reach the same statistical uncertainty (or increase efficiency) without altering the result produced (Verhaegen and Seuntjens, 2003). Common methods include:

- **Transport Cut-offs:** An energy below which electrons or photons will not be transported, but will be assumed to deposit all energy where they are. Due to the higher range of photons compared with electrons, photon cut-offs are usually lower than electron cut-offs. Since cut-offs can introduce bias, they are not strictly a variance reduction technique;

- **Bremsstrahlung splitting and Russian Roulette:** For particles of particular interest, such as Bremsstrahlung photons in a target, for every Bremsstrahlung event many Bremsstrahlung photons will be modelled, each with a reduced weight. Russian Roulette involves removing a proportion of secondary charged particles, due to the large numbers produced.

- **Electron Range Rejection:** For electrons that do not have high enough energy to reach a plane of interest, the electrons are terminated, for example electrons with insufficient energy to escape the linac head. As electrons may give rise to Bremsstrahlung photons, a energy threshold must be set, above which no rejection is performed. As with cut-offs, this method can introduce bias, so is strictly not a variance reduction technique;

- **Interaction Forcing:** Interactions that are relatively uncommon, but give rise to particles of interest, such as contaminant electrons near patient/phantom surfaces, may be under-sampled even in long simulations. If the interactions are forced to occur, but the resulting particles have lower weight, the statistics in these situations can be improved.

2.5.1 Monte Carlo codes

2.5.1.1 ETRAN, ITS and MCNP

The electron gamma modelling code developed by Berger and others was released as a package called ETRAN. This underwent many developments, before becoming the ITS code. The ITS electron transport code now lives on in MCNP (Brown, 2003), maintained by group at Los Alamos National Laboratory. In their review of MC codes for radiotherapy, Spezi and Lewis (2008) found that MCNP was, at the time of publication, the most commonly cited MC code in the field of nuclear science and technology, and in all fields combined. This was due to its ability to model neutrons, and photo-nuclear reactions. It is also the second most commonly used code in the field of radiology, nuclear medicine and medical imaging. However, it is not as fast as other codes, due in part to its ability to model so many situations.

2.5.1.2 EGS, BEAM and EGSnrc

Spezi and Lewis (2008) reported that the electron-gamma-shower (EGS) code, or MC packages based on it, were used in more papers in the subject areas of radiology, nuclear medicine and medical imaging between 1985 and 2007 that any other MC code, with approximately 500 papers in this period. EGS can trace its history back to Oak Ridge National Laboratory in the 1960s. Early

versions of what would become EGS were developed for shielding design, and to assist with the interpretation of experiments at Stanford Linear Accelerator Centre. Since the 1980s development of EGS has focussed on medical physics applications, starting with improvements in transport in the energy range 10 keV to 30 MeV, which were included in EGS4/PRESTA.

Rogers (2006) highlights the difficulties of modelling the response of a parallel plate chamber accurately. Kawrakow and Rogers (2000) used a version of EGS, EGSnrc, which included improvements to particle transport in order to resolve some issues seen when using EGS4/PRESTA for this purpose. EGSnrc included improvements in the CH technique, including a new electron step algorithm, often called PRESTA-II, and an exact boundary crossing algorithm. It also included improvements in physical processes, such as improvements to the electron elastic scattering mechanism to include relativistic spin effects, and inclusion of National Institute of Standards and Technology (NIST) Bremsstrahlung cross sections as recommended by the International Commission on Radiation Units and Measurement (ICRU) (ICRU, 1984).

The BEAM code was produced by the Ottawa Madison Electron Gamma Algorithm (OMEGA) project, which was set up in response to many users of EGS wanting to apply EGS to linac modelling (Rogers et al., 1995). It is a tool to allow clinical physicists to perform MC simulations with EGS, by being more accessible than EGS. Although it was possible to specify the components of a linac head in EGS, due to its flexibility, it was not easy to do so. BEAM is written in MORTRAN3, a preprocessor for Fortran77 (i.e. MORTRAN3 macros can be used to generate Fortran77 code). BEAM uses Component Modules (CMs) to describe each component of the linac head as an independent 'building block', in which the geometry and composition of the part of the linac is described. For example the CM SLABS can be used to model any rectangular object in the linac head, by specifying the material and certain co-ordinates of the slab to be modelled.

The boundaries and materials required by the EGS code will be generated from the parameters supplied to the CM. BEAM allows generation of 'phase space files', which are files containing the type, energy, position and direction of travel of particles at a plane, at any level of the linac. A simulation may be performed of the radiation transport through the 'patient invariant' portion of the linac, i.e. the portion excluding wedges, jaws and MLCs, which does not change for a given energy, to generate a phase space file above the 'patient specific' portion of the head, i.e. the wedges, jaws and MLCs. This phase space file can then be used as an input to many different simulations of the patient specific portion of the head, which avoids repeatedly simulating invariant situations. The BEAM code was also well documented to help users understand the system. The BEAM code could be extended by users who wanted to add new CMs (such as DYNVMLC, used to model dynamic MLC motion for the Varian Millennium 120-leaf MLC, by Heath and Seuntjens (2003)). BEAM also provided a number of source routines, such as a simple isotropic point source, simple parallel beams, elliptical beams to represent an electron beam incident on the target in a typical linac, and more complex sources. Again, users have written their own sources for specific purposes, some of which have been incorporated into the BEAM distribution (Lobo and Popescu, 2010).

2.5.1.3 PENELOPE

PENetration and Energy LOss of Positrons and Electrons (PENELOPE) is a general purpose MC code for modelling electron-photon interactions in complex geometries (Baró, 1995). It allows regions to be specified by quadric surfaces (i.e. surfaces described by second order equations), and provides a 'visual debugger' to visualise the surfaces and resolve issues with the geometry specified. It is well suited to simulation of low energy particles, due to detailed cross-sections for low energy interactions.

PENLINAC (Rodríguez, 2008) is an extension to PENELOPE, which provides a simple way for users to define the geometry of a linac, in much the same way as BEAMnrc does for EGSnrc. PENLINAC does not require the user to interact with the quadratic surfaces directly, and can be considered as a layer of abstraction between the powerful but complex geometrical modelling of PENELOPE and the user.

PRIMO (Rodríguez et al., 2013) is a tool-kit built on top of PENELOPE, which simplifies the simulation process for new users. PRIMO is unique in that is includes details of Varian and Elekta linac geometry in its code. The advantage of this approach is that once a user has the code, a realistic simulation can be run almost straight away with no configuration required. The downside is that the user cannot add linacs, components and techniques not yet included in PRIMO, but must wait for the authors to update PRIMO, whereas uses of the underlying PENELOPE system could configure new geometries themselves (with sufficient time, effort and expertise). PRIMO has some variance reduction techniques tuned for simulating linac based treatment plans. PRIMO also has a number of graphical tools for the evaluation of doses calculated from MC, such as isodoses and DVH plots.

2.5.1.4 GÉANT

GÉANT was developed at CERN for particle physics applications (Agostinelli et al., 2003). It is comprised of a series of C^{++}[1] toolkit libraries, containing functions for simulating different processes. All the MC codes discussed above use Fortran[2]. GÉANT4 is able to include the effects of external electromagnetic fields, so is used in proton beam radiotherapy. Earlier versions implemented the hadron

---

[1]C^{++}. Available at: http://www.open-std.org/jtc1/sc22/wg21/ (Accessed: 8 December 2016).

[2]Fortran. Available at: http://www.nag.co.uk/sc22wg5/index.html (Accessed: 8 December 2016).

event generator from Fluka (section 2.5.1.5). Along with MCNP, GÉANT4 is one of the few codes to include simulation of photo-nuclear reactions. It is very flexible, but this flexibility means it is slower than EGSnrc unless much effort is devoted to adapt it for the energy ranges and materials of interest in medical physics. Seco and Verhaegen (2013) state that the number of publications using GÉANT for medical applications is approaching the number of publications using MCNP for medical applications (note MNCP is used extensively in non-medical applications).

2.5.1.5 Fluka

FLUKA is another general purpose code from CERN (Ferrari et al., 2005). As with GÉANT4, it has the ability to include the effects of external electromagnetic fields on charged particle transport which makes it suitable for proton beam radiotherapy, as well as photon or electron beam radiotherapy. This ability will also be advantageous with advent of MRI-linacs. FLUKA can transport photons and electrons over an energy range of 10^{12}, 1 keV to 1 peta-electron volts (PeV). It has detailed cross-section data at very low and very high energies, and is well suited to shielding calculations for linacs and proton accelerators, due to ability to model radionuclide production (Fassó et al., 2000).

2.5.1.6 Accelerated codes

There are a number of 'accelerated' MC codes, which aim to reduce computation time while not significantly altering the doses calculated. They achieve this by employing approximations and simplifications for processes that are less important for energies and materials relevant to radiotherapy. One of the best known of these is the Voxel Monte Carlo (VMC) program by Kawrakow et al. (1996), which was written for electron dose calculations only. This targeted the energy

range 1-30 MeV, and densities in the range 0-3 g/cm^3, and sacrificed accuracy outside these ranges for efficiency savings within them. Further efficiencies were found by only supporting dose deposition in rectangular grid geometries - as found in CT scans. This enabled a decrease in calculation time of 35 compared with EGS4. VMC is available in commercial treatment planning systems, e.g. it is implemented in Oncentra as the eMC dose calculation method.

The VMC code was extended to include photon dose calculations, in a code called XVMC (Fippel, 1999). This demonstrated a increase in calculation speed of 15-20 compared to EGS4, without significant loss in accuracy except in lung substitute. XVMC is available within the BrainLab iPlanNET[3] TPS.

XVMC, which was written in Fortran, was re-implemented in C^{++} by Kawrakow and Fippel (2000), as VMC^{++}. The Object Oriented nature of C^{++} allows Simultaneous Transport of Of Particle Steps (STOPS), in which particles with the same energy but different positions, directions and even statistical weights are transported simultaneously, to increase efficiency. VMC^{++} has a reported speed up compared to EGS4 of 150 times, while agreeing to within 1% with EGSnrc.

Other accelerated codes include DPM by Sempau et al. (2000), PEREGRINE by Hartmann Siantar et al. (2001) and MCDOSE by Ma et al. (2002). DPM and PEREGRINE use a method in which transport through a number of voxels can be modelled in a single step, increasing calculation speed. PEREGRINE uses a source model in which a phase space file is produced by modelling the patient invariant portion of the linac head. The distribution of energies, directions and positions is then approximated by histograms. Particles are generated by sampling data from the histograms, which produces similar data to the phase space file without having to store a large phase space file. VMC^{++} remains the fastest accelerated MC code, although MCDOSE uses extensive variance reduction techniques to achieve calculation speeds within a factor of two of VMC^{++}

---

[3]BrainLab (2013) BrainLab Physics Technical Reference Guide rev 1.7. Germany.

(Chetty et al., 2007).

The DMLC code (Keall et al., 2001; Siebers et al., 2002) is notable in that it only models the MLCs. It takes a phase space file positioned above the MLCs as it input. It then creates a phase space file beneath the MLCs, having simulated the passage of the radiation through the MLCs using a simplified transport method, but which is capable of accounting for the dynamic motion of the MLCs. This approach has been superseded by BEAMnrc CMs that are capable of including MLC motion, e.g. DYNVARMLC, or source models capable of accounting for motion in the linac head (Lobo and Popescu, 2010).

2.5.2 High Performance or High Throughput Computing

Due to the computationally demanding nature of MC, many workers have used High Performance Computing (HPC) or High Throughput Computing (HTC) to reduce the execution time for MC simulations. HPC is the field devoted to executing tightly coupled computing tasks, often requiring a high number of Floating Point Operations Per Second (FLOPS). Examples of HPC hardware include supercomputers and clusters. Supercomputers are computers with a high level of computational capacity compared to a consumer computer. They are frequently large, expensive, and have a large number of processors, which generally share memory and a filesystem (e.g. Sunway TaihuLight[4], currently the record holder for the number of FLOPS, has 40,960 64-bit RISC processors and cost over \$270M). Computer clusters are sets of computers, which may be loosely or tightly coupled, which are used together and may be viewed as a single system. An example of a type of cluster is a Beowulf cluster[5], which is a computer cluster built from standard, consumer-grade computers that share a filesystem, and act

---

[4]Sunway TaihuLight. Available at: http://demo.wxmax.cn/wxc/soft1.php?word=soft&i=46 (Accessed: 9 December 2016).

[5]Beowulf Clusters. Available at: http://www.beowulf.org/ (Accessed: 9 December 2016).

as a single system in which processing tasks are shared amongst the constituent computers.

In contrast, HTC systems are not designed for tightly coupled problems, and may not use hardware that is faster than consumer hardware. In HTC the key is performing many calculations at once, often utilising many computers, which may or may not be distributed over a large area. Examples include the Folding@home[6] and SETI@home[7] projects, which both involve volunteers downloading analysis programs and data to their computers; the analysis program is configured to run while the computer is idle. Once a batch of data has been analysed, the results are uploaded for the researchers to access. The 'virtual supercomputer' formed by participant computers in the SETI@home project holds the record for the largest number of *total* floating point operations in a computation (10^{21}). This is despite Sunway TaihuLight being capable of performing floating point operations at 50 times the rate of SETI@home.

To illustrate when HPC and HTC are appropriate, consider the simulation of many coupled oscillators; a change in the position of one oscillator will affect the forces experienced by other oscillators. Therefore the simulation cannot be separated into independent tasks, as all nodes involved will need to be able to access information on all oscillators as the simulation progresses. Communication between calculation processes will be required, therefore shared resources such as common memory and/or a shared filesystem will be required, so a supercomputer or cluster would offer a significant advantage over HTC resources. In contrast, MC has been described as an 'embarrassingly parallel' computational problem, i.e. not only can simulations be separated into different dependent 'batches' and run on different computers, but the task is so well suited to this approach that it would be embarrassing not to do so. HTC is therefore a more natural fit for

---

[6]Folding@home. Available at: https://folding.stanford.edu/ (Accessed: 9 December 2016).
[7]SETI@home. Available at: https://setiathome.berkeley.edu/ (Accessed: 9 December 2016).

MC, as the parallel nature of HPC is not needed or used in MC simulations - but HPC has been used.

2.5.3   Monte Carlo on HPC or HPC resources

A number of authors have published work involving running MC simulations on clusters or distributed resources or other non-standard computing architecture in order to reduce the time needed to perform a simulation.

2.5.3.1   MCRTV

The Monte Carlo for Radiotherapy Treatment plan Verification (MCRTV) is a system to perform MC on a Linux cluster consisting of 14 nodes, with two 3.2 GHz Intel Xeon CPUs per node (Yamamoto et al., 2007). MCRTV takes DICOM-RT files from the TPS, and produces input files for the EGS4 simulations. The simulations are parallelized via the Message Passing Interface (MPI), a set of library routines for programming a variety of parallel computing architectures. An initialisation command from the MPI library is used to send the EGS4 input files to the nodes, with different random numbers for each input file. The nodes execute the EGS4 programs, and further MPI commands are used to collect and sum the outputs. The simulation is generally split into three parts, each using EGS4 - the patient invariant portion of the head, the patient specific portion of the head and the patient/phantom geometry. Phase space files may be used in place of the simulation of the patient invariant portion of head. The voxel size in the simulation is matched to that of TPS resolution, to facilitate comparisons. Analysis of MC calculated dose and and comparison with TPS dose is performed within MCRTV, using tools written for this purpose.

2.5.3.2 VIMC

The Vancouver Island Monte Carlo (VIMC) system (Zavgorodni et al., 2007) is used to run simulations on an HTCondor pool[8], consisting of nine single-core 2.2 GHz computers, which has been extended for RapidArc treatments (Bush et al., 2008). The simulation is divided into three parts, the head of the linac excluding the MLCs, the MLCs alone, and the patient/phantom geometry. BEAM-nrc is used for the linac head simulation, DMLC is used for the MLC simulation, and either DOSXYZnrc or VMC^{++} may be used for the patient/phantom geometry. Batch files are created to perform a simulation for each beam, or each control point for arcs, individually. These batch files are then submitted to an HTCondor pool for execution. Simulations are performed based on the required statistical uncertainties, rather than a set number of histories. VIMC can convert dose to medium into dose to water, before converting the data into DICOM-RT format, which can be imported into a TPS, to utilise the analysis tools of the TPS. The system is fully automated, and is suitable for routinely verifying doses for RapidArc plans. The system has been developed in recent years to incorporate the ability to perform Electronic Portal Imaging Device (EPID) image predictions. Further developments added the ability to perform simulations based on the linac parameters used during delivery, using the positions recorded by the linac into files known as Dynalog files or Trajectory Log files (Popescu et al., 2015).

---

[8]HTCondor. Available at: https://research.cs.wisc.edu/htcondor/ (Accessed: 21 November 2016).

2.5.3.3 SMCP

Swiss Monte Carlo Plan is a framework for performing MC on a Linux based cluster (Fix et al., 2007). It is interfaced to research version of the Varian Eclipse[9] TPS, so can benefit from tools such as outlining, isodose visualisation, and DVH comparisons. In Eclipse, the SMCP is chosen as a calculation type, which launches a Graphical User Interface (GUI) written in QT4. This GUI is used to specify the many options available - the details of the plan and patient geometry are provided by Eclipse; SMCP is registered as a dose calculation, and may be selected within Eclipse just as the AAA algorithm is. The MC simulation is split into three parts, the plan-independent portion of the linac head, the portion of the linac head comprising the beam modifiers and the patient geometry. For each of these, different MC transport methods for can be used, depending on the requirements for accuracy or speed. For the plan-independent portion of the linac head, options for the transport method include phase space files, histogram based source models and full simulation using VMC^{++}. The results of the source modelling may be used to write a phase space file to input to the beam modifier simulation. However, the particles are more commonly passed in memory to the next stage of the simulation, thanks to the use of interface classes. SCMP can run simulations for a given number of histories or a required statistical uncertainty, and can continuously sample linac positions, and correlate these to MU delivery, for dynamic applications. The beam modifier simulation may be performed using EGSnrc, VMC^{++} or an in-house transport code called Photon Interaction (PIN), which allows beam modifiers to be modelled as:-

- completely absorbing

- absorbing based on attenuation only

---

[9]Varian Eclipse Treatment Planning System. Available at: https://www.varian.com/en-gb/oncology/products/software/treatment-planning/eclipse-treatment-planning-system (Accessed: 19 December 2016).

- absorbing using attenuation and first order Compton

- absorbing using attenuation including all Compton

- 'flat' geometry, in which modifiers are collapsed onto a single plane, and assumed to be completely absorbing.

The patient/phantom simulation maybe performed using EGSnrc or VMC^{++}. An in-house algorithm is used to re-sample the CT to the desired resolution. The advantages of SMCP include the range of transport methods, the minimal user input required to perform a simulation so the simulations are robust, and the use of isodose visualisation and DVH tools within Eclipse for plan comparisons.

SMCP has been used to facilitate a variety of research projects. These projects include investigating the effective dose rate throughout irradiated volume from VMAT plans (Mackeprang et al., 2016), and an investigation of Modulated Electron Beam Therapy, in which the photon MLC is used to modulate electron beams (Henzen et al., 2014).

2.5.3.4 MMCTP

The McGill Monte Carlo Treatment Planning (MMCTP) system runs on multiple operating systems (Windows, Linux, Macintosh), and provides a GUI written in REALBasic (now called Xojo[10]) to facilitate the execution of MC simulations on a computing cluster (Alexander et al., 2007). The GUI runs on a standard computer, can import from TPS data in DICOM-RT and standard TPS file formats, and generate input files for the MC simulations. It also acts as an abstraction layer, to simplify the control of the jobs on the cluster using Secure SHell (SSH)[11]

---

[10]Xojo, new name for REALBasic. Available at: https://en.wikipedia.org/wiki/Xojo (Accessed: 16 December 2016).

[11]SSH. Available at: http://www.snailbook.com/protocols.html (Accessed: 5 November 2016).

commands, and provides functions such as contouring and DVH analysis to compare the MC doses with TPS calculated dose distributions; the paper stresses the importance of comparing DVHs on the same platform, to avoid the influence of implementation details. The data is described as being anonymised prior to transmission to the cluster, although it is not clear if the anonymisation is performed by converting data to input files that hold no Personally Identifiable Information (PII). Internally, the system uses a custom file storage format, dubbed the McGill RT format (which is an extension of the Radiation Therapy Oncology Group (RTOG) format). The MC simulation is divided into a linac simulation using BEAMnrc, and a patient/phantom simulation using XVMC. The simulations are run for a large number of histories, rather than a specified statistical uncertainty, using nodes in the cluster to run batches of histories.

MMCTP has been used as a platform to facilitate various research topics. Conneely et al. (2013) used MMCTP to automatically tune MC linac parameters, performing multiple simulations with different electron beam energies. The agreement with measured data was used to select the optimum electron beam energy. Conneely et al. (2014) used MMCTP to simulate the delivery of IMRT to a Farmer chamber in a phantom, in two situations: firstly, with all beams with the gantry angle set to zero; and secondly, with the real planned gantry angles. This showed that using the real gantry angles was the preferred method, and highlighted factors such as dose heterogeneity in the Farmer chamber must be considered if performing verification measurements with all gantry angles set to zero.

2.5.3.5 RTGrid

RTGrid is a project to enable MC simulations to be run on a variety of distributed computing resources, including HTCondor pools, or the UK e-Science National Grid Service (NGS) (Yaikhom et al., 2008; Downes et al., 2009). It is implemented

using the Gridsphere portal framework, and aims to be an accessible interface for Clinical Scientists unaccustomed to distributed computing to access distributed computing resources. RTGrid may be run on a Linux or Windows server, but the soft links used by Linux file-systems allows files used by many jobs to be stored once per simulation, saving file space. The MC simulations may in principle be performed on Windows PCs via HTCondor, Linux clusters, or other computing resources, through the use of profiles to describe the simulation details, and the interaction with the computing resources. In practice the executable programs involved in the MC simulation need to be compiled for either Windows or Linux, so a profile will be associated with only one computing resource. The RTGrid is accessed through a web portal, so no software needs to be installed in the PC of the scientist using the RTGrid.

Simulations are divided into linac head simulations, performed using BEAMnrc, and patient/phantom geometry, performed using DOSXYZnrc. Input files specifying the MC simulation parameters may be uploaded through the RTGrid server, or DICOM and DICOM-RT files may be sent to a DICOM server, and input files produced based on the information bin the DICOM and DICOM-RT files. A number of independent 'jobs' are created, each suitable for executing on a node of the appropriate computing environment. The random numbers in the input files are varied between the jobs. The jobs are submitted to the computing resources, and the progress of the job execution can be monitored via the web portal. Optionally, the simulation can be completed when a certain percentage, e.g. 95% of jobs have been completed, to avoid the whole simulation waiting for a small number of jobs that have not been executed. The RTGrid itself does not provide tools for visualising or analysing the simulated doses, preferring to leave this to dedicated tools.

The RTGrid has been used in a number of research projects. Cufflin et al. (2010) used the RTGrid for MC based prediction of EPID treatment verification images.

Millin and Lewis (2010) used the RTGrid as part of a MC based verification system for verification of stereotactic radiotherapy and stereotactic radiosurgery treatment plans.

2.5.3.6 Monte Carlo on GPUs

GPUs are computer components designed for graphical operations, which contain thousands of processing units on a single chip. NVIDIA[12] markets a number of GPUs that use the Compute Unified Device Architecture (CUDA)[13], an Application Programming Interface (API) that allows the GPU to be programmed. Alternatively, the Open Computing Language (OpenCL)[14] may be used to program GPUs from different manufacturers. The cross-vendor support of OpenCL limits the tailoring of the code to the specific architecture of the hardware being used. Tian et al. (2015) found several comparisons of computation problems implemented on CUDA and OpenCL platforms. CUDA based implementations were found to be faster in most cases, but Tian et al. (2015) expect that the specific problem type and size will influence whether CUDA or OpenCL will be faster.

A common term in computing is a thread, which is the smallest sequence of instructions that can be executed independently. The processing units on a GPU execute threads in groups, known as warps. The execution of all threads in a warp is started together, and the results from a warp are read back at the same time. Reading the data back from the whole warp must wait for the last thread to complete. This is not an issue if all threads are performing very similar tasks, as is common in graphics rendering. For MC transport, different types

---

[12]NVIDIA Available at: http://www.nvidia.co.uk/page/home.html (Accessed: 28 November 2016).

[13]CUDA Available at: http://www.nvidia.com/object/cuda_home_new.html (Accessed: 28 November 2016).

[14]OpenCL Available at: https://www.khronos.org/opencl/ (Accessed: 28 November 2016).

of interactions require different calculations, which may take different lengths of time to calculate. This issue is termed thread divergence, and avoiding it is challenging, but key to maximising the performance of GPUs for MC.

Jia et al. (2010) ported the DPM (section 2.5.1.6) code to NVIDIA GPUs using the CUDA platform, which they called gDPM v1.0. To port the code they had to consider that double precision arithmetic is not available on the GPU, how the processors of the GPU accesses memory both internally and externally to the GPU, and how to accumulate dose while avoiding concurrent writing to memory. gDPM v1.0 was able to decrease calculation time by a factor of 5-6 compared to DPM on a CPU.

gDPM v2.0 (Jia et al., 2011) made further improvements, chiefly by batching the transport of similar particles together so the parallel nature of the warps could be used more effectively. Single precision was used throughout, as it was shown not to affect the accuracy of MC, a new faster random number generator was employed, and improvements in the interpolation of cross-sectional data were made. This allowed gDPM v2.0 to realise an increase in execution speed by a factor of 70-80 compared to DPM.

A further enhancement gDPM v3.0 (Townson et al., 2013) can read phase space files, and sort the particles therein, to allow simulations to be batched into warps. This allowed a head and neck case to be simulated in less than a minute, while performing a simulation for the same case with BEAMnrc and DOSXYZnrc took 2.6 hours for the linac simulation and 5.8 hours for the patient simulation. There was some disagreement between the two simulations however, with 97% of points being within either 2%, or having a point within 2 mm with the same dose.

Hissoiny et al. (2011a) developed a MC code specifically for GPUs, based on PENELOPE, called GPUMCD. This algorithm batched the simulation of similar particles to allow a warp to only consider one type of particle at a time, before gDPM v2.0 had implemented this feature. GPUMCD was able to achieve a

speed increase of 27 times over PENELOPE. GPUMCD as described by Hissoiny et al. (2011a) used only mono-energetic beams, but it was stated that further work would allow simulations using poly-energetic spectra required for realistic treatment plans, and phase space files. GPUMCD has since been used for low-energy Brachytherapy seed dosimetry (Hissoiny et al., 2011b), and adapted to dose calculations in strong magnetic fields (Hissoiny et al., 2011c).

Jahnke et al. (2012) ported the electro-magnetic transport functionality of GÉANT4 to a GPU. They used pre-calculated random numbers rather than calculating them on the fly. For electrons and positrons, steps were simulated separately, rather than simulating all steps for a particles sequentially. This avoided 'thread divergence', increasing the efficiency of simulation on a GPU.

2.6 Comparisons of Monte Carlo and analytical algorithms

2.6.1 Phantom Studies

Many groups have studied the differences between analytical algorithms and MC simulations in heterogeneous phantoms. Krieger and Sauer (2005) studied the differences in the doses calculated in a heterogeneous slab phantom, for PB and CC algorithms implemented in Helax[15], the forerunner of Oncentra, and MC doses calculated with XVMC (section 2.5.1.6). For point doses, they found CC and MC agreed with measurements within 3%, but the PB algorithm predicted doses up to 14% higher than expected in regions of low density (polystyrene). For profiles through the phantom, the PB algorithm generally over-estimated the dose in low density regions, although it under-estimated the dose near to water-density regions surrounding the low-density regions. In water-equivalent material, near

---

[15]Nucletron, Veenendaal, Netherlands, superseded by Oncenrta MasterPlan

low density regions, the PB algorithm generally over-estimated the dose. The CC algorithm agreed well with measurement in water-equivalent regions, even near low density regions. In the low density regions, CC generally under-estimated the dose, but was closer to measurement than was the PB dose. MC agreed well with measurement throughout.

Knöös et al. (1995) compared the EGS4/PRESTA code with the PB algorithm in a heterogeneous phantom. For beams passing parallel to regions of low density (cork, used as a substitute for lungs) the PB dose was 2-5% higher than the MC dose. This was attributed to the inability of the PB algorithm to account for changes in lateral scatter - due to the spatially invariant pencil beam kernels used in the algorithm. For beams passing through the low density regions, the PB algorithm over-estimated the dose by up to 14% for 18 MeV beams. It was however noted that the dose calculated by PB returned to the correct level shortly after the beam re-entered the water density material.

Arnfield et al. (2000) compared doses calculated by a power law method (section 2.4.2), an implementation of the CC algorithm and the PEREGRINE MC dose calculation system, in a heterogeneous phantom. Differences in low density regions for the CC algorithm of 5% were reported.

2.6.2 Clinical Studies

Several workers have published comparisons of dose distributions generated by conventional radiotherapy planning systems and those produced by various MC dose calculation packages. However, due to the large computational load involved in producing each MC plan, the number of cases cited in each study is relatively small and therefore the statistical weight of these studies is low. For example, Ma et al. (2000) initially demonstrated agreement between MC and measurements in various 'phantom' test objects within 2%. They then compared PB with MC

for two clinical example cases (prostate and spinal column) and demonstrated reasonable agreement (within 4%) in homogeneous tissue regions but significant discrepancies in the presence of inhomogeneities (e.g. 20% to the spinal cord). Wang et al. (2002) compared PB with MC in nine cases - four head and neck cases and five lung cases. In the lung cases, calculated doses to tumour volume and lungs were up to 10% lower in MC than PB, whereas doses to the spinal cord were in generally good agreement. Francescon et al. (2000) compared CC with MC in two anatomical cases – mediastinal and breast. They reported good agreement (within 3%) between these methods in the cases studied, but acknowledged that the plans cited involved fairly large field sizes (~10 cm x 10 cm). Greater discrepancies would be expected if smaller fields were involved, due to the sensitivities of these fields on the focal spot size used in the dose calculation method.

2.6.3   Oesophageal Studies

While there have been hundreds of studies comparing analytical algorithms with MC simulations, there have been relatively few performed for oesophageal treatment plans. To demonstrate this, a search was performed of the bibliographic database SCOPUS[16] for abstracts containing the phrases "treatment plan", "Monte Carlo" and "lung"; over 380 articles were found to match these search criteria. Replacing "lung" with "head AND neck" found 150 articles. However, replacing "lung" with "*esophag*" found only 26 articles. Of the articles found, only four (Wills et al., 2009; Irvine et al., 2004; Haga et al., 2014; Jang et al., 2006) involve comparisons of doses from different calculation methods for oesophageal RT.

Wills et al. (2009) compared 15 oesophageal plans calculated with the PB and CC algorithms on Oncentra MasterPlan, as well as comparing plans re-optimised using the CC algorithm against those optimised using the PB algorithm. The dose to the PTV decreased by less than 1% when re-calculating the plans using

---

[16]SCOPUS. Available at: https://www.scopus.com/ (Accessed: 20 August 2017).

the CC algorithm, and OAR doses were found to agree to within 2%. For *type B* algorithms, a proposal was made to reduce the volume of the PTV that was required to receive 95% of the prescription dose (V95%), based on the percentage of PTV overlapping the lungs. This was due to the dependence of the change in PTV V95% on the overlap between PTV and lung. No comparison was made with MC simulations.

Irvine et al. (2004) compared 10 lung plans and 10 two-phase oesophageal plans calculated with the PB and CC algorithms implemented in Helax version 6.1, but did not use MC calculated doses. In the two phase oesophageal technique used, the first phase plan used two beams, one anterior and one posterior; the second phase consisted of three beams, an anterior beam, and two posterior lateral obliques. It was found that the minimum dose to the PTV decreased by up to 23% when the PB plan was re-calculated with the CC algorithm (using the MU from the PB calculation). The under-dosage was attributed to low density tissue within the PTV, a decrease in the scattered dose from the low density region surrounding the PTV, and a secondary build up at the intersection of low and high density materials. No comparison was made with MC simulations.

Haga et al. (2014) compared the doses calculated by the superposition/convolution algorithm in the Pinnacle TPS with the XVMC MC package implemented within Monaco v3.3 (Elekta), for 15 patients treated with VMAT plans. Of the 15 patients, one had oesophageal cancer, five were prostate cancer cases, five were lung cancer cases, three were head and neck cases and one was a rectal cancer case. For the oesophageal case, the agreement between the analytical algorithm and the MC simulation was within 2% for the PTV and lung, but the analytical algorithm underestimated dose to the spinal cord by approximately 5% compared to MC. As this study involved only a single oesophageal case, it is unclear how applicable to other oesophageal cases this result is.

Jang et al. (2006) compared doses for a number of thoracic plans, including four lung and four oesophageal IMRT plans, as calculated by the Corvus system[17] version 4 or Pinnacle versions 6.0i or 6.2b with doses from MC simulations. For the lung and oesophageal IMRT cases Pinnacle agreed with measurement to within 5% or 5 mm in the high and medium dose regions, but differences of 10% were found in lose dose regions in the lungs. The MC simulations showed no differences (compared with measurement) in the low dose regions, and agreement was within 5%/5 mm for the majority of measurement points. For comparisons of point doses between Pinnacle and MC simulation, agreement was within 2%. Comparisons of DVHs for Pinnacle and MC simulated doses were presented for two lung IMRT cases. The DVHs were very similar for the cord and combined lungs for both cases. In one case the DVH for the PTV showed MC calculated higher doses than Pinnacle. The differences found between Pinnacle and measurement in the low dose regions were attributed to differences in the modelling of MLC transmission.

The proximity of the lungs and vertebrae to the oesophagus present difficulties in accurate dose calculation for this treatment region. However, there are relatively few published comparisons of analytical algorithms and MC simulations for oesophageal RT treatment plans, in contrast to other treatment regions. In those studies that did examine oesophageal plans, fewer than five cases calculated with MC were considered. Studies using larger numbers compared different analytical algorithms, without using MC simulations. With the development of a Grid-based MC computation facility discussed in section 3.2.2, it will be feasible to undertake a higher powered study to compare dose distributions generated by both PBE and CCE with the MC alternatives, in an effort to provide a definitive study of these dose calculations in oesophageal RT.

---

[17]North American Scientific, Chatsworth, CA, USA

2.7   Summary

In this chapter, the basic interactions between radiation and matter that are important in radiotherapy energy ranges have been reviewed. The characteristics of photon beams from linacs have been presented, to illustrate the difficulties in modelling these analytically. Various analytical approaches to dose calculations have been reviewed, and the PBE and CCE algorithms as implemented in Oncentra v4.3 have been discussed. An overview of common MC codes has been presented, and methods to use HPC or HTC resources to decrease execution times for MC simulations have been reviewed. Finally, differences found between analytical algorithms and MC simulations has been presented.

BEAMnrc and DOSXYZnrc are specifically aimed at RT, and the physical models used have been optimised for the energies and materials of interest in RT. As such, they are the most used MC codes in RT, based on the number of publications. Its accuracy and active user base make it an ideal choice for this work. The RTGrid involves no specialist hardware, such as high end GPUs, or dedicated computing clusters. It does not require costly software, unlike SMCP, which depends on Varian Eclipse. The RTGrid requires only a Linux server, and Windows Desktop PCs using HTCondor. It has proven validity following work undertaken and published by our group (Cufflin et al., 2010; Millin and Lewis, 2010). The RTGrid is therefore a good choice for a computing platform on which to perform the MC simulations for this work. The next chapter will describe the use of the RTGrid to perform MC simulations with BEAMnrc and DOSXYZnrc.

Chapter 3

Pilot study

3.1 Chapter aims

In this chapter the protocol followed in our centre for production of oesophageal RT plans will be described. The implementation of the RTGrid in Velindre Cancer Centre for performing MC on distributed computing resources will be described, as it was configured at the start of the project. The changes that were made to the RTGrid to allow this work will then be discussed. The production of a CT ramp, OesRAMP, to convert CT data into the material and density information for MC simulations will be described. The adaptation of the existing BEAMnrc linac model to allow plans with EDWs to be simulated will be discussed. The approach to removing CT contrast from CT data will be presented. Finally, the simulation of 12 oesophageal cases will be described, and the results of these simulations will be used to calculate the number of oesophageal cases to simulate in the main study, for the required statistical power.

3.2 Materials and methods

3.2.1 Oesophageal planning protocol

In this work oesophageal plans produced following the Velindre protocol for radical oesophageal plans, based on the SCOPE1 protocol, have been used. The protocol issued to centres involved in the SCOPE1 trial called for a RT prescription of 50 Gy in 25 fractions, each of 2 Gy, delivered over five weeks (i.e. one fraction per weekday, with no RT over the weekend). For this work, only oesophageal plans with a prescription of 50 Gy in 25 fractions were selected.

The SCOPE1 protocol also included detailed guidance on outlining OARs such as the heart, as well as the Gross Tumour Volume (GTV) and margins to be applied to the GTV. The GTV was to be delineated on a CT scan, and the use of PET was encouraged. A number of margins were to be applied to the GTV, including two Clinical Target Volumes (CTVs). A CTV is the GTV plus a margin to account for the spread of the disease that cannot be detected, so will not be included in the GTV. First, the CTV-A was produced by outlining the oesophagus on slices 2 cm superior and up to 2 cm inferior of the GTV (but no further than the gastro-oesophageal junction, if this was within 2 cm of the GTV). From the CTV-A, the CTV-B is produced by growing a 1 cm margin in all directions within the plane of the CT slice. Finally, the PTV is produced by growing a margin of 1 cm superior and inferior, and 0.5 cm in all directions in the plane of the CT slice of the CTV-B. A margin known as a Planning Reference Volume (PRV) was to be added to the spinal cord to account for the uncertainties in the position of this OAR. Again, only oesophageal plans with these volumes were selected for this work, to exclude palliative plans, or plans produced to a different protocol.

Prior to SCOPE1, oesophageal cancer was commonly treated by a two-phase RT

plan. Initially, a two-beam plan was delivered, with the beams irradiating the patient from the front ('anterior') and the back ('posterior'), known as the first phase plan. This would be delivered for a certain number of fractions, then the treatment plan would be changed to a three field plan avoiding the spinal cord, known as the second phase plan, to be used for the remaining fractions. By using a single phase, multi-beam technique, the dose to the heart could be reduced more easily (Cominos et al., 2005). Also, a single phase plan was postulated to be better for the OARs, as the dose they received would be spread over the whole course of radiotherapy. In the two phase approach, the spinal cord would receive the dose just in the first phase, reducing the amount of recovery time for the healthy tissue. The SCOPE1 protocol mandated a single phase approach, but allowed the number and arrangement of beams to be determined on a case by case basis. This is the approach adopted at Velindre for all CFRT oesophageal treatment plans, so all plans included in this work are single phase plans.

The DVH parameters that are reported for oesophageal plans in Velindre are: -

- **PTV V95%:** The volume, as a percentage, of the Planning Target Volume receiving 95% of the prescription dose;

- **PTV D99%:** The dose, in Gray, received by 99% of the Planning Target Volume;

- **GTV D100%:** The dose, in Gray, received by 100% of the Gross Tumour Volume;

- **heart V80%** The volume, as a percentage, of the heart receiving 80% of the prescription dose;

- **Cord PRV D1cc:** The dose, in Gray, received by 1 cubic centimetre of the Planning Organ at Risk Volume for the Spinal Cord;

- **Combined Lung V20 Gy:** The volume, as a percentage, of the combined lungs receiving 40 Gy;

- **Liver V30 Gy:** The volume, as a percentage, of the liver receiving 30 Gy (if included in the treatment field);

- **Individual Kidney V20 Gy:** The volume, as a percentage, of the each kidney receiving 20 Gy(if included in the treatment field);

Since the liver and kidneys are not included in the treatment field, or indeed the CT scan for all oesophageal treatments, these were excluded from consideration. All other DVH parameters were investigated in this work.

3.2.2 RT Grid at Velindre

The RTGrid was introduced in section 2.5.3.5, as a system designed to allow MC simulations to be run on distributed computing platforms. Details of how the RTGrid is used in Velindre will be presented, along with adaptations made to aid in this work.

The RTGrid at Velindre uses HTCondor[1] as the distributed computing resource on which MC simulations, using BEAMnrc, are run. HTCondor (previously known as Condor) is a workload management system that executes programs on a pool of computers while those computers are idle. It is an example of a HTC resource, as discussed in 2.5.2. HTCondor uses the concept of a slot; an executable is run within a slot, the more slots available, the more jobs can be run simultaneously. For computers with Intel CPUs each core of the CPU provides two slots to HTCondor, due to the Hyper-Threading technology of Intel CPUs. Therefore, a PC with a quad-core Intel CPU will be able to run eight HTCondor

---

[1]HTCondor. Available at: https://research.cs.wisc.edu/htcondor/ (Accessed: 21 November 2016).

jobs simultaneously. The HTCondor pool in Velindre Cancer Centre currently has approximately 200 slots, so can in theory run 200 jobs simultaneously. If a single MC simulation was split between all slots, this would provide a increase in efficiency that would exceed even that of VMC^{++} $\big($which demonstrates an increase in efficiency of 150 compared to EGS4 (section 2.5.1.6)$\big)$. The availability of these slots varies depending on how many PCs are in use, or switched off. It is common that only a few tens of slots will be available during working hours, but 200 will be available overnight, or at weekends.

There are five key elements to the RTGrid, illustrated in Figure 3.1:

1. Web portal

2. Web services

3. MySQL database

4. Experiment manager

5. Computational resources

The Web services, MySQL database and the *experiment-manager* run from a single server. The Web portal can be accessed on a web browser from the users' computers and is used to stop and start experiments. The distributed resources are computers throughout the Medical Physics department at Velindre Cancer Centre, although the HTCondor Manager is also installed on the RTGrid server.

All communications between the Web Services and *experiment-manager* are performed by updating the MySQL database. The use of a database for communication facilitates the updating of information about many different experiments concurrently.

The RTGrid provides four 'portlets', for the administration of Users, Resources (i.e. HTCondor), Profiles and Experiments (an individual simulation is termed
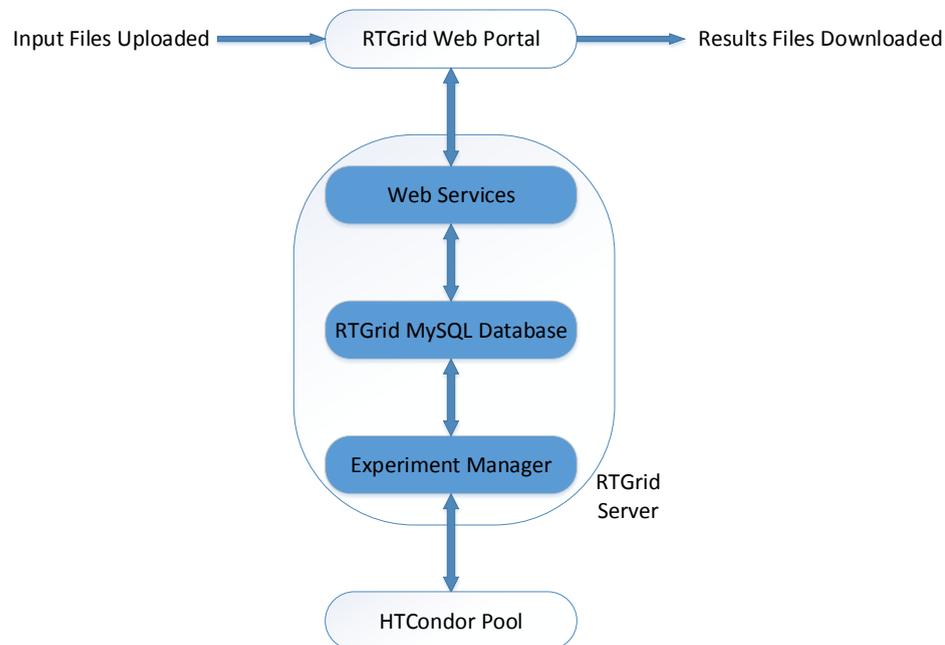
Figure 3.1: Original configuration of the RTGrid at Velindre Cancer Centre.

an experiment). The RTGrid is designed to support many different types of MC simulation on many different types of computation platform. The profiles describe the specifics of a simulation with a given MC package (i.e. BEAMnrc), using a given resource (i.e. HTCondor). Profiles are defined by a Job Submission Description Language (JSDL) document. This is an XML file describing the programs to be executed, the parameters passed to the executables, the resources they can be run on etc. In addition two shell scripts, *pre.sh* and *post.sh*, are run at the start and end of the simulation respectively, to prepare the experiment, and to perform any post-processing steps. A second XML file, called a profile extension, may be used to specify information that is outside the scope of the JSDL file. An example used by some RTGrid profiles (though not employed in this work) would be to specify DICOM peer information in the extension.xml file, used by a *post.sh* to export results to, in DICOM format.
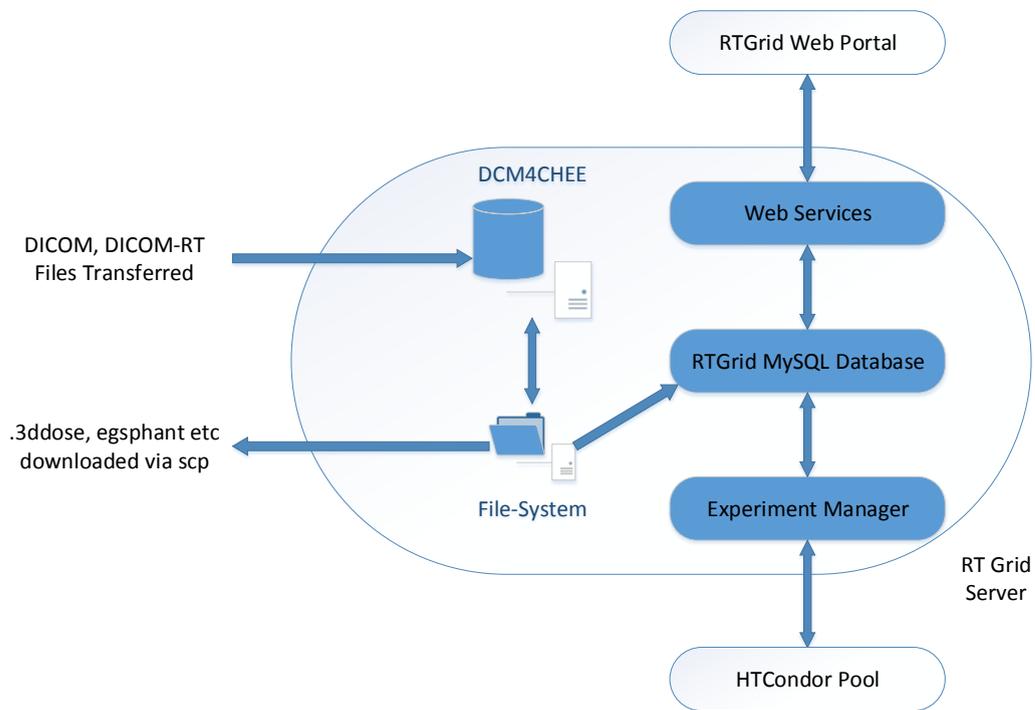
Figure 3.2: RTGrid, as configured for this project.

When the RTGrid is started, the *experiment-master* process is started, to monitor changes in the RTGrid database, and start an *experiment-manager* process for each individual experiment. This also updates the database, for changes in the state of the related experiment.

In order to perform a simulation, a 'watch' process is started, which monitors the file system for new DICOM and DICOM-RT files. The DICOM and DICOM-RT files describing a treatment plan are sent from a client PC to a dcm4chee[2] server, which writes the DICOM and DICOM-RT files to the file system (Figure 3.2). The creation of these files is detected by the watch process, which creates an entry in the database for the new experiment. The *experiment-master* updates the web portal, so that the experiment is visible in the web portal. In the current

---

[2]dcm4chee Image Manager and Archive. Available at: http://www.dcm4che.org/ (Accessed: 19 December 2016).

implementation, nothing further will happen without user interaction - the user must start the experiment via the web portal.

The *experiment-master* 'spawns' an *experiment-manager* for each new experiment. The *experiment-manager* executes the *pre.sh*, which will create the jobs to execute on the HTCondor pool, to match the job as laid out in the JSDL file. In this profile, this is achieved via a number of Python[3] scripts in the profile, which create the BEAMnrc and DOSXYZnrc input files from the DICOM and DICOM-RT data. The *experiment-manager* will update the database with the number of jobs created, and jobs are copied to the HTCondor resources. As the jobs are completed, the HTCondor process on the client computer will return the output to the HTCondor server, and the database will be updated, so that the *experiment-manager* can update the web portal to show progress (Figure 3.3). Initially a 'pre-test' simulation is performed, with one job per beam and 10000 histories per job. The results of these jobs are used to calculate the number of histories that will be needed for 2% statistical uncertainty at the isocentre. From this, the number of jobs that the simulation should be split into to keep the jobs to approximately two hours each in calculated. Typically 30-50 jobs are used per oesophageal plan. Since the time to simulate 10000 histories in the pre-test is used in the calculation, this can be affected if the PC used at this stage is unusually slow or fast, so jobs may not take two hours as intended. As for the pre-test, during the main simulation the web portal will be updated with progress as the jobs move from submitted, to running to finished.

HTCondor will suspend the job being executed on a PC when that PC starts to be used, i.e. HTCondor detects keyboard or mouse activity; if it cannot be started on the same PC within a set time, the HTCondor server will submit the job to a different available slot. As all jobs must finish before a simulation is completed, the completion of a simulation may be delayed by a single job. A setting in the

---

[3]Python Software Foundation. Python Language Reference, version 2.7. Available at http://www.python.org

RTGrid profile can be used to complete the simulation if a set percentage of the jobs, e.g. 95%, have been completed. This feature was not used in this work, in case the outstanding 5% of jobs all related to the same beam, and this resulted in poor statistical uncertainties for this beam. As most simulations were performed overnight or at weekends, not enabling this feature did not result in problems in this work.
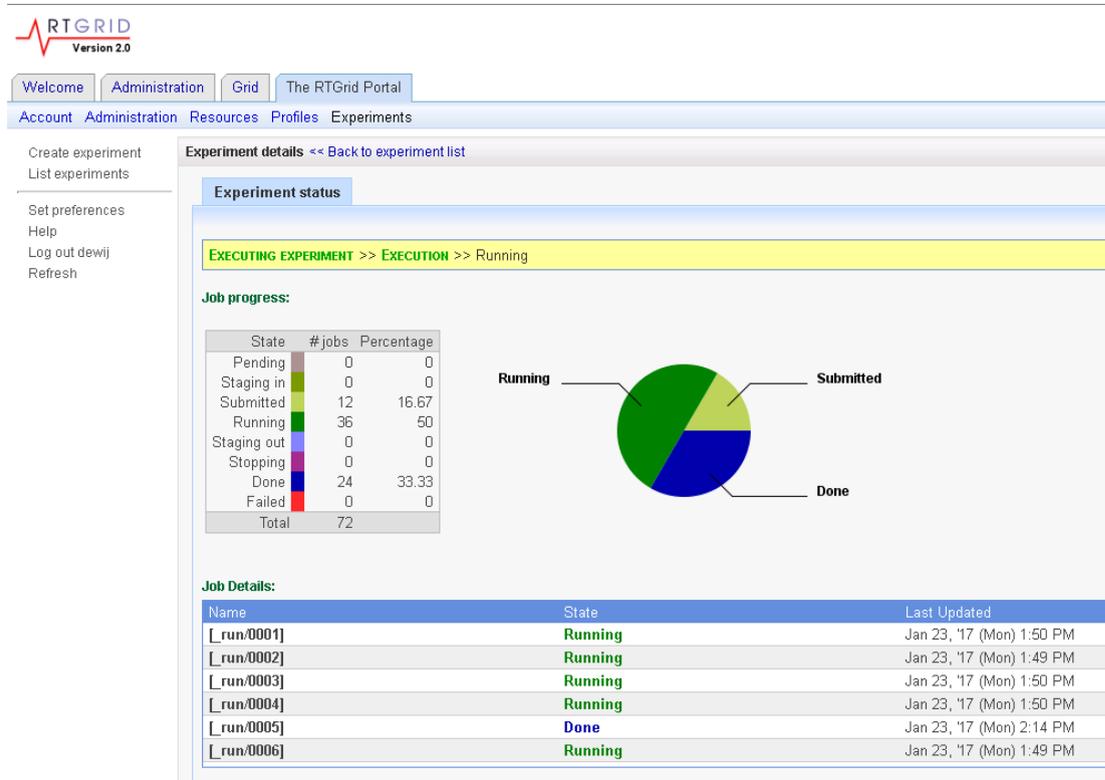
Figure 3.3: Screenshot of RTGrid, showing the progress of an experiment.

After simulation, the SSH is used to download the .3ddose files, egsphant files and logs. SSH is called from a Python script, so the same files are downloaded each time for consistency. The 3ddose is then calibrated in terms of dose, correcting for back-scatter. The dose is converted from Monte Carlo, Dose To Material (MC D_{Mat}) to Monte Carlo, Dose To Water (MC D_{Wtr}), and both set of dose files converted to DICOM-RT. Python scripts used to call CERR (Deasy et al., 2003) scripts to create .cerr files for PBE, CCE and both types of MC dose. This

facilitated the production of DVH data using the same method, as recommended by authors such as Alexander et al. (2007). Using CERR for the DVH production also allowed the DVH production to be automated, which would not have been possible in Oncentra.

3.2.3 BEAMnrc on RTGrid

DOSXYZnrc is capable of using various inputs, called sources, which may be point sources, phase space files, simulations of linac from BEAMnrc or others. The profile used on the RTGrid for this work use ISOURCE9, which uses a BEAMnrc linac compiled to a .dll file. DOSXYZnrc is able to call the .dll file to 'request' an input particle. As with SMCP in section 2.5.3.3, this avoids copying large phase space files to HTCondor slots.

The jobs write .pardose files, rather than .3ddose files, which are a more compact way of storing dose data. Once the .pardose files are copied back to the RTGrid/HTCondor server, they are converted to .3ddose files.

3.2.3.1 Naming convention for structures

In order to facilitate analysis, it was decided to enforce a naming convention for all Regions Of Interest (ROIs) in the DICOM-RT Structure Set file prior to MC simulation. A script was written in Python to check each ROI for each case prior to submitting it to the RTGrid. If a case was found to have an ROI that did not match the desired naming convention, the code either tried to substitute a recognised name based on a look up table, or submission would be prevented until a suitable substitution was added to the look up table. For example, if an ROI named 'Left Lung' was encountered, and there was an entry in the look up table relating 'Left Lung' to 'Lung_Left' (the preferred name for the left lung ROI), the ROI would be renamed in the DICOM-RT file to 'Lung_Left'. If however a

ROI with a variation that had not been predicted, e.g. 'Haert', the DICOM and DICOM-RT files would not be copied to a staging area used to store data prior to uploading to the RTGrid, and an error logged. Once the look up table had been amended to link 'Haert' to 'Heart', the substitution would be performed, and the DICOM and DICOM-RT files copied to the staging area.

3.2.4 Developmental work required

In order to simulate oesophageal plans on the RTGrid, the following changes needed to be made to the profile: -

1. changes to allow MC simulations using Velindre CT scans;

2. changes related to the oesophageal radiotherapy technique used at Velindre;

3. changes relating to reproducing the Oncentra approach to dose calculation, to allow appropriate comparisons to be made.

3.2.4.1 CT Ramp

Radiotherapy dose calculations use CT scans of patients, which are maps of Hounsfield Units (HU), a quantity related to relative attenuation coefficients (Equation 3.1). BEAMnrc calculations require the simulation geometry to be represented as maps of the material composition and density, e.g. in an egsphant file that describes the geometry using two matrices, one listing the material for each voxel, and the other listing the mass density (rather than electron density) for each voxel. (A separate file lists the interaction cross sections for different energy ranges for the materials specified in the egsphant files). MC treatment planning therefore requires a method to convert HU to materials and densities — this is known as a CT ramp. ICRU (1992) lists the material compositions and interaction data for approximately 100 tissues found in the human body. The CT ramp
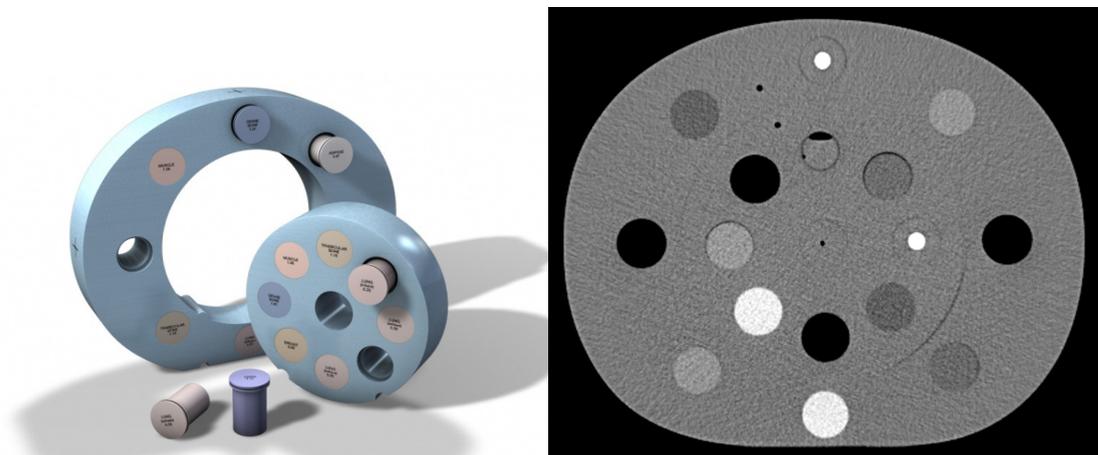
(a) Illustration of the CIRS 062M phantom.   **(b)** CT Scan of the CIRS 062M phantom.

Figure 3.4: The CIRS 062M phantom, which features inserts which mimic the electron densities of tissues of the body.

used in this work was created based on the Stoichiometric method of Schneider et al. (1996), as described by Vanderstraeten et al. (2007). The ramp in this work will be termed OesRAMP.

Stoichiometry is the calculation of relative quantities of chemicals in a compound or mixture. The basic steps in this approach are:-

1. Scan a range of test materials of known elemental composition with the CT scanner protocol of interest, to determine the HU of the test materials;

2. Determine the attenuation coefficients of the test materials;

3. Use an iterative fitting method to determine parameters in the relation between attenuation coefficients and material composition;

4. Use these parameters to predict the HU for standard tissues, based on their elemental composition.

$$H = 1000\left(\frac{\mu}{\mu_{water}} - 1\right) \tag{3.1}$$

The first step was to scan the CIRS 062M phantom[4] in the CT scanner at Velindre Cancer Centre, using the CT protocol used for oesophageal patients – 'RTChest-Contrast'. This CT scan included the HU for inserts of known densities and elemental composition (Figure 3.4). Equation 3.1 shows the relation between HU and the attenuation coefficient, μ. Since both μ and μ_{water} vary with the energy spectrum of the incoming radiation beam, the HU values will vary too. Equation 3.2 describes the total attenuation coefficient, μ, of a material that is a mixture of elements with elemental weights w_i, atomic numbers Z_i and atomic masses A_i, where: -

- ρ is the mass density;

- N_A is Avagadro's Constant, 6.022×10^{23} mol^{-1};

- K^{ph}, K^{coh} and K^{KN} are energy dependent parameters that characterise different interaction cross-sections, K^{ph} for photoelectric absorption, K^{coh} for coherent scattering including the binding correction, and K^{KN} for the Klein-Nishina coefficient,

$$\mu = \rho N_A \sum_{i=1}^{n} \left(\frac{w_i}{A_i}\left(K^{ph}Z_i^{4.62} + K^{coh}Z_i^{2.86} + K^{KN}Z_i\right)\right) \tag{3.2}$$

Equations 3.1 and 3.2 may be applied to each insert of the CIRS 062M phantom. Since we are only interested in the attenuation relative to water, Equation 3.2 can be rearranged as Equation 3.3, leaving two fitting parameters, $k1 = K^{ph}/K^{KN}$ and $k2 = K^{coh}/K^{KN}$. Code was written in Python to extract the HU for the various inserts from the CT scans, look up the appropriate elemental weights, atomic

---

[4]Computerized Imaging Reference Systems, Inc.

number and atomic masses of the inserts, and calculate the fitting parameters $k1$ and $k2$. This was performed by a function to minimise the difference between the published densities of the CIRS materials and the values calculated from Equation 3.3. The values that this method yielded were $K^{ph}/K^{KN} = 2.14 \times 10^{-5}$ and $K^{coh}/K^{KN} = 7.42 \times 10^{-4}$. These compare favourably with the values reported Vanderstraeten et al. (2007) (Table 3.1). In addition, the value of K^{coh}/K^{KN} calculated is a positive number, and Vanderstraeten et al. (2007) state that a negative value of K^{coh}/K^{KN} is unphysical.

$$\mu = \rho N_A K^{KN} \sum_{i=1}^{n} \left(\frac{w_i}{A_i} \left(\frac{K^{ph}}{K^{KN}} Z_i^{4.62} + \frac{K^{coh}}{K^{KN}} Z_i^{2.86} + Z_i \right) \right) \tag{3.3}$$

The densities of the different inserts were plotted against the calculated HU (Figure 3.5). This yielded three distinct portions of the graph; air and lung, soft tissue and bony tissue. As the equations fitted to the first two portions of the graph were very similar they were combined into a single fit covering air, lung and soft tissue (Equation 3.4). The R^2-squared value for this equation was 0.9998, showing that there was little deviation between the equation and the measured data. The CT ramp was therefore defined by the two equations, applying Equation 3.4a below HU value of 55 and Equation 3.4b above 55, where 55 is the HU value where the two equations give the same mass density value. This enabled HU to be converted to a density.

$$\text{Mass Density} = 9.931 \times 10^{-4} \times HU + 1.0185 \text{ for HU } \leq 55 \tag{3.4a}$$

$$\text{Mass Density} = 5.287 \times 10^{-4} \times HU + 1.0445 \text{ for HU } > 55 \tag{3.4b}$$

To close the loop, for the inserts of CIRS 062M phantom, the HU were calculated from Equation 3.4, and the calculated HU plotted against the mean measured HU. If the approach is correct, the data should lie along the line y=x. To test this, the equation y=x was fitted to the data, and the R^2-value calculated. The R^2-value

CT Scanner	Location	K^{ph}/K^{KN}	K^{coh}/K^{KN}
GE ProSpeed	Inselspital Radioonkologie, Berne, Switzerland	1.11×10^{-5}	4.79×10^{-3}
Philips Gemini GLX PET/CT	Ghent University Hospital, Belgium	1.99×10^{-5}	9.26×10^{-4}
Siemens Somatom Sensation Open	University Hospital Tuebingen, Germany	2.11×10^{-5}	3.70×10^{-4}
OesRAMP		$\mathbf{2.14 \times 10^{-5}}$	$\mathbf{7.42 \times 10^{-4}}$
Siemens Somatom Sensation Open	Velindre Cancer Centre, Cardiff, UK	2.30×10^{-5}	3.49×10^{-5}
Siemens Somatom Sensation Cardiac 64	Clinica Quadrantes, Lisbon, Portugal	2.65×10^{-5}	-2.84×10^{-4}
GE HiSpeed QX/i	Royal Marsden Hospital, London, UK	2.78×10^{-5}	-3.98×10^{-4}
Toshiba Xvision/EX	HUV Macarena, Seville, Spain	3.89×10^{-5}	-8.37×10^{-4}
Siemens Somatom Plus 4	Ghent University Hospital, Belgium	4.24×10^{-5}	-1.77×10^{-3}

Table 3.1: Comparison of results of stoichiometric calibration with published data, adapted from Vanderstraeten et al. (2007).
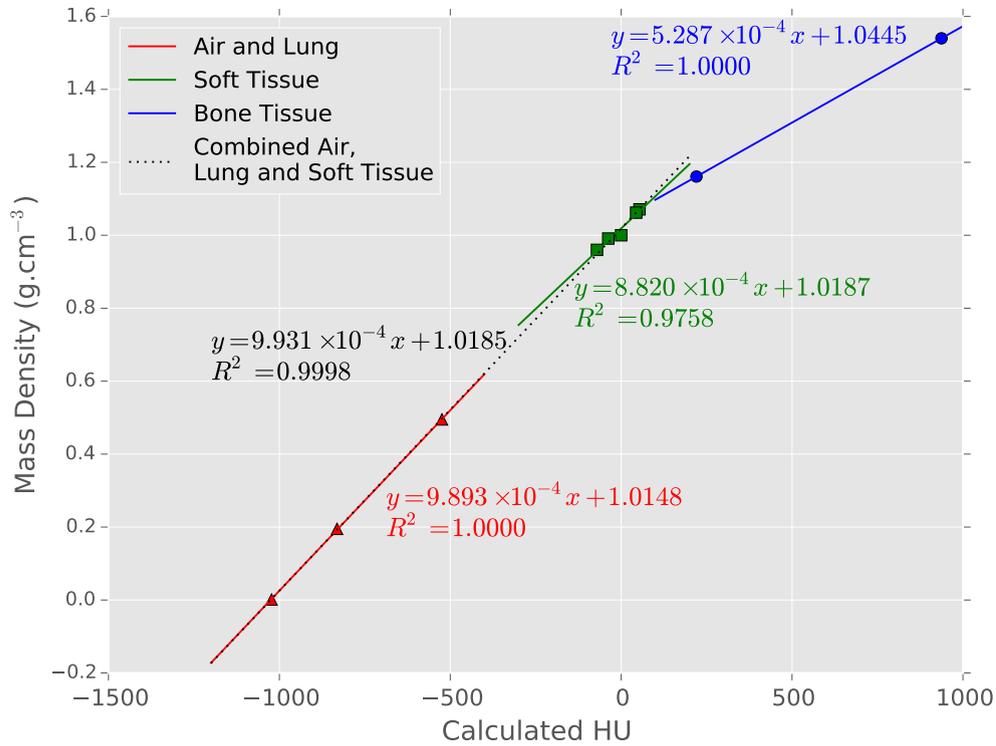
Figure 3.5: Hounsfield Unit calibration curve, for the RTChestContrast protocol, showing fitting parameters for the air and lung region, the soft tissue region, the bone tissue region, and the combined air, lung and soft tissue regions.

was 0.9992, demonstrating the ability of the method to recover the data used to generate the CT Ramp, verifying that the work was done without significant error (Figure 3.6). For each insert in the CIRS 062M phantom there was a range of HU values measured. The error bars in Figure 3.6 are the standard deviations of the ranges of HU measured for each insert.

Equation 3.4 is a continuous function converting HU to *density*. A method to convert HU to *material* is still required; this will not be a continuous function, but a conversion of a continuous variable to a discrete series of materials. Vanderstraeten et al. (2007) recommend 14 discrete materials, or bins, in the ramp, and this approach was followed in this work.
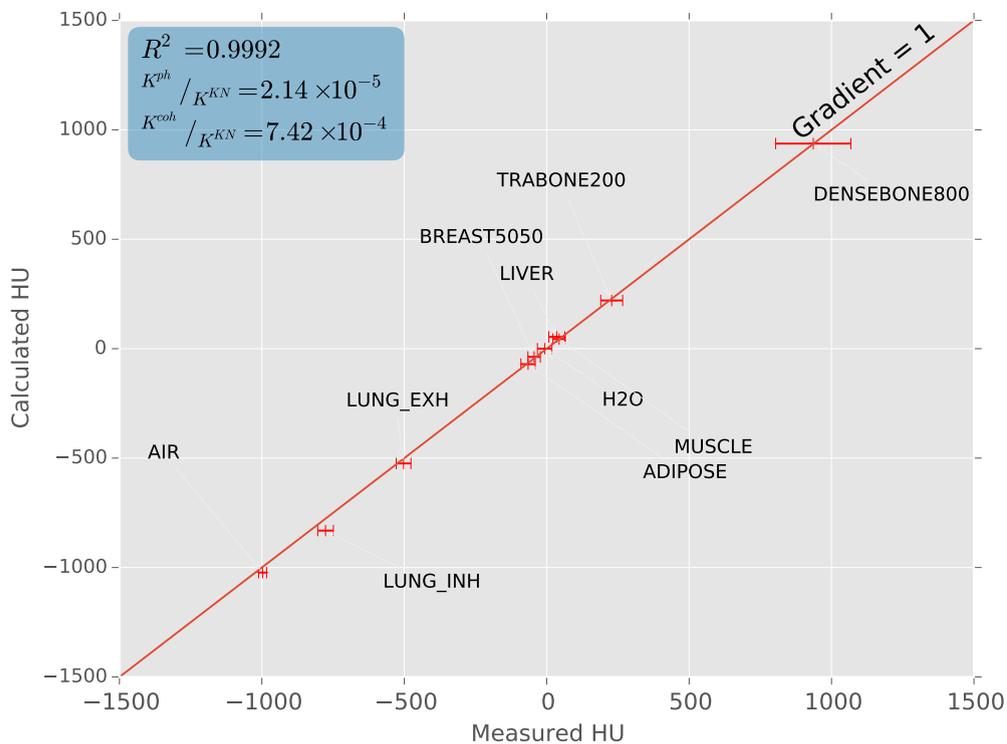
Figure 3.6: Comparison of calculated and measured HU for CIRS 062M phantom. The error bars are the standard deviations of the ranges of HU measured for each insert.

The HU boundaries as given in Table 3 of Vanderstraeten et al. (2007) were initially used to determine the boundaries to be used in this work, by converting HU to densities using the ramp in Vanderstraeten et al. (2007), then converting these back to densities using OesRAMP. These HU values were used to visualise which regions of a CT slice would be treated as each material, for the central CT slice of each patient in the pilot study, and the CT of the CIRS phantom. It was seen that the conversion to material was not always correct. This is illustrated in Figures 3.7 and 3.8. Both show the density (a) of the voxels in a slice through an egsphant file, and then for AIR (b), LUNG (c), ADIPOSE (d), MUSCLE (e) and the ten bone bins combined (f). The voxels converted to the material of interest are shown in white, and all the other voxels are shown as black.

In Figure 3.7(c) it can be seen that many voxels outside the lung have been set to LUNG, whereas in Figure 3.8(c) the only voxels set to lung are in the lung, or a small 'halo' around the surface of the patient. This is due to the CT scan having a higher resolution than the egsphant file, which matches the resolution of the dose calculation, 3 mm × 3 mm. This causes the the voxels at the patient surface to include some air and some soft tissue, and appear to be an average of the two HU values, which is set to LUNG. As this affects only a single layer of voxels, the effect on the dose calculation can be neglected. Other cases exhibited other related issues, such as voxels in the lung being set to AIR, and voxels in bones being set to MUSCLE. To correct these incorrect assignments, the HU values used as the boundaries were adjusted until AIR, LUNG, ADIPOSE, MUSCLE and BONE1 showed no obvious issues.

The width of the bone bins from Vanderstraeten et al. (2007) was approx 0.9 g.cm^{-3}, which is equivalent to 170 HU for OesRAMP (c.f. 150 HU for the ramp of Vanderstraeten et al. (2007), which has a different gradient). The boundaries of the ten bone bins were set based on the interface of MUSCLE and BONE1 as the starting point, then regularly spaced at 170 HU. The upper limit of BONE10 was set to 4000, as this is above the maximum HU the scanner would produce, to avoid the code encountering a runtime error if it tried to interpolate a value outside the defined boundaries.

The HU and densities defining OesRAMP were then included in the source code for ctcreate.exe. This is not the ctcreate from BEAMnrc, but a program written in C[5] as part of RTGrid development (Downes, 2010), which creates the egsphant file from CT files. When this ctcreate.exe is called with the appropriate option, OesRAMP is used in the creation of the egsphant file.

---

[5]The C Programming Language. Available at: http://www.open-std.org/jtc1/sc22/wg14/ (Accessed: 25 November 2016).
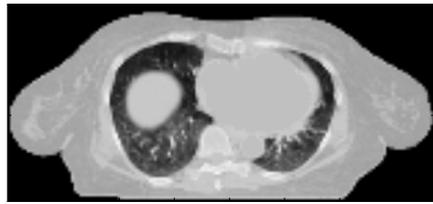
(a) Plot of density in sample slice.

(b) Voxels converted to Air shown as white, other materials shown as black.

(c) Voxels converted to Lung shown as white, other materials shown as black.

(d) Voxels converted to Adipose shown as white, other materials shown as black.

(e) Voxels converted to Muscle shown as white, other materials shown as black.

(f) Voxels converted to any bone bins shown as white, other materials shown as black.

Figure 3.7: Plots showing material assignment based on incorrect boundaries, plus plot of density for comparison.
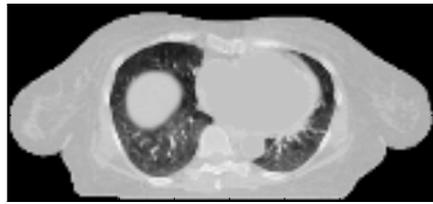
(a) Plot of density in sample slice.

(b) Voxels converted to Air shown as white, other materials shown as black.

(c) Voxels converted to Lung shown as white, other materials shown as black.

(d) Voxels converted to Adipose shown as white, other materials shown as black.

(e) Voxels converted to Muscle shown as white, other materials shown as black.

(f) Voxels converted to any bone bins shown as white, other materials shown as black.

Figure 3.8: Plots showing material assignment based on final boundaries, plus plot of density for comparison.

Vanderstraeten Ramp		Mass Density (g.cm^{-3})		OesRAMP—initial		OesRAMP—adjusted	
HU Lower	HU Upper	Lower	Upper	HU Lower	HU Upper	HU Lower	HU Upper
	-900		0.104		-921	-1024	-950
-900	-100	0.104	0.904	-921	-115	-950	-165
-100	20	0.904	1.024	-115	6	-165	0
20	100	1.024	1.104	6	113	0	100
100	250	1.104	1.165	113	228	100	270
250	400	1.165	1.255	228	399	270	440
400	550	1.255	1.345	399	569	440	610
550	700	1.345	1.435	569	739	610	780
700	850	1.435	1.525	739	909	780	950
850	1000	1.525	1.615	909	1079	950	1120
1000	1150	1.615	1.705	1079	1250	1120	1290
1150	1300	1.705	1.795	1250	1420	1290	1460
1300	1450	1.795	1.885	1420	1590	1460	1630
1450		1.885		1590		1630	4000

Table 3.2: Table of the boundary values used to distinguish different tissues types, for OesRAMP.

3.2.4.2 EGS4DAT file

To complement the CT Ramp, a file containing the interaction cross-sections for the materials in the ramp is required. For BEAMnrc/DOSXYNnrc, this information is held in a file with the extension .egs4dat. BEAMnrc is supplied with sample .egs4dat files for elements commonly found in linac heads, and certain human tissues, matching those used by the standard BEAMnrc CTRAMP, i.e. air, lung, soft tissue, and bone. In order to perform simulations using the 14-bin ramp, an .egs4dat file containing interaction cross-sections for the 14 materials of interest had to be created. This was done using 'egs gui', a program within BEAMnrc (Figure 3.9). For AIR, LUNG, ADIPOSE and MUSCLE the ICRU density files were used. For the bone bins, the approach of Schneider et al. (2000) was used, in which the weight of elements for bone tissue was interpolated between red/yellow marrow and skeleton cortical bone, for the densities of the centre of the bone bins (Table 3.2).

3.2.4.3 SPR calculation

The absorbed dose to a given material, D_{mat} may be converted to the absorbed dose to water, D_{wtr}, using Bragg-Gray cavity theory, defined by Equation 3.5, where SPR$_{wtr,mat}$ is the unrestricted Stopping Power Ratio (SPR), water to material (a restricted SPR is one restricted to energy losses below a certain cut-off energy). Bragg-Gray cavity theory is applicable when a small volume of water, small enough not to perturb the fluence of charged particles, is substituted for a volume of the material being irradiated.

$$D_{wtr} = D_{mat} \times SPR_{wtr,mat} \tag{3.5}$$

SPRRZnrc (Kosunen and Rogers, 1993) was used to perform simulations using the
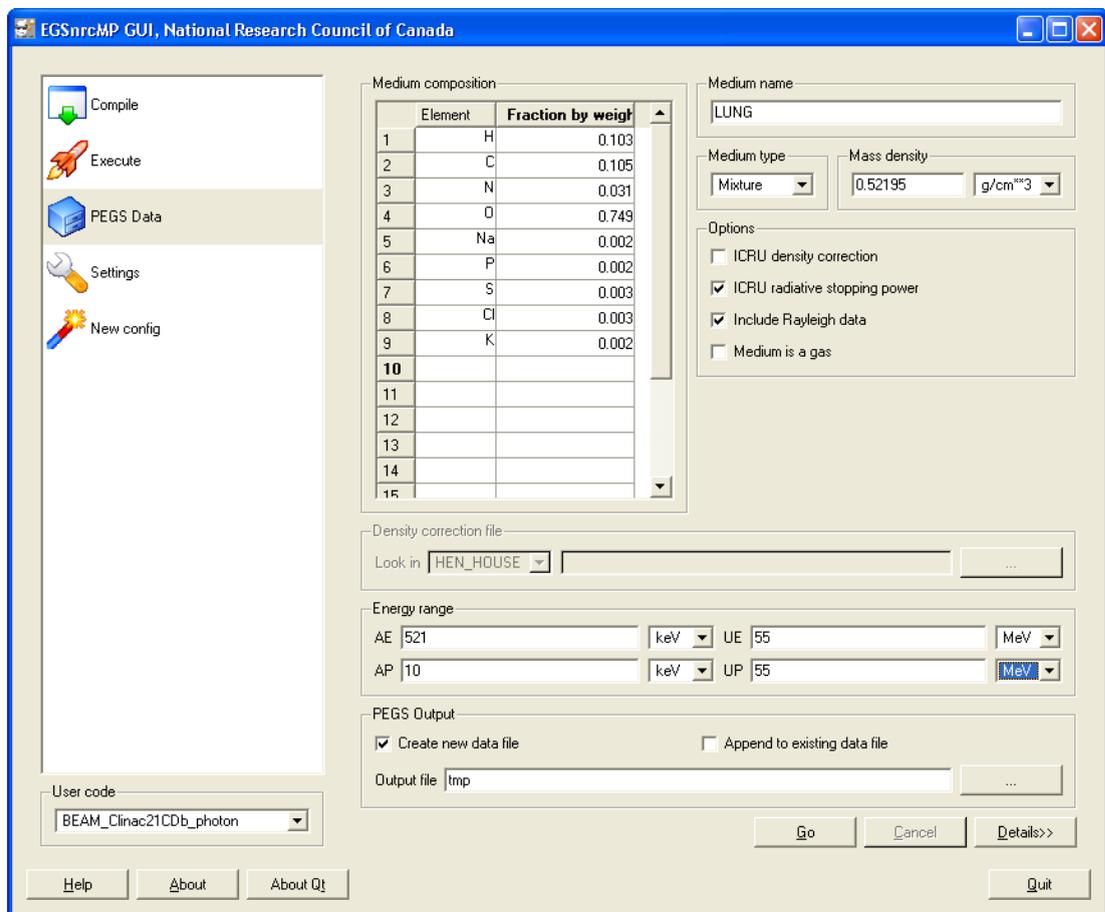
Figure 3.9: Screenshot of egs_gui, used to create .egs4dat file.

same MC linac model used for the oesophageal simulations, to calculate the SPR values for the 14 materials used in this work. These were compared with standard values taken from Siebers et al. (2000), the results are shown in Figures 3.10 and 3.11.

3.2.4.4 Simulation of Enhanced Dynamic Wedge

EDW fields have been simulated on the RTGrid using the component module 'EDW', by Downes (2010). The component module was written for BEAM-nrc 2007, and BEAMnrc 2010 was used in this work. Therefore the differences between the cm.macros and cm.mortran files for the Jaws CM for BEAMnrc
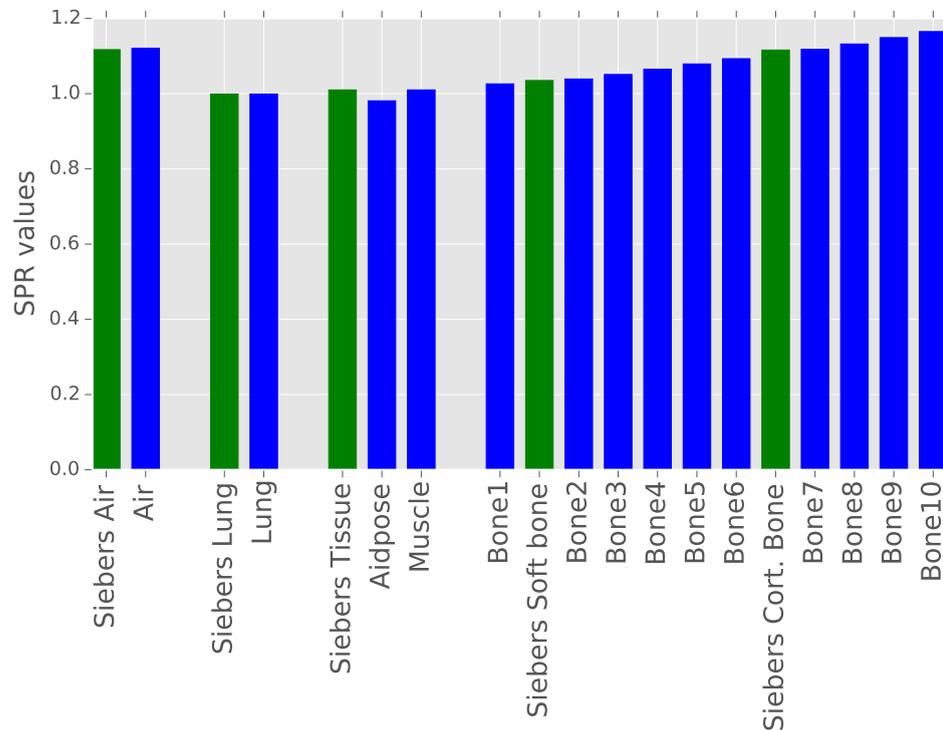
Figure 3.10: Comparison of calculated and published SPR values, for a 6 MV beam from the Varian 2100 linac model.

2007 and BEAMnrc 2010 were studied, and the changes to the cm.macros and cm.mortran were reproduced for the EDW Component Module.

3.2.4.5 Calibration of linac model

The values produced in the .3ddose file from DOSXYZnrc are energy deposited per unit mass, normalised to the number of incident particles, i.e. electrons incident on the target. It is therefore common to perform a simulation for 100 MU, find the energy deposited per unit mass, per incident particle, at a reference point at which the dose from 100 MU is known, and use this value to produce a calibration value. This calibration value can then be applied to subsequent simulations to convert from the values in the .3ddose file to Gy per MU. Chetty et al. (2007)
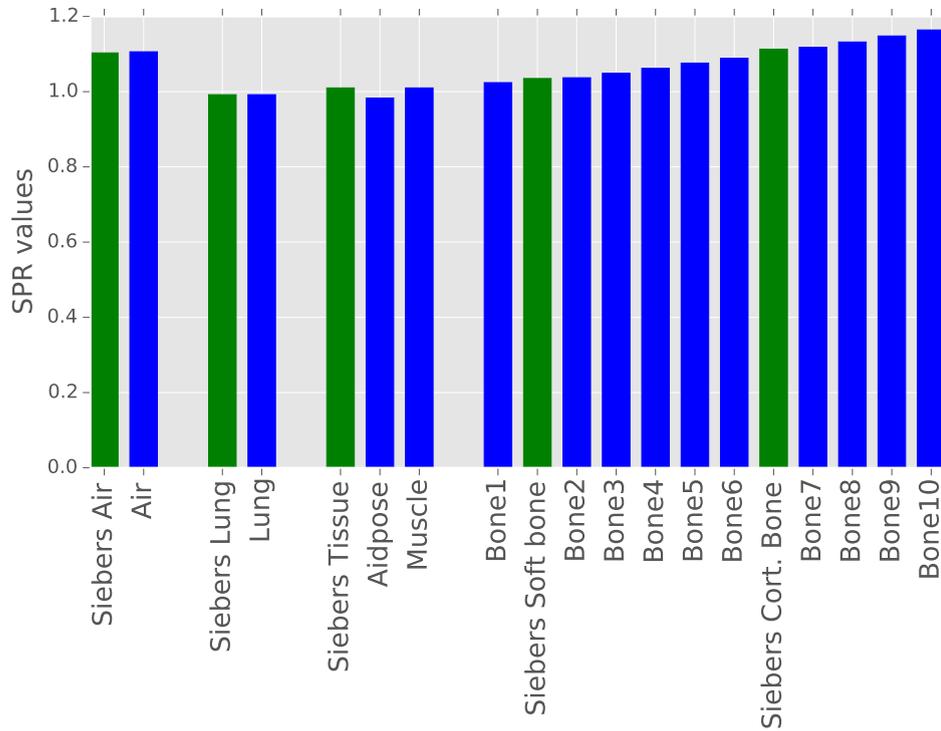
Figure 3.11: Comparison of calculated and published SPR values, for 10 MV beam from the Varian 2100 linac model.

caution against using a single voxel to produce the calibration value, to avoid the calibration value being affected by the statistical uncertainty of the value in a single voxel. In this work, simulations were performed with 100 MU delivered to a cubic water phantom of side 20 cm, at 90 cm Focus to Source Distance (FSD) for a 10 cm by 10 cm field. The dose values were plotted between 5 cm and 15 cm deep on the central axis. A curve was fitted to the values, and the value of the curve at 10 cm deep used in the production of the calibration values. By fitting a curve to many voxels, the effect of the statistical uncertainty will be averaged out.

3.2.4.6 Backscatter correction

The method used to calibrate the linac model described in section 3.2.4.5 will not account for radiation scattered from the jaws into the MU chamber. Consider the definition of Output Factor (OPF) as the ratio of dose at a reference point from a given field to that from a 10 cm × 10 cm field (British Institute of Radiology, 1996); this FSF will have three components, head scatter, phantom scatter and backscatter. The head scatter is the change in radiation at the reference point due to the collimators occluding more or less of the effective radiation sources in the head (such as target, flattening filter, collimator faces). The phantom scatter is the change in radiation at the reference point due to the change in the volume of the phantom being irradiated as the collimators are moved in or out of the beam. The backscatter effect is due to radiation being scattered from the upper surface of the collimators into the monitor chamber, changing the signal measured by the chamber relative to the 10 cm × 10 cm collimator settings. Closing the collimators will lead to an increased signal, so 100 MU will be reached with less radiation fluence impinging on the chamber from above. The backscatter effect will not be accounted for in the calibration method, so a method must be found to correct for this in order for the calibration to give accurate doses across field sizes.

Verhaegen et al. (2000) found that the backscatter contribution was $2 - 3\%$ for static fields, for linacs that were affected by it, but stated that EDW fields would show larger backscatter contributions. Furthermore, Verhaegen et al. (2000) found that the source of the backscatter effect was backscattered electrons from the secondary collimators, not the MLCs.

Yu et al. (1996) and Duzenli et al. (1993) found that Varian 600C linacs have little or no backscatter effect, due to the copper plated steel exit window of the mica ionisation chamber, which absorbs the backscattered radiation; Varian 2100 linacs, which have a kapton monitor chamber, did demonstrate a backscatter

effect.

Hounsell (1998) showed that Elekta Precise linacs do not demonstrate a backscatter effect, due to the presence of steel absorber above the flying wedge.

Measuring the contribution of backscatter is challenging. Some authors used the telescopic technique (Yu et al., 1996), to isolate all radiation but narrow beam, so that the measured dose is proportional to Bremsstrahlung generated in the target. Others use variation of charge integration or pulse counting technique, to analyse the signal from the monitor chamber as a function of pulses from Pulse Forming Network (Lam et al., 1998).

For linacs that do demonstrate the backscatter effect, there are a few approaches to dealing with it. Popescu et al. (2005) placed a scoring plane above the collimators, and performed a simulation which describes the radiation leaving the patient invariant portion of the linac head (i.e. the components above the phase space file are fixed for all photon fields for a given energy). The phase space file produced is then used as a input to subsequent simulations, in which the energy deposited in the monitor chamber from the backscattered radiation is produced, as well as the dose deposited in a phantom or patient geometry. For a given field, the dose in the monitor chamber due to radiation incident from above can be found from the initial simulation, combined with the dose due to the backscattered radiation. This can be compared with the dose accumulated in the monitor chamber for reference conditions, and so the absolute dose from the MC can be determined. This is an attractive approach, as it mimics what occurs in the linac. However, the simulation based on a phase space file above the jaws does not fit with the ISOURCE9 beam library based approach used in this work.

Liu et al. (2000) performed simulations of a Varian 2100 linac head with various collimator settings, then used the results to produce equations describing the backscatter from each pair of jaws, and the masking effect of the upper jaws on backscatter from the lower jaws. This method was able to predict the change in

output due to backscatter to within 0.5% of measured data. The equations are applicable (after suitable validation) to other Varian 2100 linacs without having to reproduce the simulations. The approach of Liu et al. (2000) was adopted in this work. Simulations were used to create three equations. Equation 3.6 describes the backscatter from each $y-$jaw, $R_y(y)$, which follows a third order polynomial. Equation 3.7, is a linear equation describing the backscatter from the $x-$jaws, $R_x(x)$. Equation 3.8 is another third order polynomial describing the shielding effect of the $y-$jaws on the backscatter from the $x-$jaws, $P_y(y)$. These can be used to calculate the relative backscatter, $S_{c,b}(x, y)$, from any field of collimator settings x and y, and Equation 3.9 used to calculate the change in dose due to backscatter, relative to a reference field.

$$R_y(y) = 1.54 \times 10^{-2} - 8.45 \times 10^{-2}y + 4.47 \times 10^{-5}y^3 \tag{3.6}$$

$$R_x(x) = 0.40 \times 10^{-3} - 1.87 \times 10^{-2}x \tag{3.7}$$

$$P_y(y) = 3.95 \times 10^{-2}y - 3.55 \times 10^{-5}y^3 \tag{3.8}$$

$$S_{c,b}(x, y) = \frac{[1 + R(x_0, y_0)]}{1 + R(x, y)} \tag{3.9}$$

3.2.4.7 Reproducing TPS approach

The changes to the RTGrid profile to perform the MC simulation in a way that mimics the TPS are:

- intra-venous contrast must be treated as water equivalent;

- a 5 cm extension must be added superiorly and inferiorly to account for scatter outside the CT scan;

- the CT couch top, and other objects outside the external contour must be treated as air.

Intra–venous contrast is frequently used to aid in the delineation of the GTV in oesophageal radiotherapy. An iodine–based contrast agent is injected shortly before the CT scan is acquired, so that the contrast agent enters the circulatory system. Due to the iodine content, the contrast agent has a higher electron density than that of soft tissue, so regions with a high concentration of iodine will be enhanced (appear more white) on the resulting CT scan. For certain tumours, including oesophageal tumours, the increased blood flow to the tumour will help the radiation oncologist to outline the tumour on the CT scan. It is common for larger blood vessels, notably the aorta, and the heart to have a high concentration of iodine, and appear enhanced on the resulting CT scan. As the iodine is filtered from the body, the kidneys and bladder may also have a high concentration of iodine, and appear enhanced.

In addition to changing the appearance of the CT scan, the change in the electron density in regions of high iodine concentration will alter the attenuation of radiotherapy treatment beams as calculated by treatment planning systems. This would be acceptable if the iodine was present during treatment, but it will have been removed from the blood stream by the body long before treatment starts. Therefore, the presence of the iodine must be corrected for prior to radiotherapy dose calculations, or the dose calculated will be incorrect. The standard method for doing this in Oncentra (and other treatment planning systems) is to outline regions of high enhancement, and mark these areas as having a bulk density representing soft tissue. In Velindre, a bulk density of 1 g/cc is used (i.e. the enhanced regions are assumed to be water density). An alternative approach
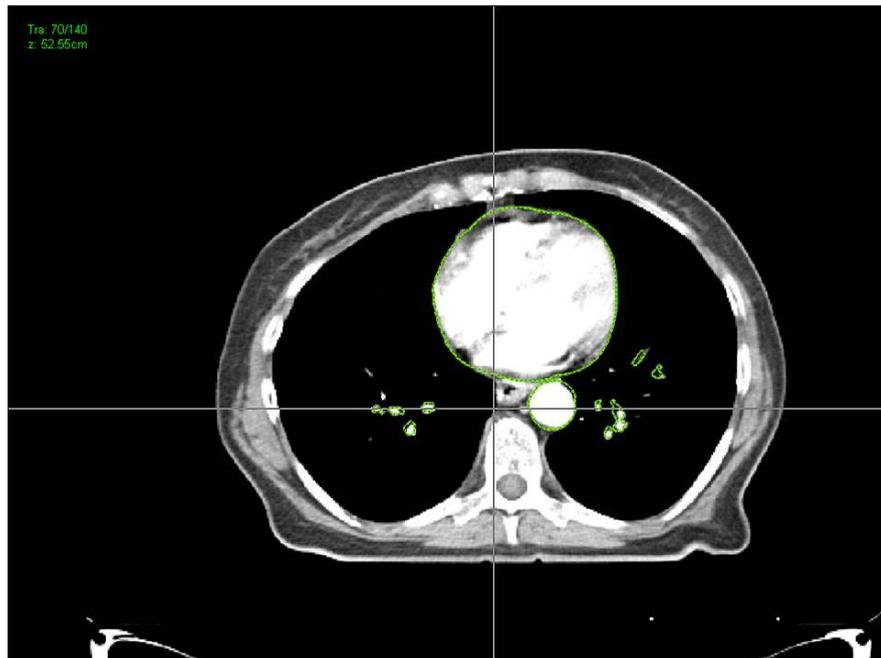
(a) CT slice before masking has been performed. Heart has regions of enhancement, due to presence of contrast, with variations in pixel values.
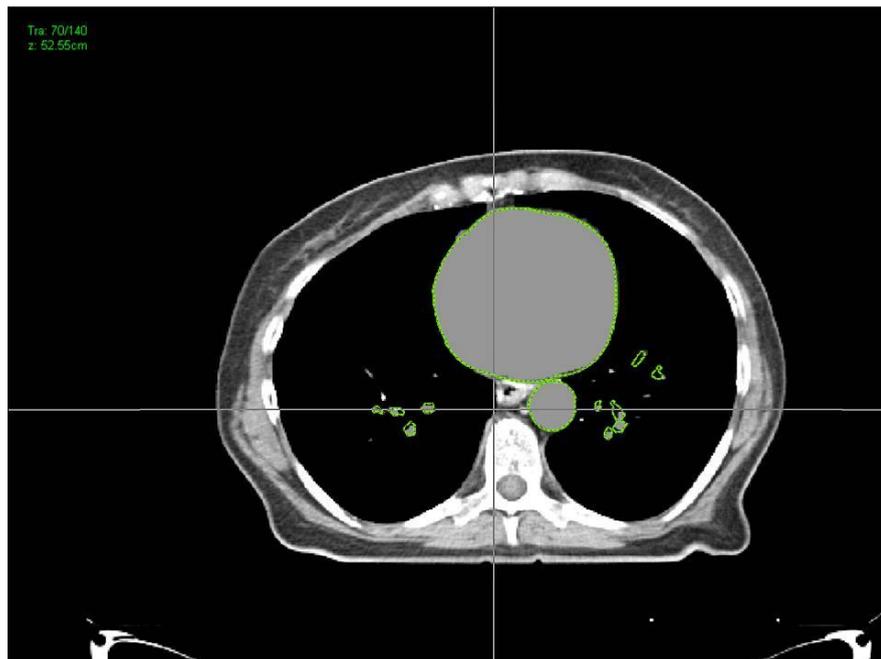
(b) CT slice after masking has been performed. Pixel values have been replaced with a value representing water density. Masked regions appear uniform.

Figure 3.12: Masking of intra-venous contrast in CT slice.

might be to use a slightly higher bulk density, more representative of soft tissues such as the heart.

In Oncentra, the bulk density can be thought of as being applied internally to the code; i.e. regions outlined as contrast will be treated as water equivalent only during the dose calculation, the CT DICOM files are not altered by this process. Since this work aims to perform MC based on DICOM and DICOM-RT files exported from the treatment planning system, a method was needed to treat regions of contrast as water equivalent in the MC, with ideally no user involvement.

One possibility would be a windowing technique, in which CT pixels within a certain HU range could be set to water equivalence. This method would be difficult, as the HU of the contrast regions coincides with that of trabecular bone (approximately 1250), as shown in Figure 3.13, so the method could treat bone as water equivalent, which would not be appropriate. This shows that any reliable approach would need to incorporate some knowledge of the organs or tissues in which the regions of enhancement were located. It was therefore decided to use the regions of contrast outlined in the planning system to determine what regions should be treated as contrast in the MC simulation. This approach has the advantage that exactly the same voxels will be treated as water in each dose calculation (Figure 3.12). The downsides are that the regions outlined as contrast are defined within a DICOM-RT Structure Set file, which the RTGrid has not needed in order to perform simulations before.

The basic methodology used is as follows:

- include the DICOM-RT Structure Set in the files uploaded to the RTGrid for a simulation;

- alter the RTGrid profile to read the Structure Set file;
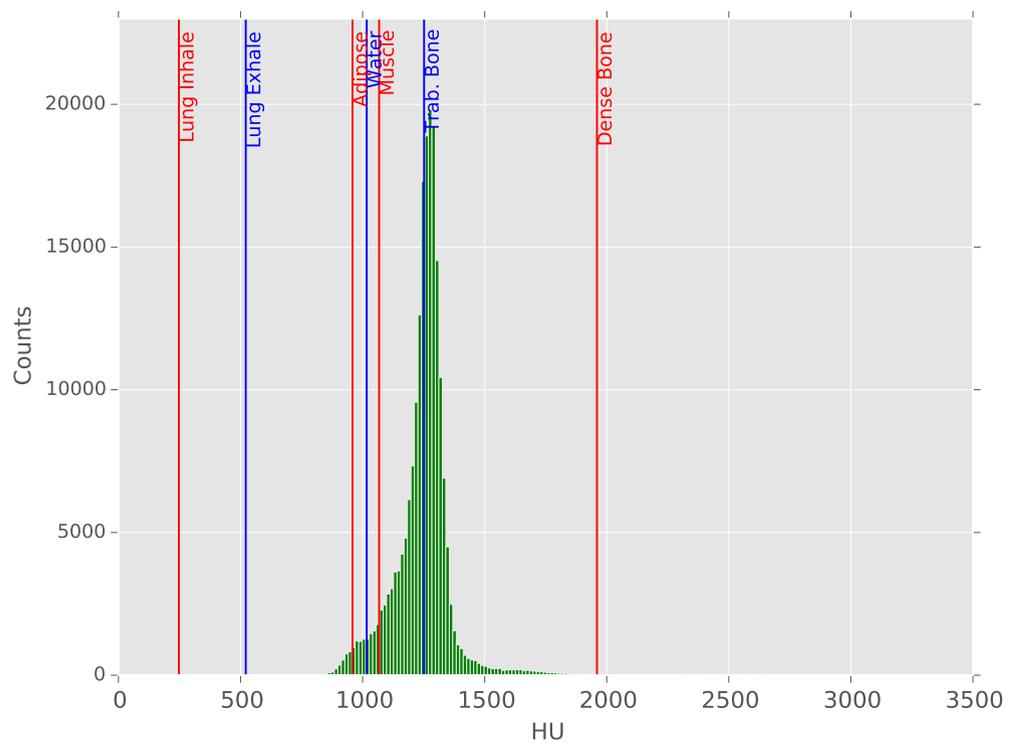
Figure 3.13: Histogram of Hounsfield Unit values within region of intra venous contrast, with HU of material inserts from CIRS 062M phantom for reference.

- for each structure with the bulk density set to 1 g.cm^{-3}, iterate over the contours;

- for each contour, determine which CT pixels are within the contour, and change the pixel value to the pixel value appropriate for water density;

- save the modified CT files;

- use the modified CT files to create egsphant file, not the originals.

As an aside, this approach would also allow bolus defined at the planning stage to be accounted for, as the bolus will have a bulk density of 1 g.cm^{-3} applied. This is not relevant for oesophageal plans, but is common in breast radiotherapy, and other treatments in which the PTV is close to the skin surface.

3.2.4.8 Superior and inferior extension

When performing a CT scan to be used for RT planning, it is common to only scan a portion of the body. This reduces the radiation dose from the CT scan to healthy tissue away from the tumour. However, this means that the treatment planning system will not be able to account for tissue outside the CT scan provided. Any tissue that will be irradiated, and any nearby OARs, must be included in the CT scan, to allow the dose in this region to be calculated, and the dose assessed by an oncologist. Some of the absorbed dose in the PTV or OARs may be due to radiation scattered from the tissue at the extremes of the scanned region. The calculated dose must be independent of the superior and inferior boundary of the CT scan, i.e. including additional slices should not alter the quoted dose within the PTV due to the additional scatter.

It may be difficult to set the boundaries of the CT scan to achieve this independence, while not scanning tissue unnecessarily. The solution provided by Oncentra is to allow a 'virtual' superior and inferior extension to be applied. In these

extended regions, the superior or inferior slice is 'copied', to provide additional scatter, out to a user defined distance. If, as for Head and Neck radiotherapy, the CT scan includes the superior aspect of the head, the superior extension would be set to 0 cm. For scans of the thorax and pelvis, it is common practice to set both the superior and inferior extensions to 5 cm. Oncentra will not calculate dose inside this 'virtual' region, but will account for scatter from this region into the scanned region. This approach may not be accurate in regions where the superior (or inferior) slice is not representative of the tissue outside the scan. For example, if the superior slice includes the shoulders, but outside the CT scan the patient outline narrows sharply to the neck. However, using an unrepresentative cross-section is expected to be better than not accounting for tissue outside the CT scan at all.

As with assigning a bulk density the iodine contrast, this correction is 'internal' to the dose calculation of Oncentra, the superior and inferior extension is not included in the files exported from Oncentra. Therefore a method was needed to mimic this approach for the MC simulations. Code was written in Pythonto parse the egsphant file, alter the number and boundaries of the voxels in the Superior-Inferior direction. The superior-most slice of the density matrices is then copied and appended to the top of the matrix. The inferior-most slice is similarly copied and appended to the bottom of the density matrix. This process was repeated for the material matrix. The modified data was written to the egsphant file, and the modified file used in the MC simulations. A complementary function was written to reverse these changes after the simulation, and to alter the .3ddose files (which take the same dimensions as the density and material matrices from the .egsphant file). This ensured that the dimensions of the files from the simulation match those of the DICOM-RT Dose files from the TPS.

3.2.4.9   Accounting for the treatment couch

The CT scan will invariably include the couch top of the CT scanner – the CT images will show the couch that the patient lay on during the CT scan. This will not be present during treatment (the patient will be lying on a different couch on the linac), so must be 'ignored' during the dose calculation (and the effect of the linac couch top should be included). For Oncentra, the CT couch is commonly removed from consideration by outlining the patient, but not the couch, and defining this outline as the external volume. The external volume is a special ROI; all pixels outside this volume are assumed to be air by Oncentra, regardless of the HU values.

Again, the solution that most closely mimicked the approach of Oncentra was felt to be the most suitable. The external contour was read from the DICOM-RT Structure Set file, a mask was constructed that was unity for all values inside the external contour, and zero for all pixels outside the contour (following the method used for masking contrast). If there is more than one contour per slice, a mask is generated per contour, and then the masks combined using an element-wise OR statement. The mask was then inverted, and used to change the pixel values for those pixels that are outside the external contour to -987, the mean of the upper and lower boundaries of the AIR bin shown in Table 3.2.

The standard practice in Velindre for accounting for the linac couch is to apply a simple attenuation correction factor to the MU for beams passing through the treatment couch. This is performed outside the planning system, so will not be included in either the doses calculated by Oncentra, or the doses from the MC simulations.

3.2.4.10 Post processing

After the MC simulation was complete, the 3ddose files produced by DOSXYZnrc
were downloaded from the RTGrid via scripts using SSH, along with associated
files such as .egsphant file, DICOM CT files and log files. The un-calibrated
.3ddose files were calibrated by applying calibration factors derived following the
method in section 3.2.4.5, and a factor to account for backscatter as described in
section 3.2.4.6. The files, now quoting dose in terms of energy deposited per unit
mass of material, were copied, and the SPR values discussed in section 3.2.4.3
applied, to convert to dose deposited per unit mass of water. The .3ddose files
were converted to DICOM-RT files, using code provided by Tony Teke of British
Columbia Cancer Agency, Vancouver Island Centre[6].

3.2.4.11 Calculation of Dose Volume Histogram parameters

For the four sets of DICOM-RT Dose files (PBE, CCE, MC D_{Mat} and MC D_{Wtr}),
DVH parameters for the 12 cases in the Pilot Study were calculated using a
Matlab[7] script that called functions from CERR.

3.2.5 Validation

In order to ensure that the MC simulations (including the generation of BEAMnrc
and DOSXYZnrc input files from DICOM and DICOM-RT files by the RTGrid),
and the post-processing tasks were being performed correctly, validation checks
were performed. This involved performing simulations in water phantoms, and
comparing these with commissioning data acquired at Velindre Cancer Centre
for a Varian 2100 linac. It is common practice when performing such verification

---

[6]Private Communication

[7]MATLAB 7.13, The MathWorks Inc., Natick, MA, 2011

simulations to prepare an egsphant file with high resolution voxels in regions of interest (i.e. along the central axis, for producing a PDD), and coarse voxels in the regions in which the dose will not be examined, but provide scatter to the beam axis. This approach reduces calculation time, compared with using high resolution voxels throughout the phantom. However, in order to test the functionality of the RTGrid, the method used by the RTGrid to create the egsphant files had to be followed, i.e. they were produced from the DICOM CT files. Since the DICOM CT files must have equally spaced voxels within a plane, the egsphant used in this work also had to have equally spaced voxels. Even using the RTGrid, simulating a cubic phantom of side 50 cm at 1 mm resolution, for tens of test fields would have taken an excessive amount of time. It was decided to perform simulations of fields of 7 cm by 7 cm and smaller in a 150 mm by 150 mm by 150 mm cube, of voxel size 1.5 mm by 1.5 mm by 1.5 mm. For fields up to 12 cm by 12 cm, a cubic phantom of side 200 mm and cubic voxels of side 2.0 mm was used. For fields up to 15 cm by 15 cm, a cubic phantom of side 300 mm and cubic voxels of side 3.0 mm was used. This would provide sufficient phantom material around the beam edges to provide scatter. The smallest phantom could be used to confirm the field edge was modelled accurately, and the larger field sizes in the larger phantoms could be used to show that a range of field sizes were modelled accurately.

The measurements were performed for commissioning of the linac prior to this work. They were acquired in a 50 cm by 50 cm by 50 cm plotting tank (water filled tank, which uses a motorised system to position a measuring device anywhere in the water tank). The measurements were performed with Scanditronix Photon Field Detector (Scanditronix-Wellhofer, Germany), which is a *p*-type silicon diode detector.

The MC parameters used for this validation work, and all subsequent simulations are shown in Table 3.3.

Option	Parameter
Boundary Crossing Algorithm	EXACT
Electron-step algorithm	PRESTA-II
Spin effects	On
Bremsstrahlung angular sampling	Simple
Bremsstrahlung cross sections	BH
Bound Compton scattering	Off
Pair angular sampling	Simple
Photoelectron angular sampling	Off
Rayleigh scattering	Off
Atomic relaxations	Off
Electron impact ionisation	Off

Table 3.3: BEAMnrc and DOSXYZnrc transport options used in simulations throughout this work.

3.2.5.1 Open field depth dose curves

PDDs were produced at 6 MV and 10 MV, for 5 cm by 5 cm, 10 cm by 10 cm and 15 cm by 15 cm fields, in varying sizes of phantom, as described above. In order to improve the statistical uncertainty, multiple copies of the same field were produced in the same DICOM-RT plan. The RTGrid then performed a simulation for each beam, using different random numbers for each, and the resulting .3ddose files combined. For each field, a plot of the measured PDD normalised to D_{MAX} is shown as a continuous blue line, and the MC values, also normalised to D_{MAX} are shown as green crosses. A second plot, showing the percentage difference between the measured and simulated values is included. This also shows the statistical uncertainty of the MC simulation as error bars, to allow the comparison of the statistical uncertainty and any discrepancy between measurement and simulation to be performed. Results in the first 5 mm are not presented, as the measurements

in this region not reliable, due to the lack of CPE.

For the 5 cm by 5 cm fields at 6 MV and 10 MV (Figure 3.14), virtually all simulated values are within 1% of measurement, aside from the shallowest and deepest points. The difference for the very deepest points can be attributed to the simulation having no material beyond 150 mm deep to provide backscatter, whereas the water phantom used for the measurements extended to approximately 500 mm deep.

The results for the 10 cm by 10 cm fields at 6 MV and 10 MV (Figure 3.15), show that the simulations are again within 1% of the measured values, aside from the deepest 20 mm of the phantom. This behaviour in the deepest 20 mm could be due to a flaw in the simulation, or the different backscatter conditions present for the measurement. However, as both 6 MV and 10 MV plots demonstrate similar behaviour, and the PDD for the 15 cm by 15 cm fields in a larger phantom show good agreement at 200 mm deep (Figure 3.16), the backscatter is the more likely explanation.

For the 15 cm by 15 cm fields at 6 MV and 10 MV (Figure 3.16), again agreement between simulation and measurement is within 1% until approximately 50 mm from the deepest point in the simulated phantom. As this is a repeat of the behaviour from other field sizes and phantom sizes, this can be explained by the difference in backscatter between measurements and simulation.

Overall, these results show that PDD can be accurately simulated for both 6 MV and 10 MV, for field sizes between 5 cm and 15 cm, for open fields.

3.2.5.2 Open field profiles

As for the PDDs, beam profiles were produced for 6 MV and 10 MV, for 5 cm by 5 cm, 10 cm by 10 cm and 15 cm by 15 cm field sizes, in varying size phantoms. Profiles were compared at 5 cm and 10 cm deep for all field sizes. Due to the
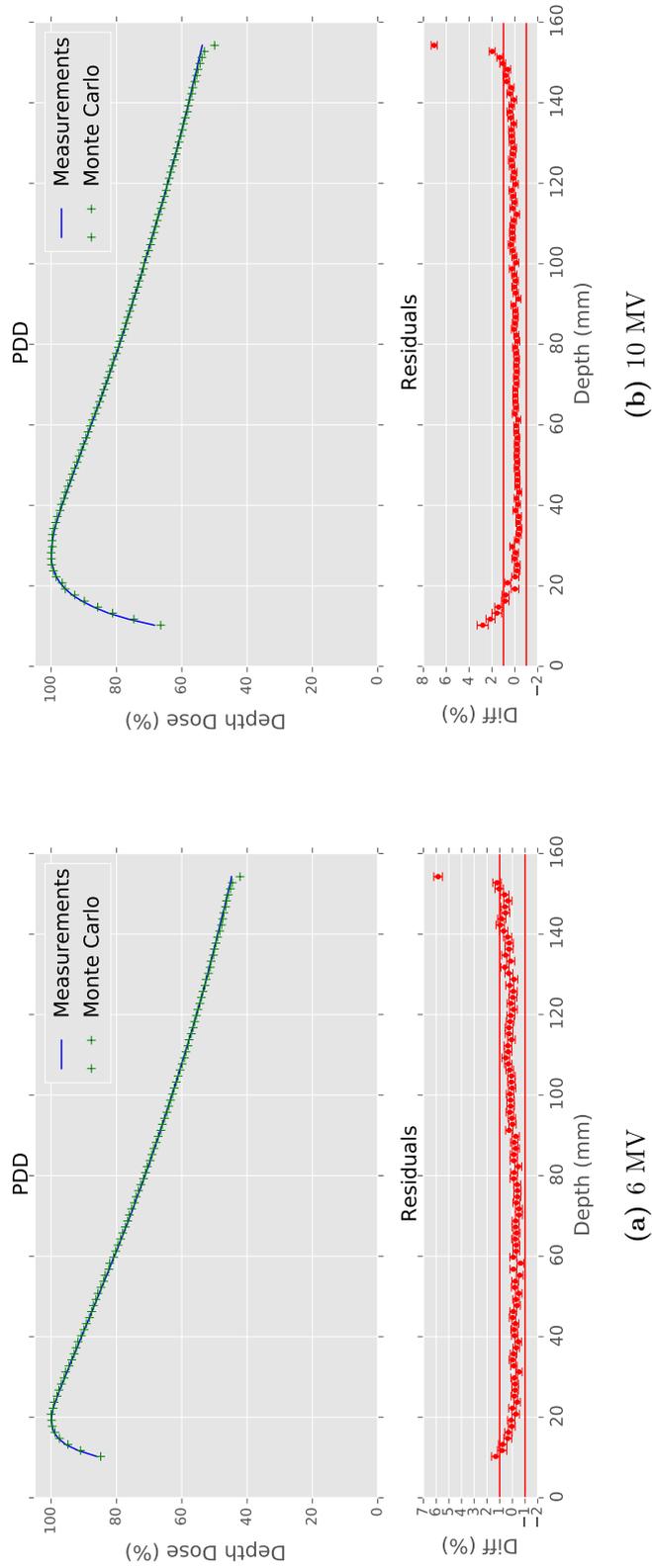
Figure 3.14: Comparison of MC calculated and commissioning PDD curves, for 5 cm by 5 cm field, at 6 MV and 10 MV.
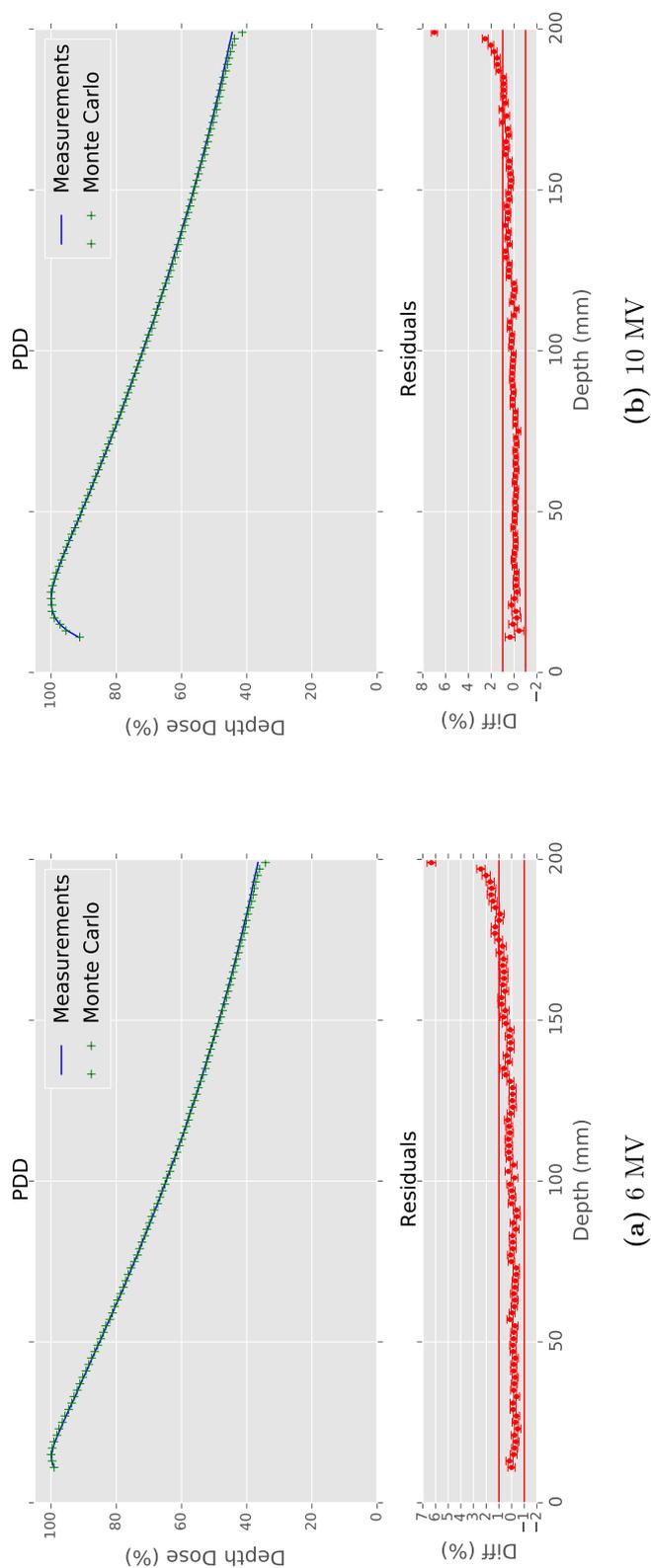
Figure 3.15: Comparison of MC calculated and commissioning PDD curves, for 10 cm by 10 cm field, at 6 MV and 10 MV.
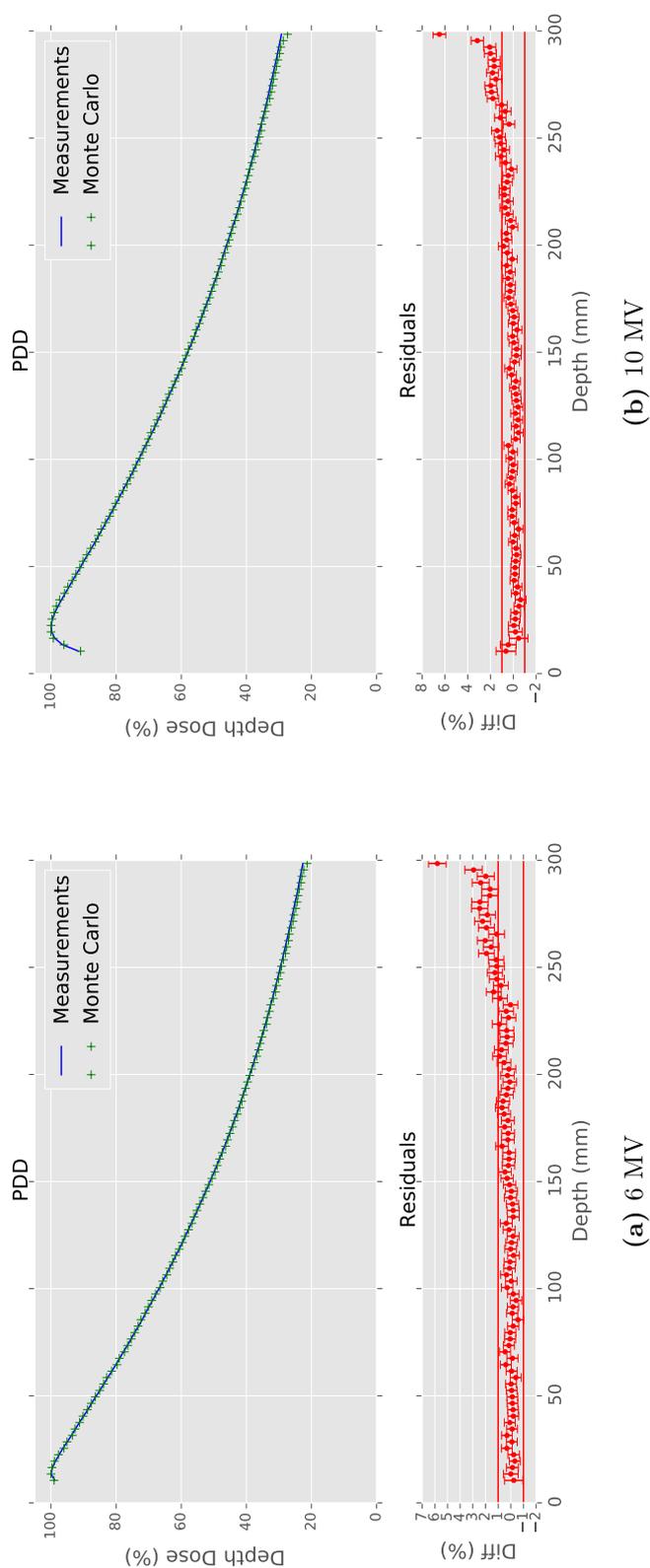
Figure 3.16: Comparison of MC calculated and commissioning PDD curves, for 15 cm by 15 cm field, at 6 MV and 10 MV.

larger phantom size used for the 15 cm by 15 cm beams, profiles at 20 cm deep are also presented. As for the PDD comparisons, for each profile two plots have been produced. The first shows measurements as a continuous blue line and MC results as green crosses, and the second shows differences between the two results, and the statistical uncertainty in the MC results.

For the 6 MV, 5 cm by 5 cm field (Figure 3.17), for both 5 cm and 10 cm deep, and in both directions (X-direction and Y-direction), the MC results are within 1% of the measurements, except in the penumbra where the dose drops sharply. It can be seen that the shape of the MC beam edge follows that of the measurements.

Figure 3.18 shows the results for the 6 MV, 10 cm by 10 cm field. The MC results are typically within 1% of the measurements, with a few points in the 'shoulders' of the profile being outside 1%, but well within 2%.

For the 6 MV, 15 cm by 15 cm field (Figure 3.19), most MC results are within 1% of measurement, even at 20 cm deep.

For the 10 MV, 5 cm by 5 cm field (Figure 3.20), the MC results are within 1% of the measurements, except in the penumbra. However, it can be seen that the shape of the MC beam edge follows that of the measurements. Since the 5 cm by 5 cm fields were simulated in a phantom with voxels of side 1.5 mm, these fields show the penumbra to the highest resolution used in this work.

Figure 3.21 shows the results for the 10 MV, 10 cm by 10 cm field. Once again, the MC results are typically within 1% of the measurements, with a few points being just outside 1%.

For the 10 MV, 15 cm by 15 cm field (Figure 3.22), most MC results are within 1% of measurement, and all are within 2%, even at 20 cm deep.
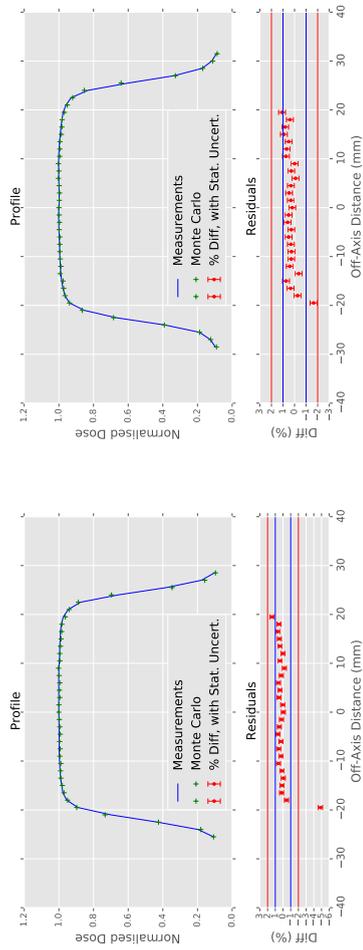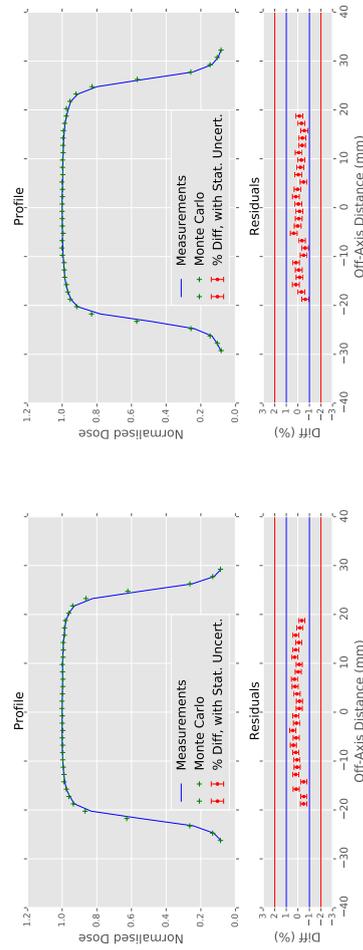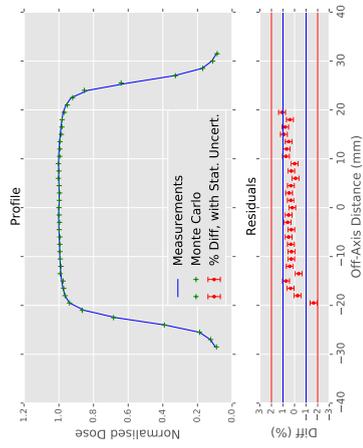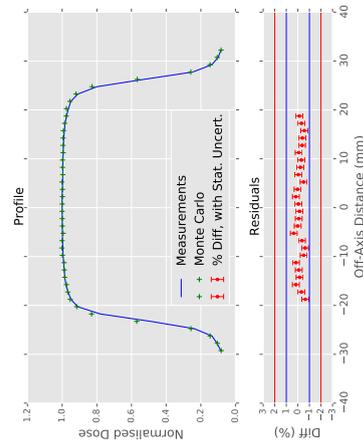
Figure 3.17: Comparison of MC calculated and commissioning profiles for 6 MV, 5 cm by 5 cm open field.
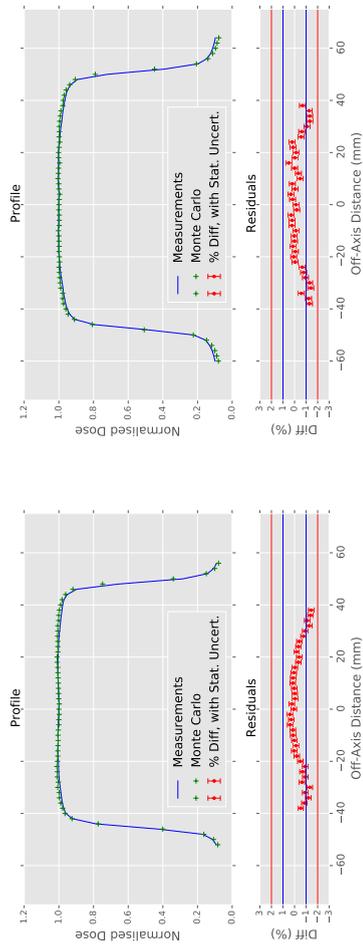
(a) 5 cm deep, X-direction
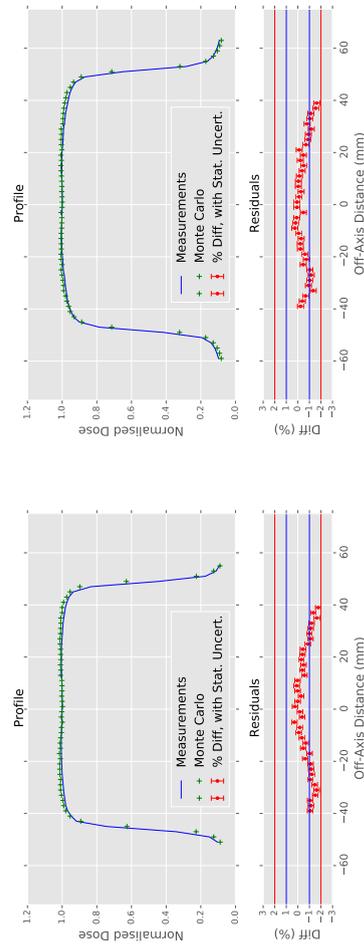
(b) 10 cm deep, X-direction
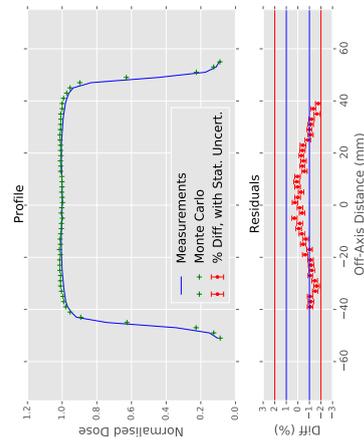
(c) 5 cm deep, Y-direction
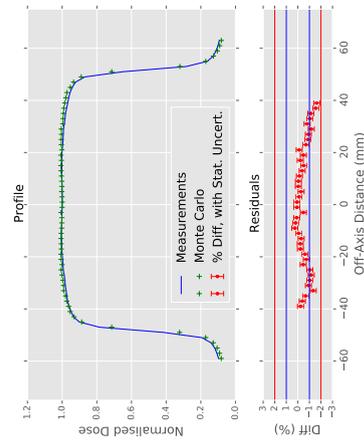
(d) 10 cm deep, Y-direction

Figure 3.18: Comparison of MC calculated and commissioning profiles for 6 MV, 10 cm by 10 cm open field.

Figure 3.19: Comparison of MC calculated and commissioning profiles for 6 MV, 15 cm by 15 cm open field.
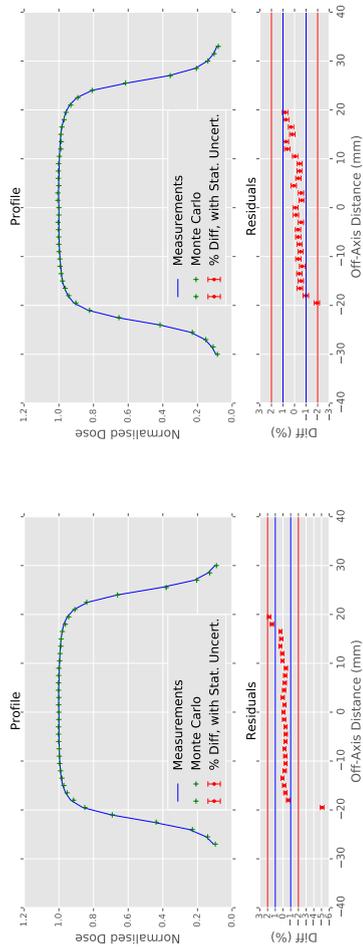
(a) 5 cm deep, X-direction
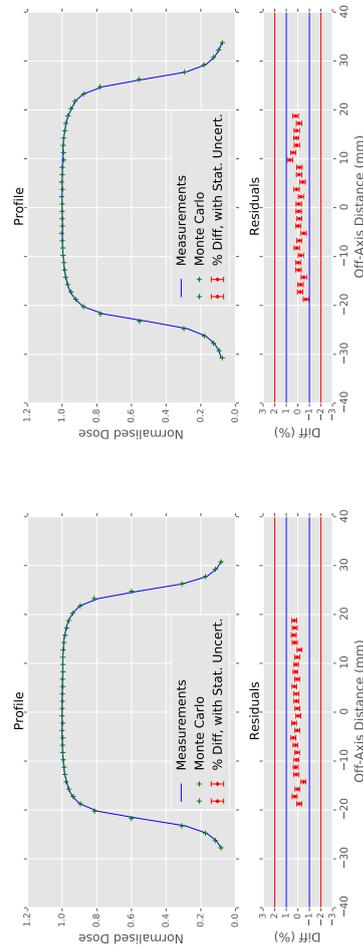
(b) 10 cm deep, X-direction
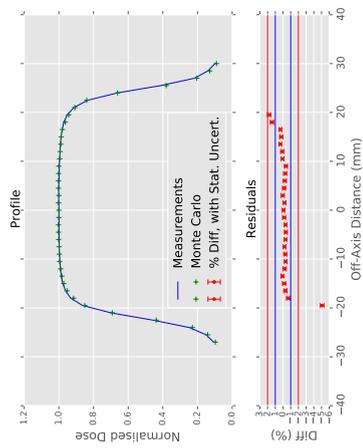
(c) 5 cm deep, Y-direction
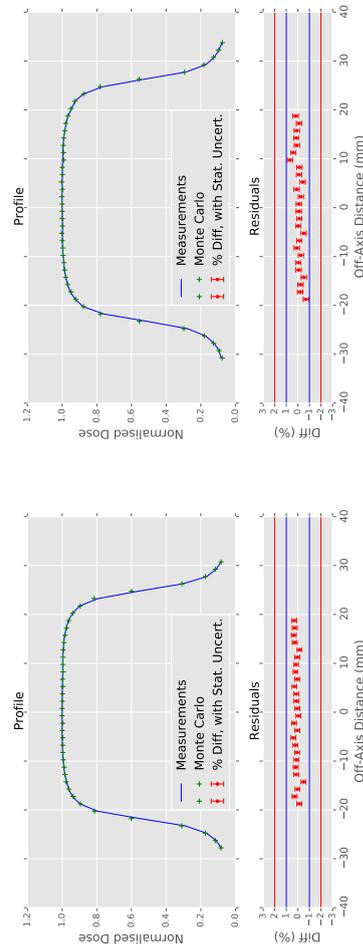
(d) 10 cm deep, Y-direction

Figure 3.20: Comparison of MC calculated and commissioning profiles for 10 MV, 5 cm by 5 cm open field.
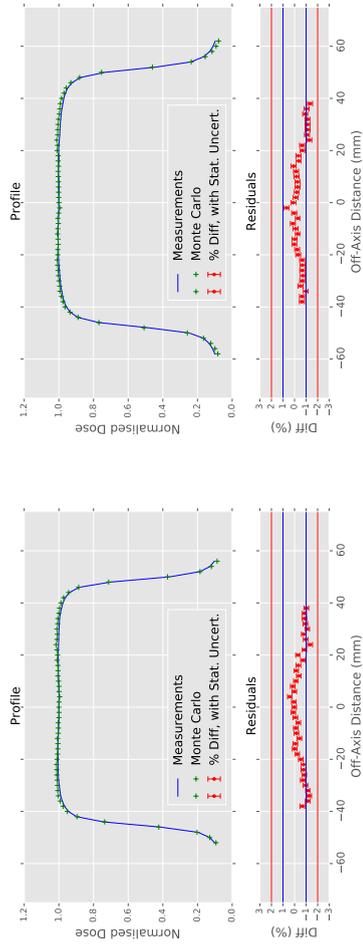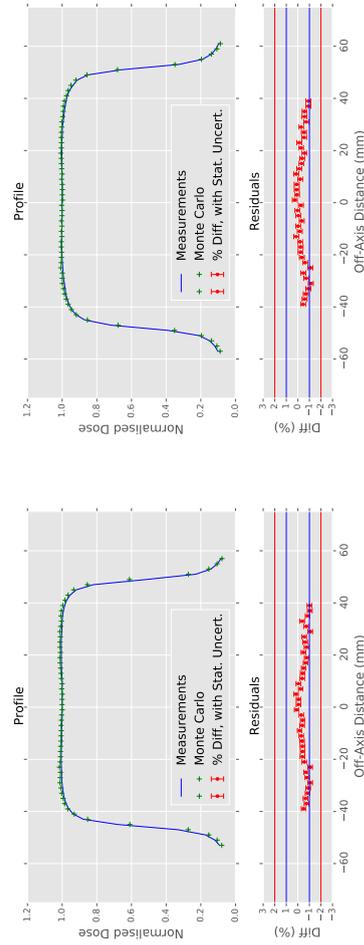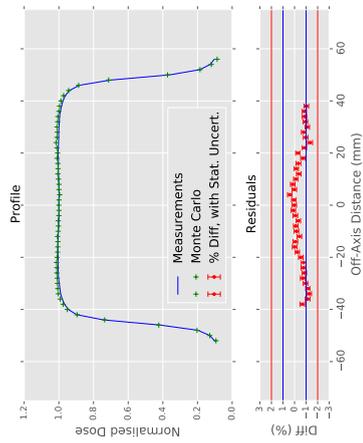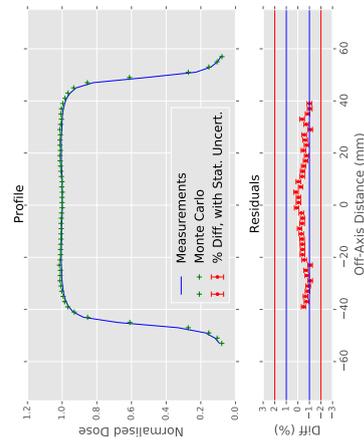
Figure 3.21: Comparison of MC calculated and commissioning profiles for 10 MV, 10 cm by 10 cm open field.
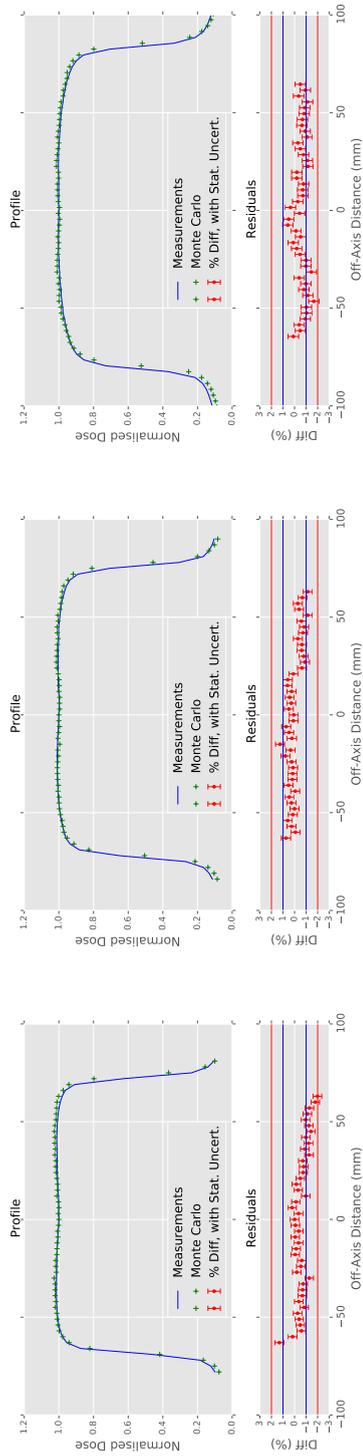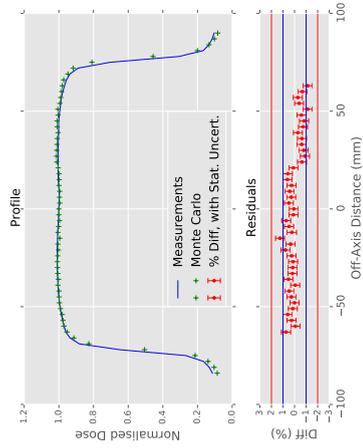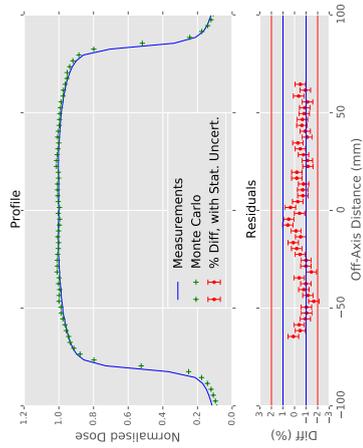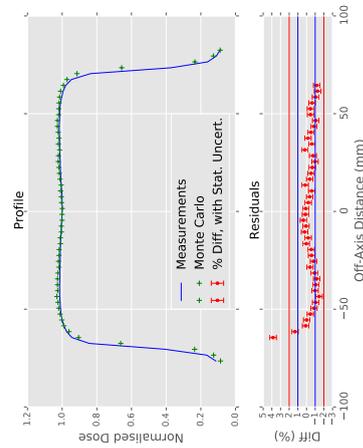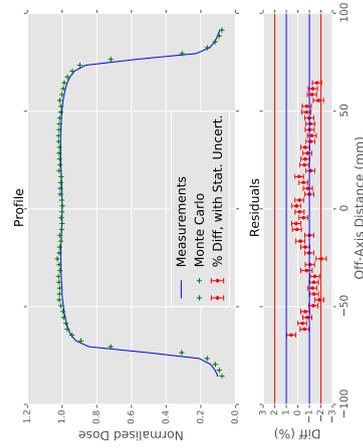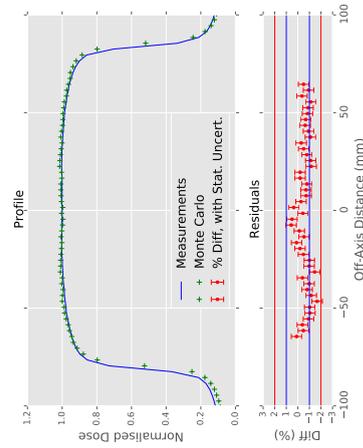
Figure 3.22: Comparison of MC calculated and commissioning profiles for 10 MV, 15 cm by 15 cm open field.

3.2.5.3   Enhanced Dynamic Wedge field depth dose curves

An advantage of EDW fields is that their PDDs are the same as those for the equivalent open field, so results will not be duplicated here.

3.2.5.4   Enhanced Dynamic Wedge field beam profiles

Beam profiles were produced for 6 MV and 10 MV, for 5 cm by 5 cm, 10 cm by 10 cm and 15 cm by 15 cm field sizes, in varying size phantoms, for the four EDW wedge angles, 15°, 30°, 45°, and 60°. For brevity, only the results for 10 cm by 10 cm fields are presented. The corresponding results for 5 cm by 5 cm and 15 cm by 15 cm fields are discussed in Appendix B. Verhaegen and Liu (2001) found the agreement between simulation and measurement to be within two standard deviations, with a standard deviation on the MC results of 1%. Therefore, agreement between simulation and measurement within 2% shall be considered the desired agreement.

For the 6 MV, 15° EDW, 10 cm by 10 cm field (Figure 3.23), results are typically within 1% of the measurements, with all points for the profiles in the X-direction being within 2% of measurements, which agrees with the results from Verhaegen and Liu (2001). For the profiles in the Y-direction, the direction of the wedge, some points in the 'toe' of the wedge are outside 2%, but this is comparable with the results of Verhaegen and Liu (2001).

Figure 3.24 shows the results for 6 MV, 30° EDW, 10 cm by 10 cm field. Again results are well within 2% aside for a small region in the toe of the wedge.

Figures 3.25 and 3.26 show the results for 6 MV, 10 cm by 10 cm fields for 45° EDW, and 60° EDW. Again results are within 2%, apart from differences in the toe of the wedges. However, overall agreement between MC and measurement is within the desired tolerance.
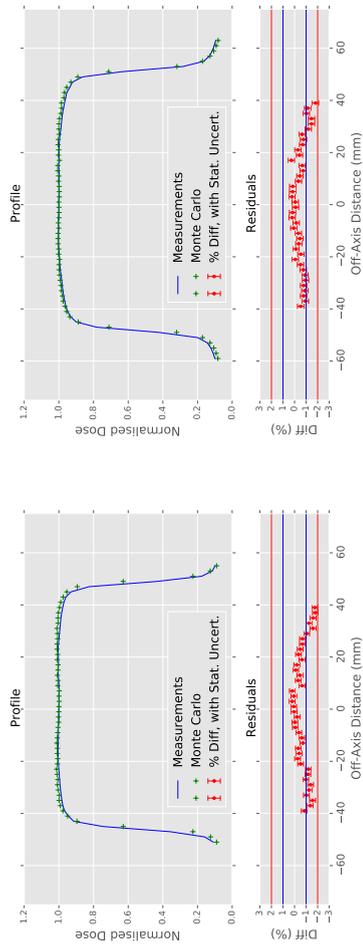
(a) 5 cm deep, X-direction
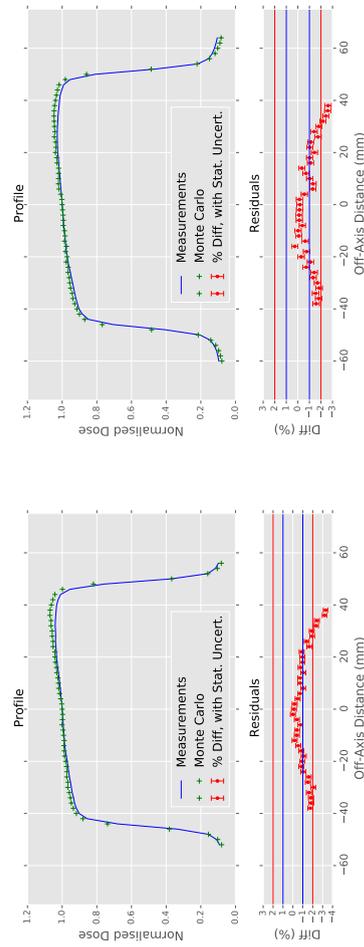
(b) 10 cm deep, X-direction
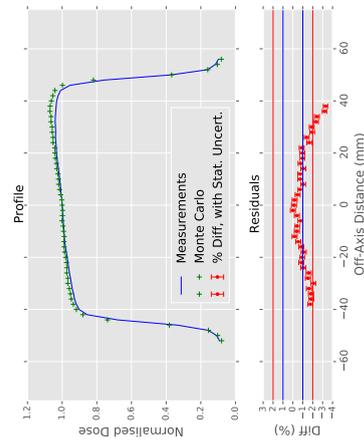
(c) 5 cm deep, Y-direction
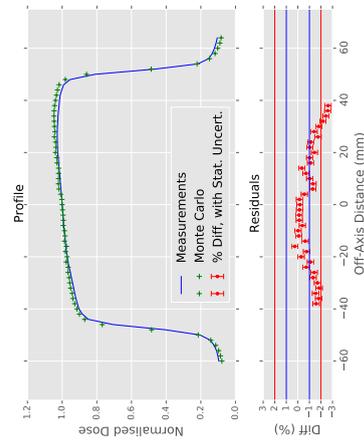
(d) 10 cm deep, Y-direction

Figure 3.23: Comparison of MC calculated and commissioning profiles for 6 MV, 10 cm by 10 cm 15° EDW field.
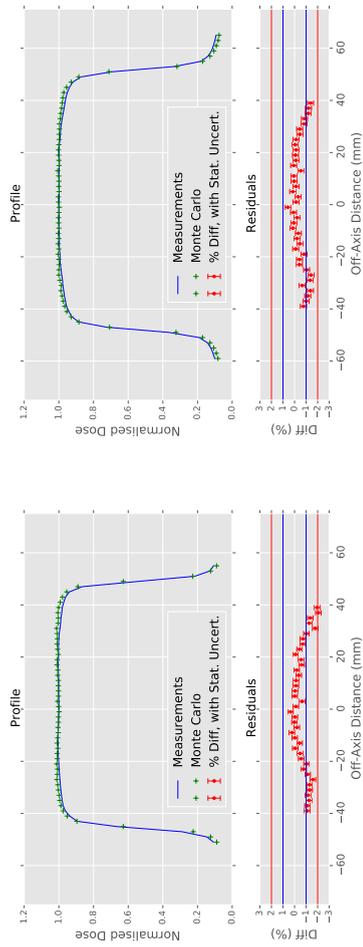
(a) 5 cm deep, X-direction
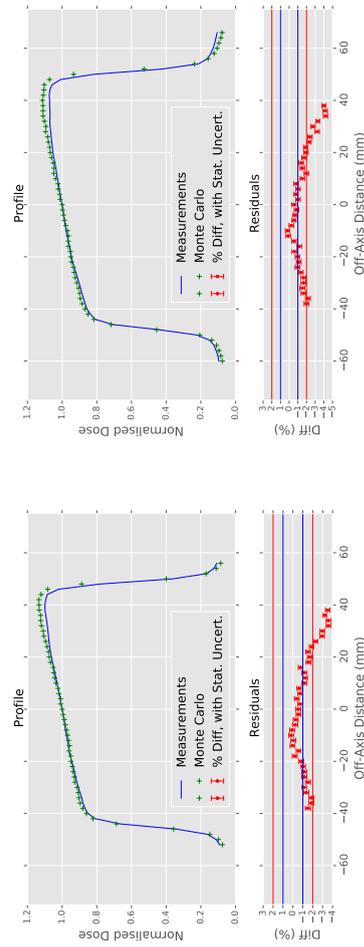
(b) 10 cm deep, X-direction
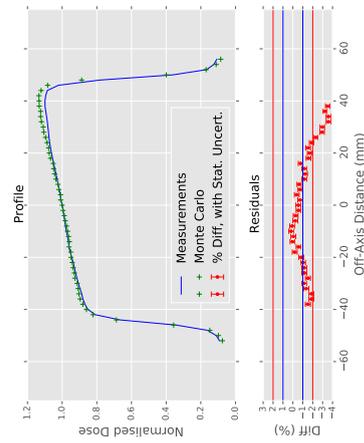
(c) 5 cm deep, Y-direction
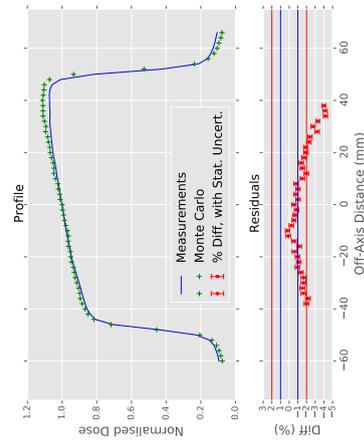
(d) 10 cm deep, Y-direction

Figure 3.24: Comparison of MC calculated and commissioning profiles for 6 MV, 10 cm by 10 cm 30° EDW field.
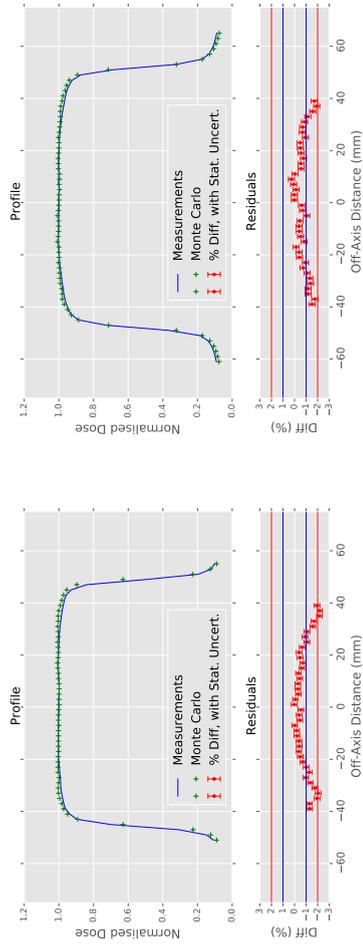
(a) 5 cm deep, X-direction
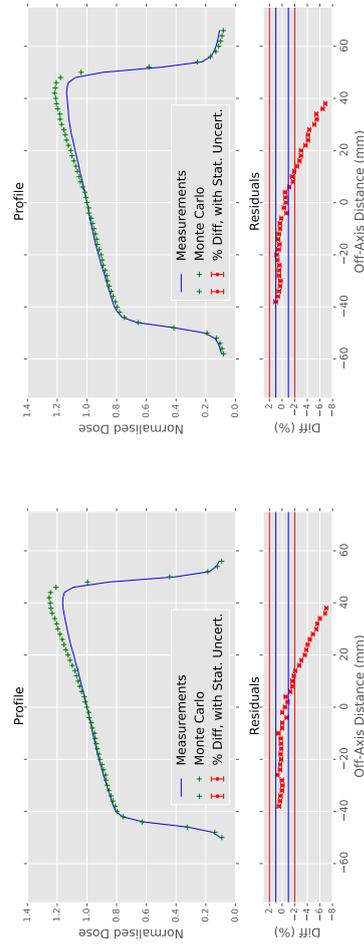
(b) 10 cm deep, X-direction
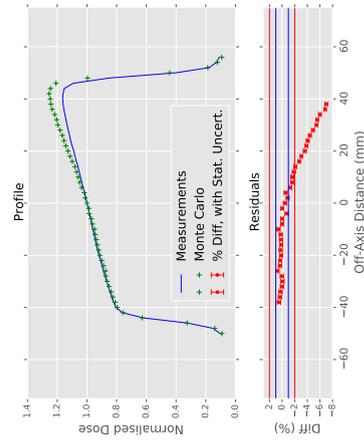
(c) 5 cm deep, Y-direction
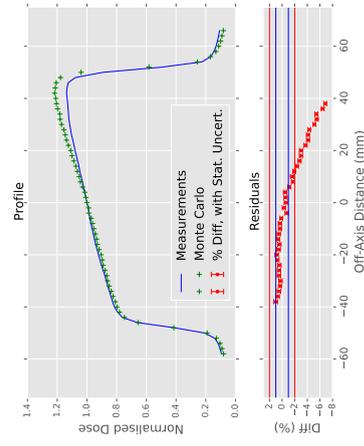
(d) 10 cm deep, Y-direction

Figure 3.25: Comparison of MC calculated and commissioning profiles for 6 MV, 10 cm by 10 cm 45° EDW field.
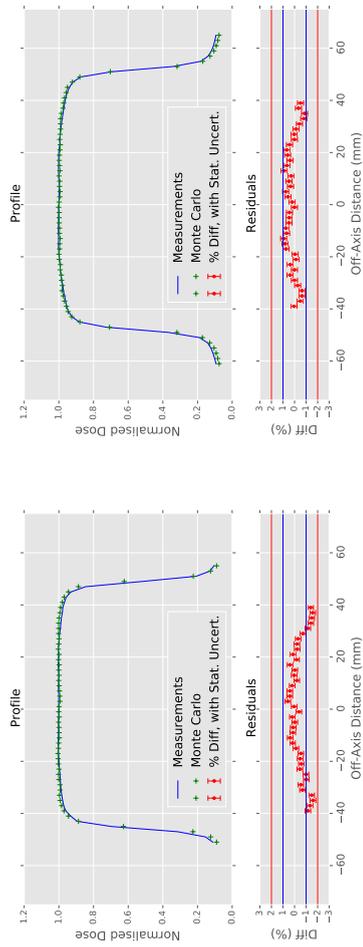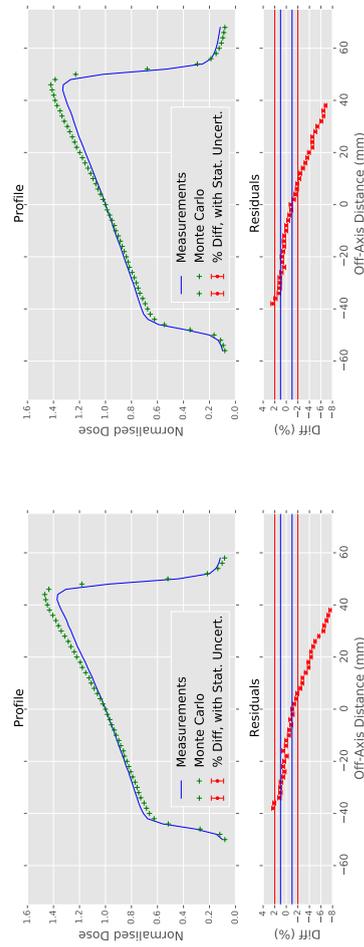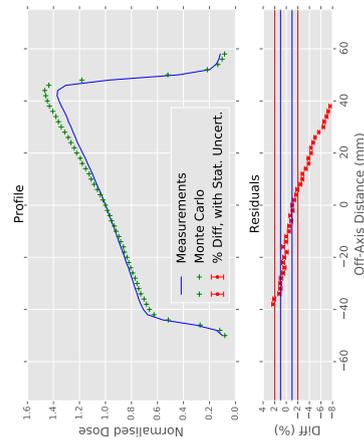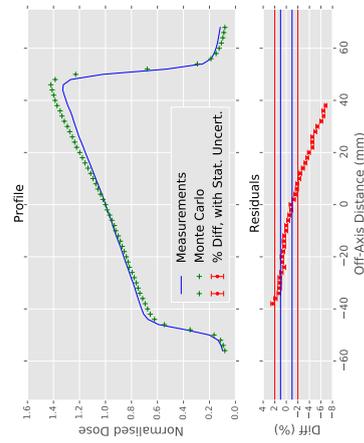
Figure 3.26: Comparison of MC calculated and commissioning profiles for 6 MV, 10 cm by 10 cm 60° EDW field.
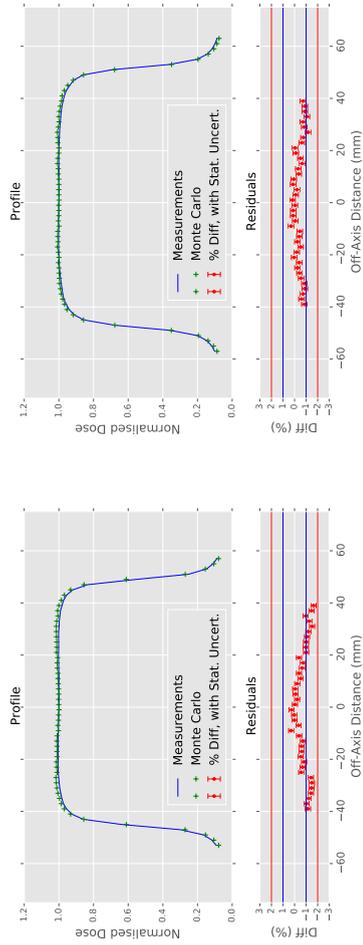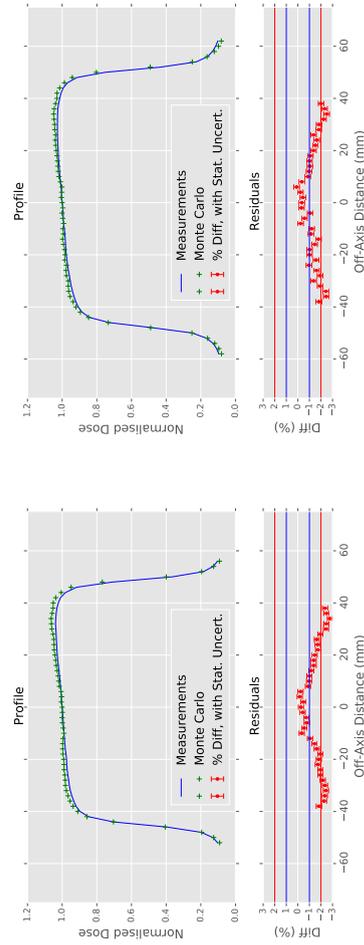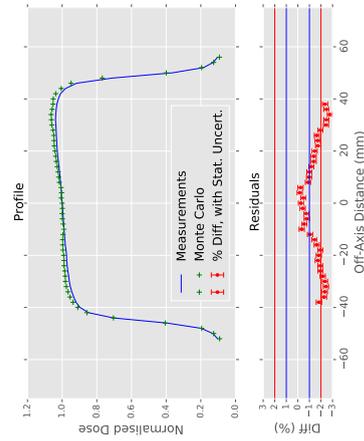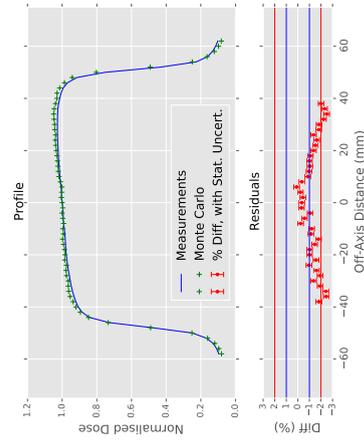
Figure 3.27: Comparison of MC calculated and commissioning profiles for 10 MV, 10 cm by 10 cm 15° EDW field.

The results for the equivalent 10 MV fields (Figures 3.27, 3.28, 3.29 and 3.30) are comparable to those for the 6 MV fields, agreement is typically within 1% - 2%, except for the toe of the wedge, with discrepancies in the toe region increasing with wedge angle. However, the results for 45° EDW, and 60° EDW at 10 MV are better than the equivalent 6 MV field sizes.

Overall, the agreement for EDW fields is comparable to the results of Verhaegen and Liu (2001), outside the very toe end of the EDW fields. This is expected from the original work on the component module used for EDW simulations (Downes, 2010). Furthermore, when different fields are combined in a clinical plan, the effect of a single field will be reduced.

3.2.5.5 Check of absolute dose

The Absolute Dose at the isocentre for most fields in Tables 3.4 and 3.5 show most fields are accurate to within 1%. Some of the smallest and largest fields show differences between 1% and 2%, with the worst result of 2.3% for the 15 cm by 15 cm 60° EDW field at 6 MV. This field is larger than a typical field used for oesophageal plans. Overall, these results show good agreement between simulation and measurement.

3.2.6   Case selection criteria

Radical RT treatment plans created in Velindre are entered into a custom task management system, which includes the treatment site of the plan, and the date the plan is due by. Structured Query Language (SQL) was used to extract a list of all oesophagus plans in this system, ordered by the due date. The pilot study comprised the first 12 cases meeting the following criteria:

- the CT was performed at Velindre Cancer Centre, using the
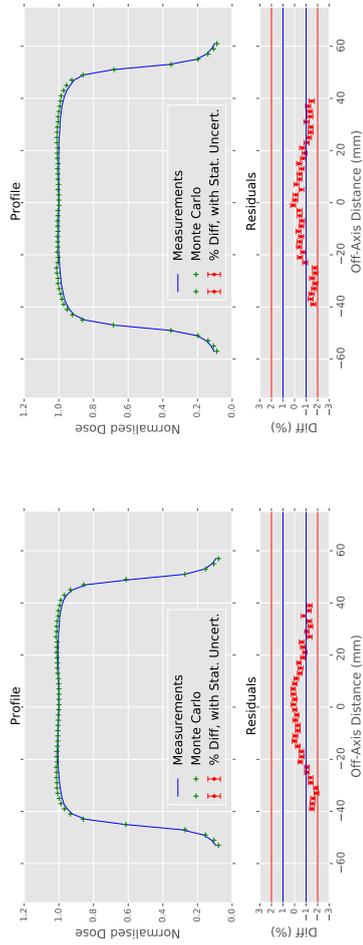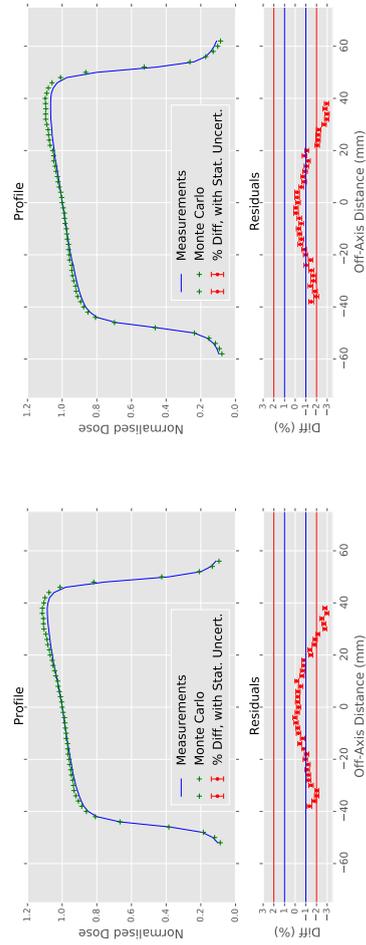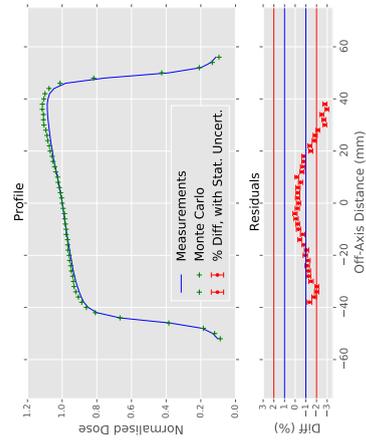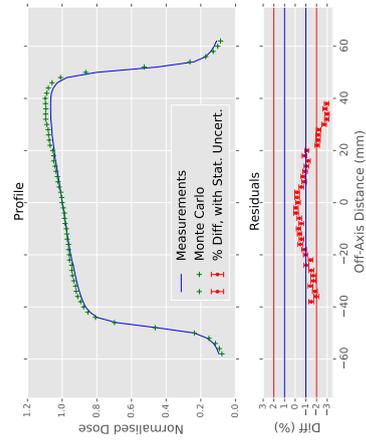
Figure 3.28: Comparison of MC calculated and commissioning profiles for 10 MV, 10 cm by 10 cm 30° EDW field.
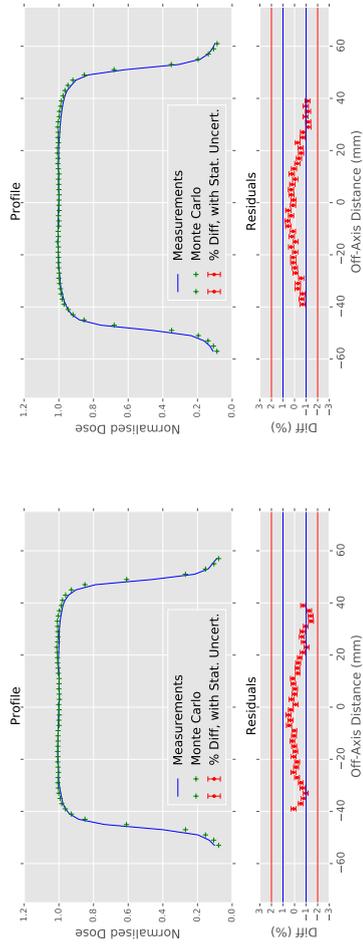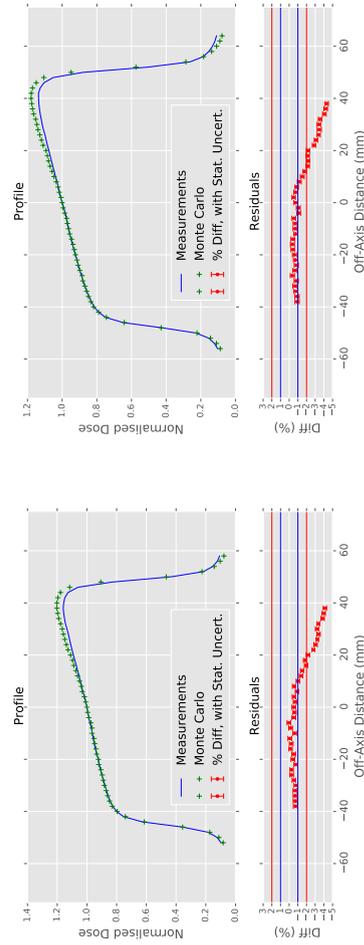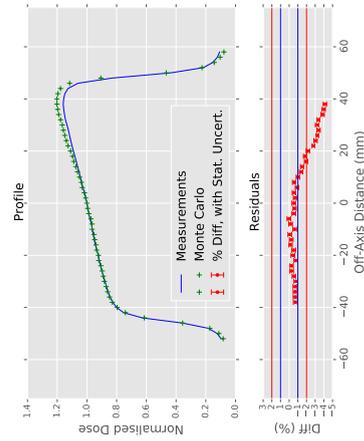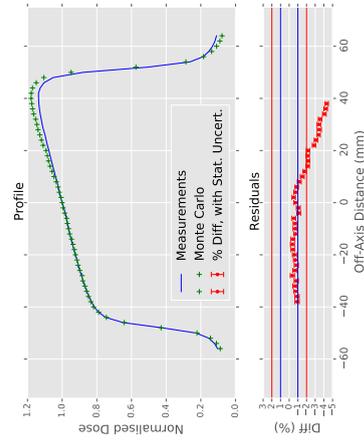
Figure 3.29: Comparison of MC calculated and commissioning profiles for 10 MV, 10 cm by 10 cm 45° EDW field.
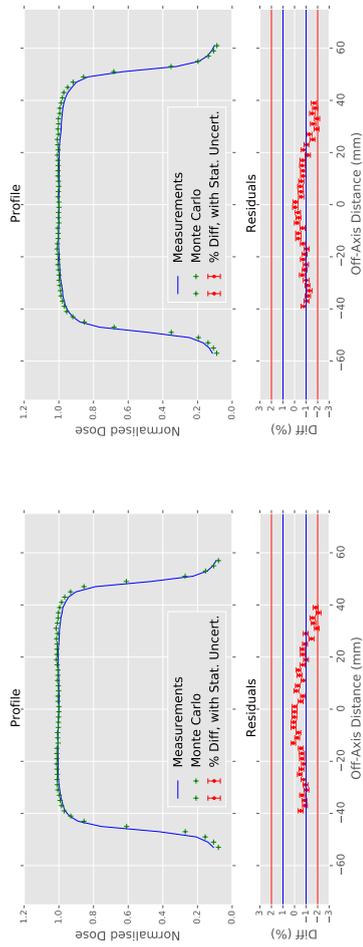
(a) 5 cm deep, X-direction

(b) 10 cm deep, X-direction

(c) 5 cm deep, Y-direction

(d) 10 cm deep, Y-direction

(a) 5 cm deep, X-direction
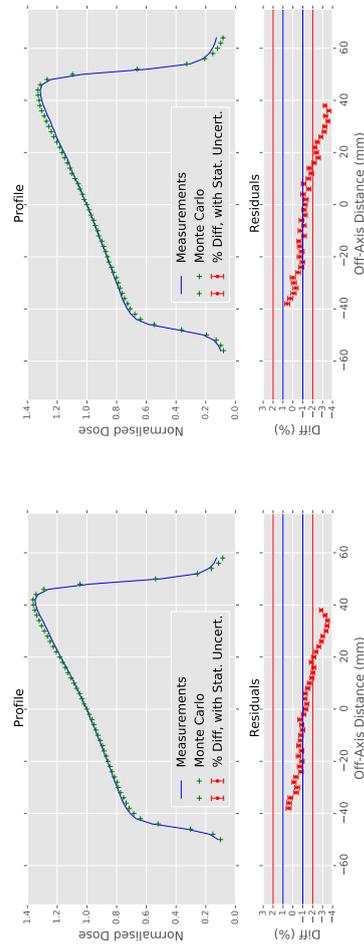
(b) 10 cm deep, X-direction
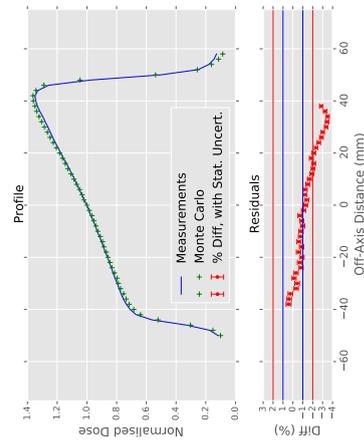
(c) 5 cm deep, Y-direction
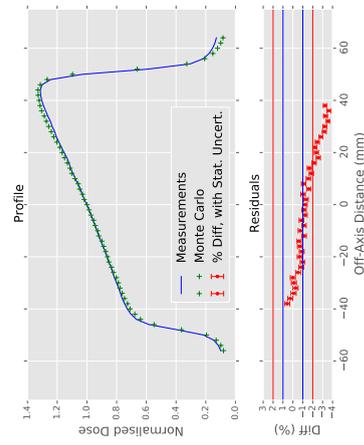
(d) 10 cm deep, Y-direction

Figure 3.30: Comparison of MC calculated and commissioning profiles for 10 MV, 10 cm by 10 cm 60° EDW field.

5 cm by 5 cm					
Field	**Raw**	**Cal Factor**	**Liu**	**MC Dose**	**Expected Dose**	**% Diff**
Open	8.76E-17	1.22E-16	1.00	0.722	0.710	1.8
15 EDW	8.41E-17	1.22E-16	1.00	0.690	0.687	0.5
30 EDW	8.01E-17	1.21E-16	1.00	0.664	0.663	0.2
45 EDW	7.60E-17	1.19E-16	1.00	0.643	0.632	1.7
60 EDW	6.91E-17	1.18E-16	1.00	0.588	0.585	0.6

7 cm by 7 cm					
Field	**Raw**	**Cal Factor**	**Liu**	**MC Dose**	**Expected Dose**	**% Diff**
Open	9.16E-17	1.22E-16	0.998	0.755	0.751	0.5
15 EDW	8.68E-17	1.22E-16	0.998	0.712	0.716	-0.6
30 EDW	8.14E-17	1.21E-16	0.997	0.675	0.679	-0.5
45 EDW	7.52E-17	1.19E-16	0.997	0.637	0.633	0.6
60 EDW	6.62E-17	1.18E-16	0.996	0.563	0.566	-0.5

10 cm by 10 cm					
Field	**Raw**	**Cal Factor**	**Liu**	**MC Dose**	**Expected Dose**	**% Diff**
Open	9.67E-17	1.22E-16	1.000	0.795	0.795	0.0
15 EDW	8.88E-17	1.22E-16	0.999	0.728	0.736	-1.1
30 EDW	8.12E-17	1.21E-16	0.999	0.672	0.677	-0.7
45 EDW	7.21E-17	1.19E-16	0.998	0.609	0.610	-0.1
60 EDW	6.10E-17	1.18E-16	0.998	0.518	0.520	-0.3

12 cm by 12 cm					
Field	**Raw**	**Cal Factor**	**Liu**	**MC Dose**	**Expected Dose**	**% Diff**
Open	9.98E-17	1.22E-16	1.000	0.820	0.817	0.3
15 EDW	8.95E-17	1.22E-16	1.000	0.733	0.742	-1.2
30 EDW	8.02E-17	1.21E-16	0.999	0.664	0.668	-0.7
45 EDW	7.00E-17	1.19E-16	0.999	0.592	0.589	0.4
60 EDW	5.71E-17	1.18E-16	0.999	0.484	0.489	-0.9

15 cm by 15 cm					
Field	**Raw**	**Cal Factor**	**Liu**	**MC Dose**	**Expected Dose**	**% Diff**
Open	1.02E-16	1.22E-16	1.000	0.840	0.844	-0.5
15 EDW	8.86E-17	1.22E-16	1.000	0.726	0.741	-2.0
30 EDW	7.66E-17	1.21E-16	1.000	0.634	0.647	-2.1
45 EDW	6.50E-17	1.19E-16	1.000	0.549	0.552	-0.6
60 EDW	5.08E-17	1.18E-16	0.999	0.431	0.441	-2.3

Table 3.4: Comparison of absolute doses from Monte Carlo compared with measurements, at 6 MV, for different fields, where Raw is the uncorrected value from the MC simulation, Cal Factor is the calibration factor, Liu is the backscatter correction for the field, MC Dose is the raw MC value multiplied by the Cal Factor and Liu Factor, the Expected Dose is the dose calculated from dosimetry tables in Velindre, and %diff is the difference between the MC and expected doses, expressed as a percentage.

5 cm by 5 cm					
Field	**Raw**	**Cal Factor**	**Liu**	**MC Dose**	**Published Dose**	**% Diff**
Open	2.24E-16	2.79E-16	0.997	0.806	0.798	1.0
15 EDW	2.17E-16	2.75E-16	0.997	0.791	0.776	2.0
30 EDW	2.08E-16	2.75E-16	0.997	0.760	0.755	0.8
45 EDW	1.99E-16	2.73E-16	0.997	0.733	0.725	1.1
60 EDW	1.84E-16	2.69E-16	0.997	0.687	0.679	1.2

7 cm by 7 cm					
Field	**Raw**	**Cal Factor**	**Liu**	**MC Dose**	**Published Dose**	**% Diff**
Open	2.35E-16	2.79E-16	0.998	0.842	0.835	0.8
15 EDW	2.24E-16	2.75E-16	0.998	0.814	0.802	1.5
30 EDW	2.11E-16	2.75E-16	0.998	0.770	0.768	0.3
45 EDW	1.97E-16	2.73E-16	0.997	0.725	0.725	0.0
60 EDW	1.76E-16	2.69E-16	0.997	0.658	0.659	-0.1

10 cm by 10 cm					
Field	**Raw**	**Cal Factor**	**Liu**	**MC Dose**	**Published Dose**	**% Diff**
Open	2.44E-16	2.79E-16	1.000	0.875	0.875	0.0
15 EDW	2.27E-16	2.75E-16	0.999	0.826	0.820	0.8
30 EDW	2.10E-16	2.75E-16	0.999	0.763	0.766	-0.3
45 EDW	1.91E-16	2.73E-16	0.999	0.700	0.702	-0.3
60 EDW	1.64E-16	2.69E-16	0.998	0.613	0.612	0.1

12 cm by 12 cm					
Field	**Raw**	**Cal Factor**	**Liu**	**MC Dose**	**Published Dose**	**% Diff**
Open	2.50E-16	2.79E-16	1.001	0.893	0.895	-0.2
15 EDW	2.29E-16	2.75E-16	1.000	0.832	0.825	0.8
30 EDW	2.08E-16	2.75E-16	1	0.757	0.758	-0.1
45 EDW	1.84E-16	2.73E-16	0.999	0.676	0.681	-0.8
60 EDW	1.56E-16	2.69E-16	0.999	0.580	0.579	0.1

15 cm by 15 cm					
Field	**Raw**	**Cal Factor**	**Liu**	**MC Dose**	**Published Dose**	**% Diff**
Open	2.55E-16	2.79E-16	1.003	0.910	0.919	-0.9
15 EDW	2.26E-16	2.75E-16	1.002	0.818	0.824	-0.8
30 EDW	2.00E-16	2.75E-16	1.000	0.729	0.738	-1.3
45 EDW	1.74E-16	2.73E-16	1.000	0.637	0.645	-1.4
60 EDW	1.41E-16	2.69E-16	1.000	0.526	0.529	-0.7

Table 3.5: Comparison of absolute doses from Monte Carlo compared with measurements, at 10 MV, for different fields, where Raw is the uncorrected value from the MC simulation, Cal Factor is the calibration factor, Liu is the backscatter correction for the field, MC Dose is the raw MC value multiplied by the Cal Factor and Liu Factor, the Expected Dose is the dose calculated from dosimetry tables in Velindre, and %diff is the difference between the MC and expected doses, expressed as a percentage.

'RTChest3mm_Contrast' protocol for which OesRAMP is defined;

- the scan orientation was Head First Supine (i.e. face up, head towards the linac gantry);

- the SCOPE trial volumes GTV, CTV-A, CTV-B and PTV were present;

- the linac model used in Oncentra was 'Varian_2100_1217', a Varian Clinic 2100CD with the Millennium 120-leaf MLC;

- the plan was prescribed to 50 Gy in 25 fractions;

- the plan used a CFRT technique, not a IMRT or VMAT technique;

- no metal implants in the CT, determined by bulk densities defined with density greater than 1

12 cases were used to ensure the pilot study would be at least 10% of the main study; it would have been impractical to identify and simulate 120 cases matching the above criteria. Over time, the linac model used for oesophageal plans in Velindre Cancer Centre changed from a Varian Clinac 2100CD with an 80-leaf MLC, to the model used for this work, to an Elekta Synergy linac.

3.3   Results

To present the results of the 12 simulations used in the pilot study, the concept of the table-graphic, or slopegraph, (Tufte, 2001) has been adapted. The table-graphic is a combination of tabulated data, and a graphical representation of the data. It allows the rank of data to be visualised for the different dose calculation methods, and the change in each dataset to be seen. Also, the slope of the lines from one dataset to the next indicates the overall trend, as well as highlighting any cases that do not follow the general trend.

For each DVH parameter as defined in SCOPE1 (aside from the kidneys and liver, because they are not always involved), two table-graphics have been produced. The first compares the DVH parameter from the PBE dose calculations with the MC D_{Wtr} calculations, as PBE reports dose to water. The second compares PBE, CCE and MC D_{Mat}. CCE reports dose to material (section 2.4.5.3), so it is natural to compare it with and MC D_{Mat}. Even though PBE and CCE report dose to different media, it has become commonplace to compare dose from these two algorithms; PBE and CCE are compared here as a control, to place the differences from CCE to MC D_{Mat} into context. In each table-graphic, lines are used to join the results for the same case calculated with different methods; the line is coloured blue if the value falls, red if the value rises, and black if the value remains the same. Grey lines are used to link labels with points - in many cases the data points are too close together to allow labels to be close to the data points.

Figure 3.31(a) shows that the PTV D99% for the MC D_{Wtr} calculations has a greater spread of values than the PBE calculations. It can also be seen that there is a general trend for the PTV D99% to be lower in the MC D_{Wtr} calculation. However, the magnitude of this change is variable, as shown by the trend-lines crossing over one another in many cases. Similarly Figure 3.31(b) shows the greater spread of PTV D99% as calculated with CCE than those with PBE. In many cases the reduction in PTV D99% when recalculating PBE with CCE is reproduced when re-calculating with MC D_{Wtr}. However, there are some exceptions to this, such as cases #5 and #12.

Figure 3.32(a) also shows the overall trend is for PTV V95% to decrease when re-calculated with MC D_{Wtr} compared to the PBE distribution. The spread in values for MC D_{Wtr} is again wider that the spread in values for PBE. Figure 3.32(b) shows the overall trend is for PTV V95% to decrease when re-calculated with CCE compared to the PBE distribution. This trend is often continued when
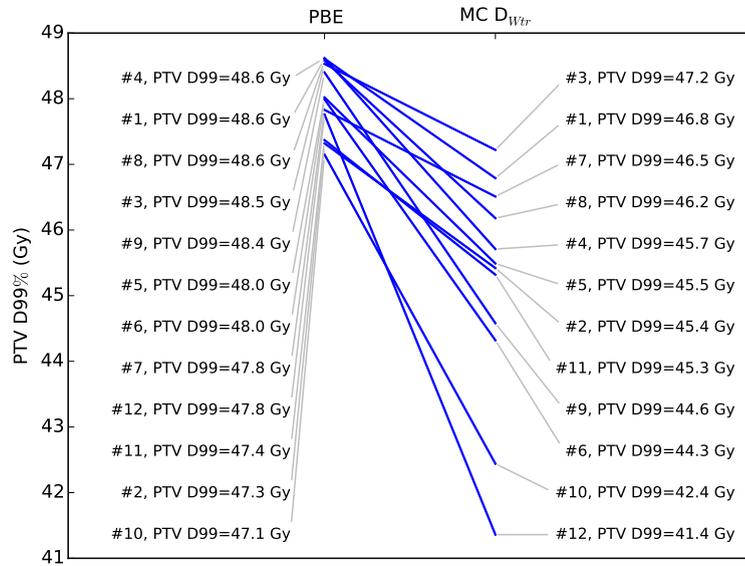
(a) PBE versus MC D$_{Wtr}$ for PTV D99%
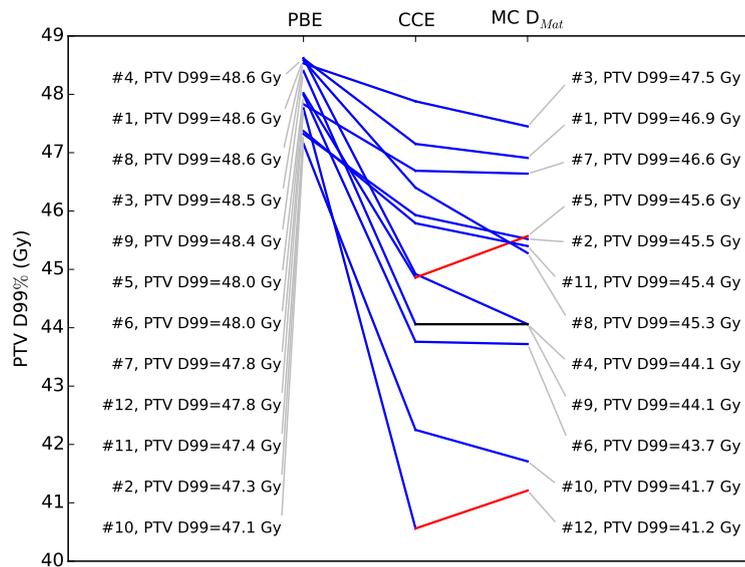
(b) PBE versus CCE versus MC D$_{Mat}$ for PTV D99%

Figure 3.31: Table-graphics comparing PTV D99% values for PBE, CCE, MC D$_{Mat}$ and MC D$_{Wtr}$.
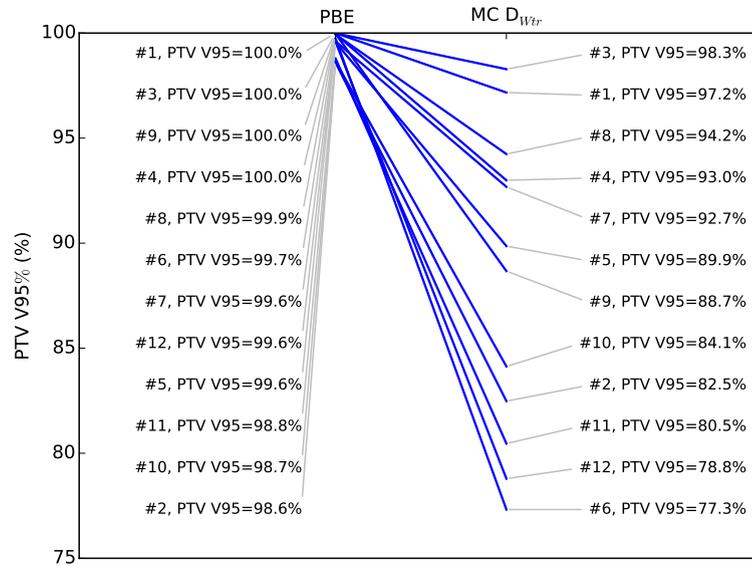
(a) PBE versus MC D_{Wtr} for PTV V95%
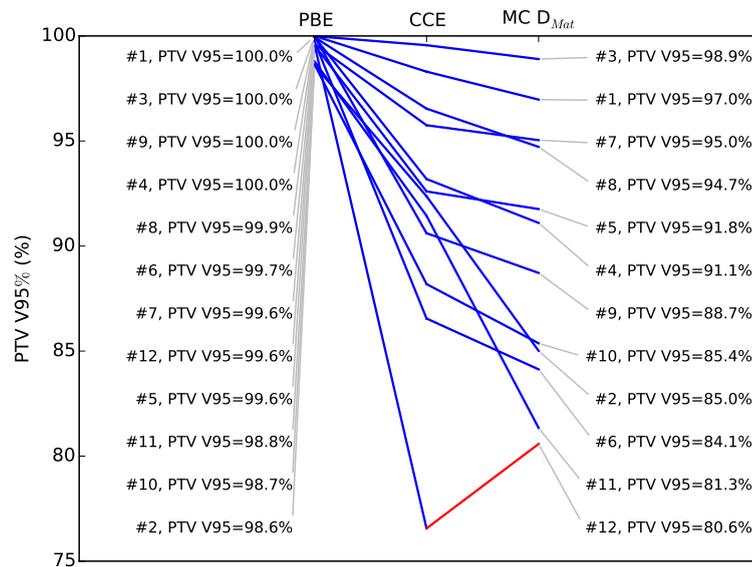
(b) PBE versus CCE versus MC D_{Mat} for PTV V95%

Figure 3.32: Table Graphics comparing PTV V95% values for PBE, CCE, MC D_{Mat} and MC D_{Wtr}.

comparing the CCE values with those from MC D_{Mat}. One case, case #12, shows an increase in PTV V95% from D_{Mat} compared to CCE (note the red line), but the overall trend is evident.

Figure 3.33(a) does not show a clear overall trend, almost as many values of GTV V100% decrease when re-calculated with MC D_{Wtr} as increase. Statistical calculations can be used (section 3.4) to reveal if there is a real difference between the two sets of data, or whether the changes are random. As there is no clear pattern emerging from these 12 cases, it may be expected that a very large number of cases would be required to demonstrate a difference. Likewise, Figure 3.33(b) does not show an obvious trend when comparing results from MC D_{Mat} with those from CCE. There does seem to be a small reduction in most values of GTV D100% when comparing results from PBE with those from CCE.

Figure 3.34(a) shows that there is very little change in heart V80% when re-calculating PBE plans versus MC D_{Wtr} plans. The order of the cases, when ranked by heart V80% for PBE is very similar to the order when ranked by heart V80% when re-calculating MC D_{Wtr}. Similarly, Figure 3.34(b) shows very little variation in the three dose calculation methods, PBE, CCE and MC D_{Mat}. This is a useful result from a clinical point of view, regardless of dose calculation method, the heart V80% is consistent (accurate). Therefore, clinical experience based on doses calculated with the PBE algorithms is still valid when using the CCE algorithm or MC methods. However, there is likely to be no interesting behaviour to investigate for heart V80% in this work.

Figure 3.35(a) shows that most values of Cord D_{1cc} increase when re-calculated with MC D_{Wtr} compared with the original PBE dose. For Figure 3.35(b), comparing PBE and CCE, there is a small reduction in Cord PRV D_{1cc} in each case. When comparing CCE and MC D_{Mat} there is not such a clear trend, a small majority of cases show a decrease in Cord PRV D_{1cc}, but some show no change and some show a small increase. Based on the variation in these results, it is
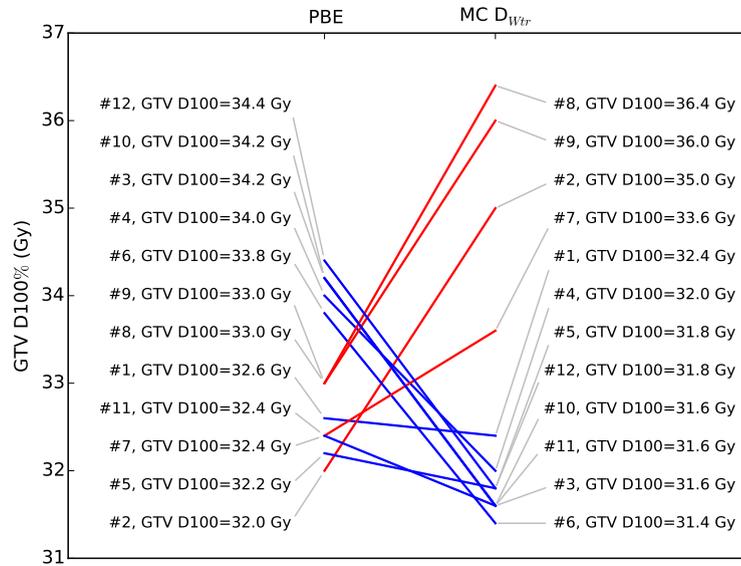
(a) PBE versus MC D_{Wtr} for GTV V100%
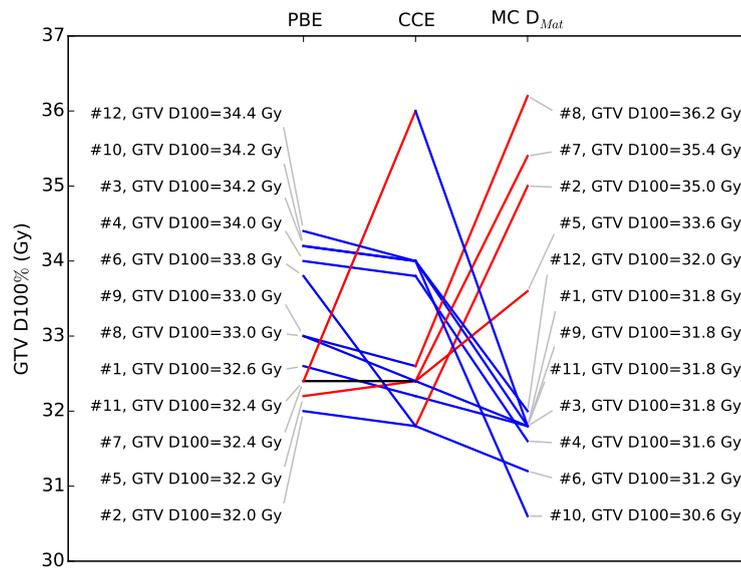
(b) PBE versus CCE versus MC D_{Mat} for GTV V100%

Figure 3.33: Table Graphics comparing GTV V100% values for PBE, CCE, MC D_{Mat} and MC D_{Wtr}.
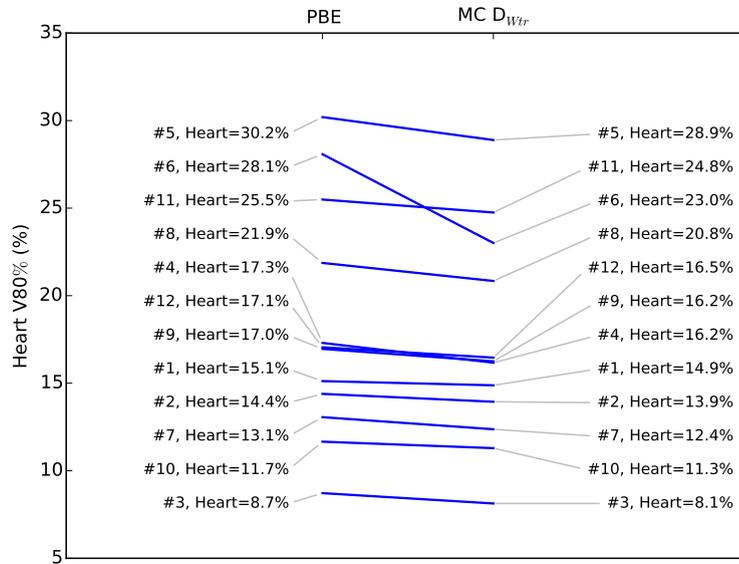
(a) PBE versus MC D$_{Wtr}$ for heart V80%
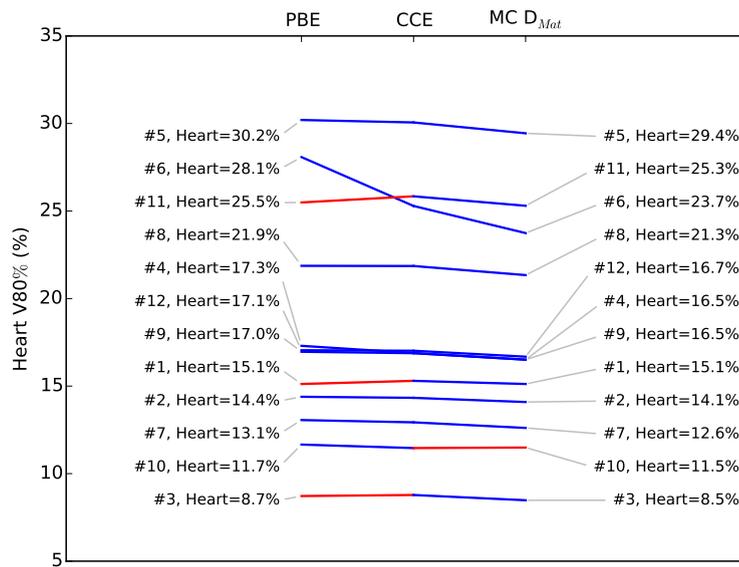
(b) PBE versus CCE versus MC D$_{Mat}$ for heart V80%

Figure 3.34: Table Graphics comparing heart V80% values for PBE, CCE, MC D$_{Mat}$ and MC D$_{Wtr}$.
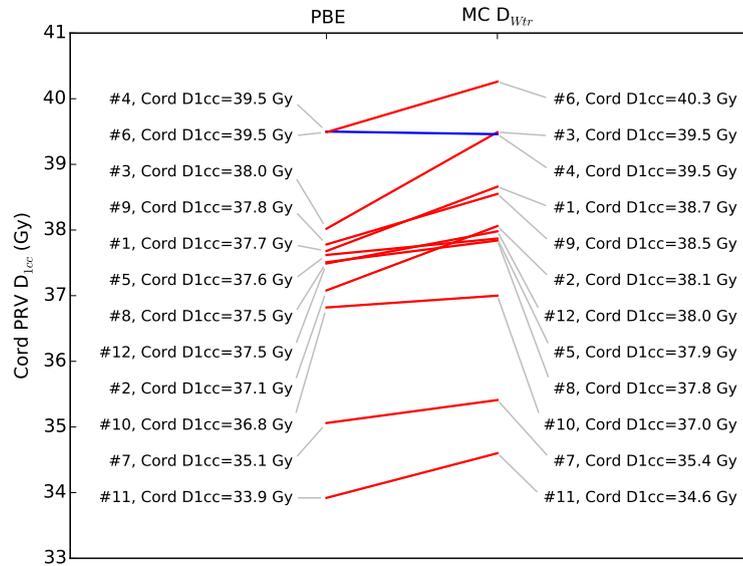
(a) PBE versus MC D_{Wtr} for Cord D_{1cc}
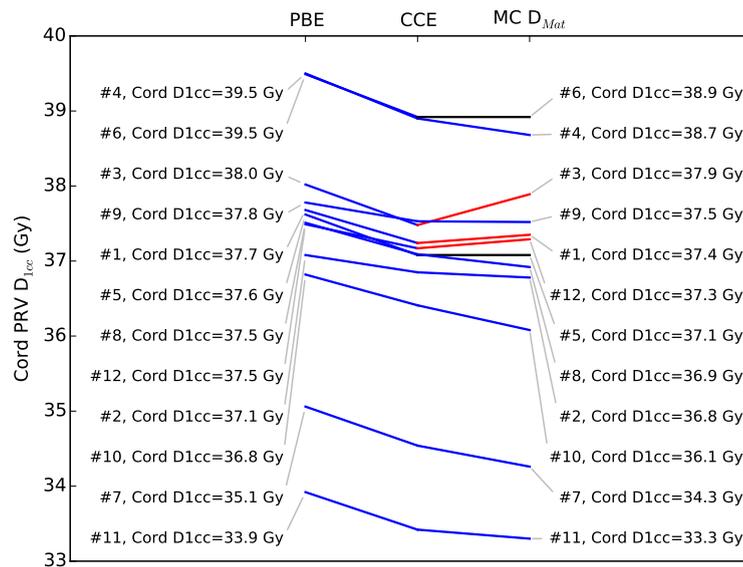
(b) PBE versus CCE versus MC D_{Mat} Cord D_{1cc}

Figure 3.35: Table Graphics comparing Cord D_{1cc} values for PBE, CCE, MC D_{Mat} and MC D_{Wtr}.

not likely that a comparison can be performed that will have a high statistical power (i.e. greater than 90%) to detect a true difference in the dose calculation methods.

Figure 3.36(a) shows that most values of lung V20 Gy reduce when produced with MC D_{Wtr} compared to the original PBE calculation. Figure 3.36(b) shows most values of lung V20 Gy reduce when recalculated with CCE compared to the original PBE, but there are a number of cases showing the opposite behaviour. The values of lung V20 Gy from the MC D_{Mat} data show small increases for most cases, but there are some with no real change, and one with a slight reduction.

3.4 Discussion

3.4.1 Discussion of selected cases

Figure 3.37 shows a comparison of the DVHs for CCE and MC D_{Mat} dose calculations, for the PTV, heart, combined lungs and the spinal cord PRV, for case #10. This case had one of the largest decreases in PTV V95% and PTV D99%. The most striking aspect of the graph is how similar the DVH plots are for the different dose calculation methods. The largest deviations between the DVH plots occur for the Cord PRV, which demonstrates differences in the region of the graph depicting doses of 5-30 Gy. For the PTV, the MC D_{Mat} plotted line is slightly to the left of the plotted line for the CCE algorithm, showing a reduction in coverage for the MC D_{Mat} case. This does not appear to be a systematic shift in the plotted line, as there is not a constant offset between the two curves. The DVH lines are very similar for the heart, which agrees with the results in Figure 3.34, which demonstrates consistent heart V80% values for the different dose calculation methods. For the combined lungs, the DVH shows that the MC D_{Mat} calculation results in a higher dose to the lungs in the 15-20 Gy regions of the
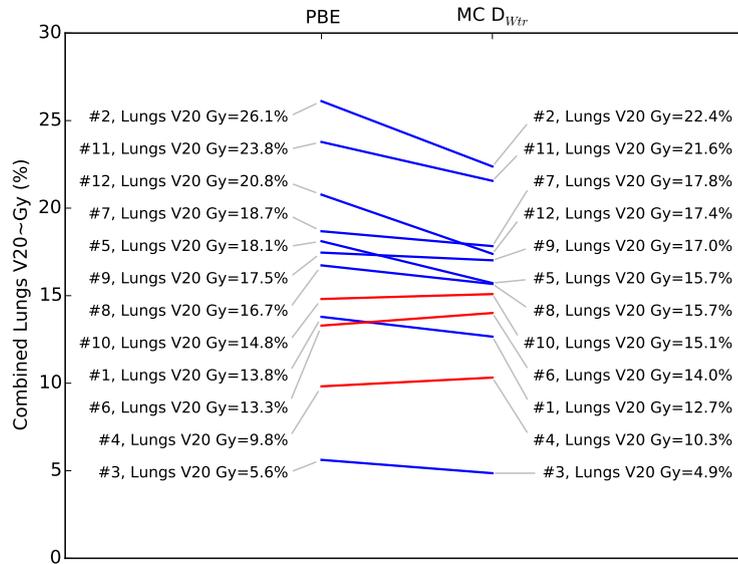
(a) PBE versus MC D$_{Wtr}$ for Combined Lung V20 Gy
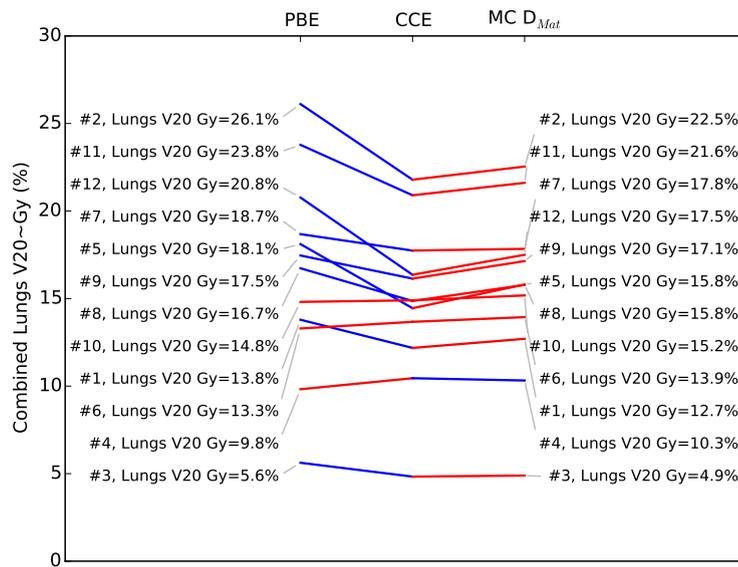
(b) PBE versus CCE versus MC D$_{Mat}$ for Combined Lung V20 Gy

Figure 3.36: Table Graphics comparing V20 Gy values for Combined Lung for PBE, CCE, MC D$_{Mat}$ and MC D$_{Wtr}$.

graph.

Figure 3.38 shows a comparison of the DVHs for CCE and MC D_{Mat}, for case #12. This was the only case to show an increase in PTV V95% for the MC D_{Mat} calculation compared to the CCE algorithm. Once more, the DVH plots for the different dose calculation methods are very similar to one another. The PTV DVH plot for the MC D_{Mat} calculation exhibits a slightly steeper curve than that for the CCE algorithm. The plotted line for the MC D_{Mat} calculation indicates a higher dose in the 60-100% volume region, but a lower dose in the region below 40% of the volume, compared to the CCE algorithm. The DVH plots for the heart are very similar. The DVH plots for the combined lungs shows that the MC D_{Mat} method is calculating slightly higher doses in the 20-40% volume region, and below 5% volume, compared to the CCE algorithm. The DVH plots for the cord PRV again show differences in the region of the plot with pronounced steps. These steps are due to the small volume of the cord PRV and the finite size of the voxels used in the DVH calculation. As the dose reduces to the right of the DVH graph, fewer voxels are treated to given dose level; if the change in the volume treated to given dose level is large compared to the total OAR volume, steps will be produced in the DVH. Similarly, the differences in the plotted lines for the two dose calculation methods in this region relate to changes in the dose to a comparatively small number of voxels.

By definition, DVH plots hold no information about where in an ROI the doses are, so comparisons of DVH plots cannot determine where the changes in dose are occurring. For this, comparisons of isodose plots must be used.

Figure 3.39 shows isodose plots for case #10, the same case for which the DVH is shown in Figure 3.37. Comparing the isodoses for CCE in Figure 3.39(a) with those for MC D_{Mat} in Figure 3.39(b), it can be seen that both demonstrate a 'hole' in the 95% isodose (47.5 Gy isodose in the legends of Figures 3.39(a) and (b)) in the bronchus, the small low density region on CT, anterior of the cross-hairs. For

the MC D$_{Mat}$ method, the smaller red contour (i.e. the 95% isodose) almost fills the area of the bronchus on this slice; as the larger red contour wholly encloses this contour, this indicates that the dose within the smaller red contour is below 95% of the prescribed dose. For the CCE algorithm, the area of the bronchus not covered by the 95% isodose is much smaller. This leads to the lower PTV V95% observed for the MC D$_{Mat}$ in this case, compared with the CCE algorithm. The dose to a small, low density regions such as this will be dominated by scattered radiation from the surrounding soft tissue regions; the difference in dose calculated by the two methods in this region is due to the improved modelling of lateral scatter of the MC method, compared with the CCE algorithm.

Figure 3.40 shows isodose plots for case #12, which is the same case for which the DVH is shown in Figure 3.38. In Figure 3.40(a), the 95% isodose resembles a T-shape. Where the PTV overlaps the heart, the lateral extent of the isodose covers the PTV. Where the PTV overlaps the lungs, the lateral extent of the 95% isodose is narrower than the PTV. In Figure 3.40(b) it can be seen that the coverage of the PTV by the 95% isodose where the PTV overlaps the lungs is improved (although the PTV is not completely covered). It is hypothesised that differences in the modelling of the forward scattered radiation between the CCE and MC D$_{Mat}$ calculations for beams entering through the heart and spinal cord cause the differences seen in this case.

3.4.2 Calculation of sample size

Whitley and Ball (2002) discuss a method for calculating sample sizes, so that the study may have the desired statistical power, following the method of Altman (1999). The three factors that determine the sample size are the *p*-value, the power i.e. the probability of correctly identifying a difference where one genuinely exists, and the size of the effect, e.g. the difference in the means of two populations. The smaller the desired *p*-value, the larger the required sample size; similarly the
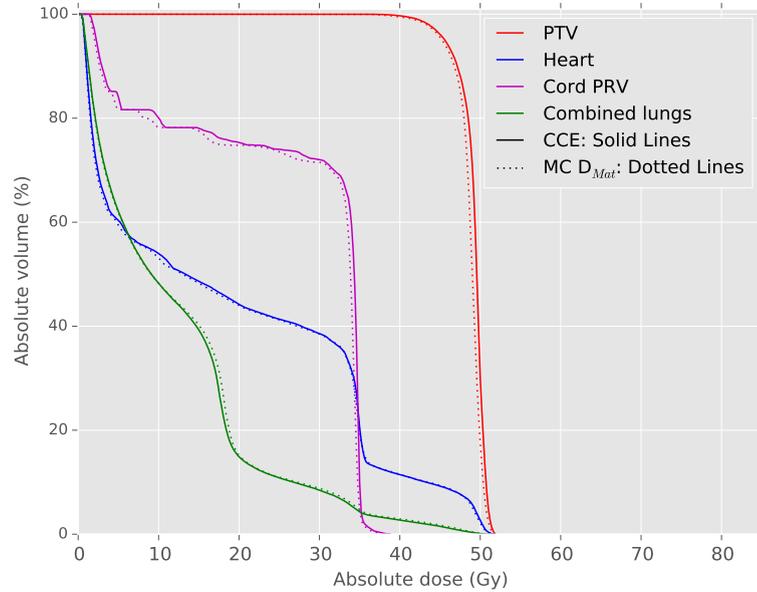
Figure 3.37: Comparison of DVH parameters for CCE and MC D$_{Mat}$, for case #10 of the 12 cases used in the pilot study.

smaller the effect that is being investigated, the larger the required sample size. Conversely, the larger the desired power, the larger the required sample size. The power of the test is the probability that the test will produce a significant difference, at a given level of significance.

Whitley and Ball (2002) relate these three parameters to sample size using Equations 3.10a and 3.10b, where d is called the Standardised Difference (the factor of 2 applicable when undertaking a paired study comparing different techniques on the same sample group). A key parameter in Equation 3.10b is the C$_{p,power}$, the critical value, defined by Equation 3.11, where PDF is the Gaussian Probability Density Function.

$$d = \frac{\text{Target Difference}}{\text{Standard Deviation}} \tag{3.10a}$$

$$N = \frac{2}{d} \times C_{p,power} \tag{3.10b}$$
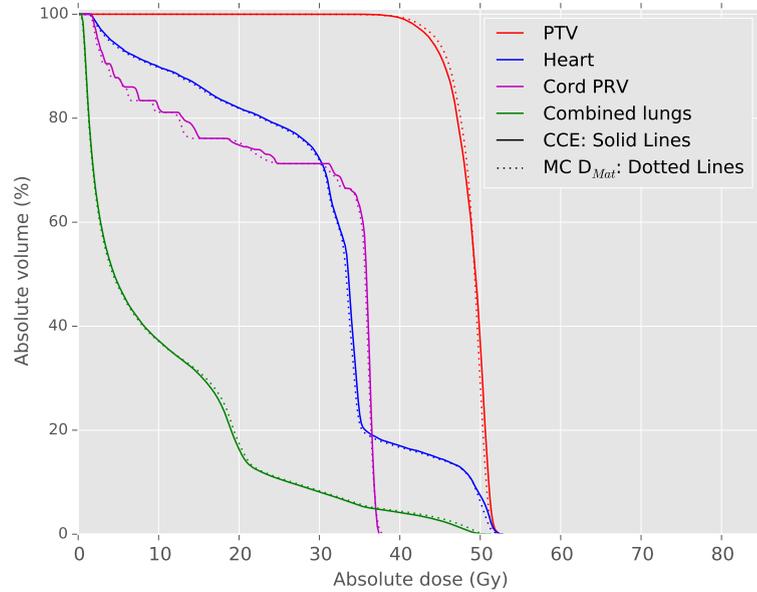
Figure 3.38: Comparison of DVH parameters for CCE and MC D$_{Mat}$, for case #12 of the 12 cases used in the pilot study.

$$C_{p,power} = \left(\int_0^{p/2} PDF + \int_0^{2(1-power)} PDF \right)^2 \tag{3.11}$$

The DVH parameters for the CCE and MC D$_{Mat}$ distributions are shown in Table 3.6. For each DVH parameter, the differences between the values from the CCE and MC D$_{Mat}$ dose distributions are calculated, then the mean and standard deviations of the 12 values found for each DVH parameter. Table 3.7 shows the sample size calculation for each DVH parameter, using the mean differences as an estimate of the target difference, using a *p*-value of 0.05 or 0.01, and powers of 0.8 and 0.9 as appropriate. The values of C$_{p,power}$ are calculated from Equation 3.11, cross-checked against Whitley and Ball (2002), where they have been tabulated. The PTV D99% has a relatively small mean differences between the values for CCE and MC D$_{Mat}$, compared with the standard deviation. Therefore, if a *p*-value of 0.05 is desired, then even the relatively modest power of 0.8 would require N=98.8, which should be rounded up to 100. This would be a very large number

PTV D99%			PTV V95%			GTV D100%			Heart V80%			Cord PRV D_{1cc}			Combined Lungs V20 Gy		
CCE	MC D_{Mat}	Diff	CCE	MC D_{Mat}	Diff	CCE	MC D_{Mat}	Diff	CCE	MC D_{Mat}	Diff	CCE	MC D_{Mat}	Diff	CCE	MC D_{Mat}	Diff
47.2	46.9	0.2	98.3	97.0	1.3	32.2	31.8	0.4	15.3	15.1	0.2	37.2	37.4	-0.1	12.2	12.7	-0.5
45.9	45.5	0.4	92.4	85.0	7.4	31.8	35.0	-3.2	14.3	14.1	0.2	36.9	36.8	0.1	21.8	22.5	-0.8
47.9	47.5	0.4	99.6	98.9	0.7	34.0	31.8	2.2	8.8	8.5	0.3	37.5	37.9	-0.4	4.8	4.9	-0.1
44.9	44.1	0.9	93.2	91.1	2.1	33.8	31.6	2.2	16.9	16.5	0.4	38.9	38.7	0.2	10.5	10.3	0.1
44.9	45.6	-0.7	92.6	91.8	0.8	32.4	33.6	-1.2	30.1	29.4	0.6	37.1	37.1	0.0	14.5	15.8	-1.4
43.8	43.7	0.0	86.6	84.1	2.4	31.8	31.2	0.6	25.3	23.7	1.6	38.9	38.9	0.0	13.7	13.9	-0.3
46.7	46.6	0.0	95.8	95.0	0.7	32.4	35.4	-3.0	12.9	12.6	0.3	34.5	34.3	0.3	17.7	17.8	-0.1
46.4	45.3	1.1	96.5	94.7	1.8	32.6	36.2	-3.6	21.9	21.3	0.5	37.1	36.9	0.2	14.9	15.8	-0.9
44.1	44.1	0.0	90.6	88.7	1.9	32.4	31.8	0.6	16.9	16.5	0.4	37.5	37.5	0.0	16.1	17.1	-1.0
42.3	41.7	0.5	88.2	85.4	2.8	34.0	30.6	3.4	11.5	11.5	0.0	36.4	36.1	0.3	14.9	15.2	-0.3
45.8	45.4	0.4	91.4	81.3	10.1	36.0	31.8	4.2	25.8	25.3	0.5	33.4	33.3	0.1	20.9	21.6	-0.7
40.6	41.2	-0.6	76.6	80.6	-4.0	34.0	32.0	2.0	17.0	16.7	0.3	37.2	37.3	-0.1	16.4	17.5	-1.1
		Mean = 0.2			Mean = 2.3			Mean = 0.4			Mean = 0.4			Mean = 0.05			Mean = -0.6
		Std Dev = 0.5			Std Dev = 3.5			Std Dev = 2.6			Std Dev = 0.4			Std Dev = 0.2			Std Dev = 0.5

Table 3.6: Comparison of DVH parameters derived from CCE and MC D_{Mat} calculated doses.

of oesophageal plans to find and simulate. Repeating the analysis for PTV V95% leads to a recommended sample size of 40 (36.1) for a power of 0.8 and 50 (48.6) for a power of 0.9, (this calculation is also depicted in the Nomogram in Figure 3.41). 40 cases on top of the initial 12 cases used for the pilot study is feasible, additional cases may not be possible due to the number of existing clinical cases available.

The sample size calculated for the GTV V100% is huge, approaching 700 (659.1). This is not unexpected, as Figure 3.33 showed no clear trend, so a very large number of comparisons would be needed to detect a signal amongst the noise, and it is unlikely that such a small difference would be clinically significant. It is not feasible to find 700 oesophageal cases using a consistent approach to simulate, as Velindre plans approx 50 oesophageal cases per year, and techniques change over time.

Performing the calculation for the heart V80% using various values of *p*-value and power shows that relatively few cases are needed for even high power and low *p*-value. Referring back to Figure 3.34, this is probably due to a consistent, if small, reduction in heart V80% when changing from the CCE to the MC D_{Mat}.

The sample size calculated for the Cord PRV D_{1cc} is also very large, around 250 (249.6). This is to be expected, as Figure 3.35 shows little variation in Cord PRV D_{1cc} between the CCE to the MC D_{Mat} dose distributions.

Due to the mean change in Lung V20 Gy being larger than the standard deviation in the mean, only 25 cases would be required for a comparison with 95% power and a *p*-value of 0.01.

3.5   Conclusions

The RTGrid has been used to simulate RT treatment plans for oesophageal cancers. To allow this, a CTRamp, OesRAMP, was produced from scans of a CIRS-

Comparison	Target Difference	Standard Deviation	Standardised Difference	P-value	Power	$c_{p,power}$	N
PTV D99% CC —MC D_{Mat}	0.2	0.5	0.40	0.05	0.8	7.8	97.5
PTV V95% CC —MC D_{Mat}	2.3	3.5	0.66	0.05	0.8	7.8	36.1
PTV V95% CC —MC D_{Mat}	2.3	3.5	0.66	0.05	0.9	10.5	48.6
GTV 100% CC —MC D_{Mat}	0.4	2.6	0.15	0.05	0.8	7.8	659.1
Heart 80% CC —MC D_{Mat}	0.4	0.4	1.00	0.05	0.8	7.8	15.6
Heart 80% CC —MC D_{Mat}	0.4	0.4	1.00	0.05	0.9	10.5	21.0
Heart 80% CC —MC D_{Mat}	0.4	0.4	1.00	0.01	0.95	17.8	35.6
Cord PRV D_{1cc} CC —MC D_{Mat}	0.05	0.2	0.25	0.05	0.8	7.8	249.6
Combined Lungs V20 Gy CC —MC D_{Mat}	-0.6	0.5	-1.2	0.01	0.95	17.8	24.7

Table 3.7: Sample sizes calculated from pilot study results, based on different *p*-values and statistical powers.

062M phantom. The EDW Component Module was incorporated into the RTGrid Profile, to allow the simulation of plans including Varian EDW fields. Code was developed to remove the high density intra-venous contrast from CT scans, ignore data on the CT scan outside the external contour. Separate code was developed to create artificial volumes at the superior and inferior extent of the CT scan, so that the MC simulation followed the approach of the TPS calculations.

A group of 12 oesophageal plans were simulated, and DVH parameters calculated from the MC dose distributions. Parameters commonly reported from the DVHs, such as PTV V95%, were compared for the CCE and MC, and sample size calculations were performed. From this analysis, 40 clinical cases will be used for the main body of the work, to investigate the coverage of the PTV with the 95% isodose. Other DVH parameters for oesophageal plans will be analysed, but may not be statistically significant at the same level.
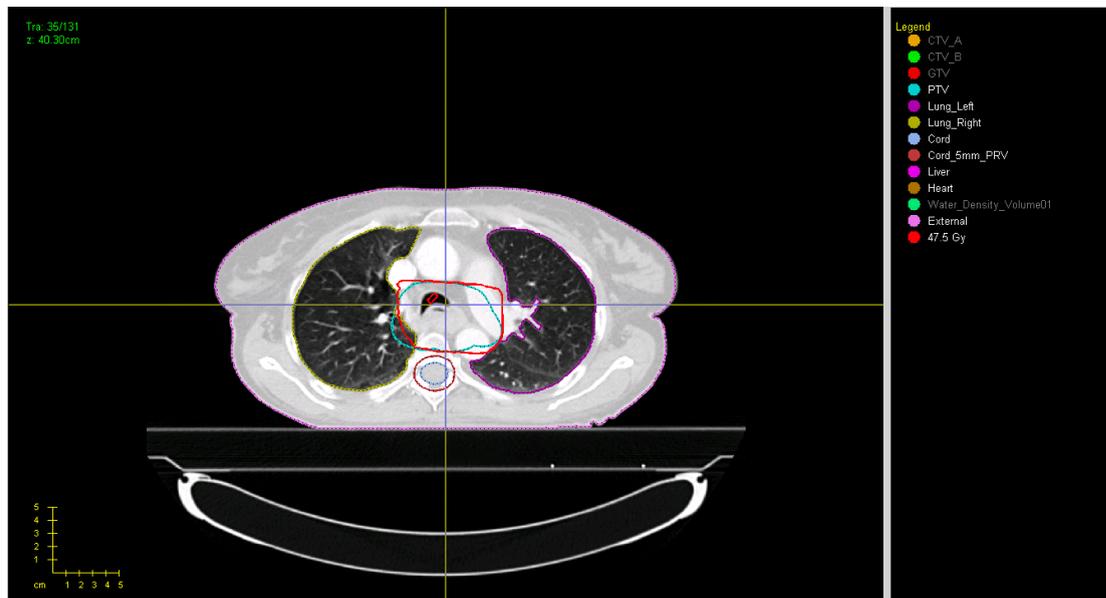
(a) 95% isodose calculated from CCE dose distribution.
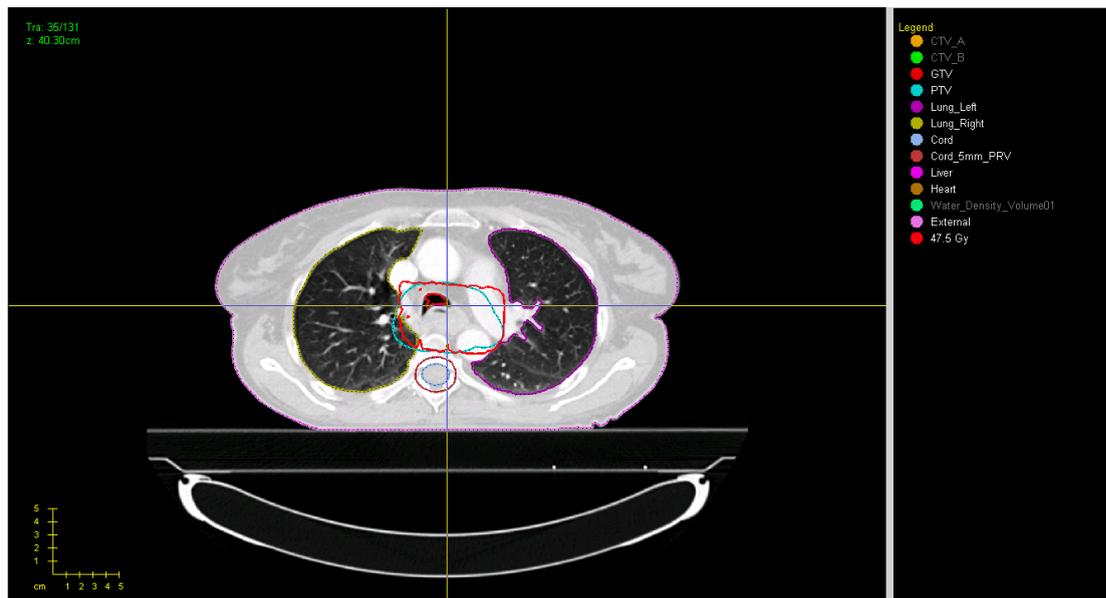
(b) 95% isodose calculated from MC D_{Mat} dose distribution.

Figure 3.39: Screen-shots comparing 95% isodose coverage of PTV in CT slice.
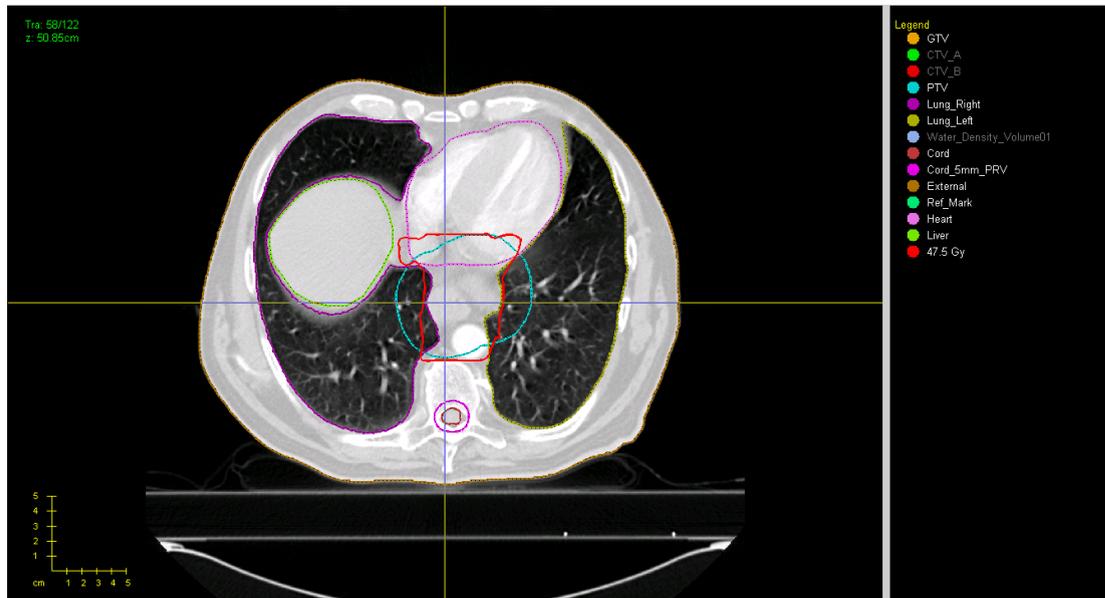
(a) 95% isodose calculated from CCE dose distribution.
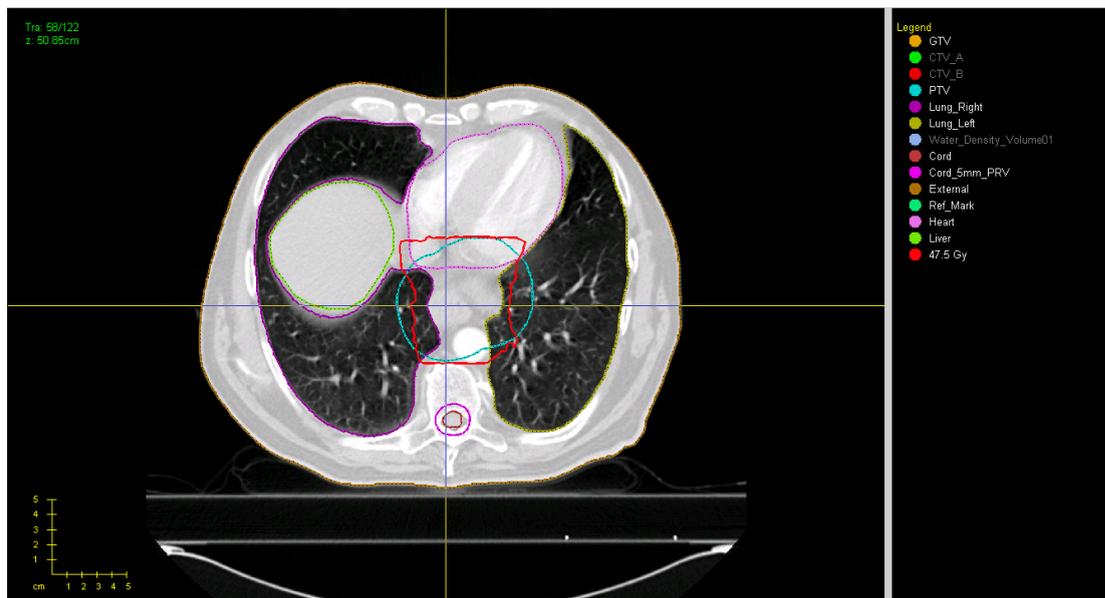
(b) 95% isodose calculated from MC D_{Mat} dose distribution.

Figure 3.40: Screen-shots comparing 95% isodose coverage of PTV in CT slice.
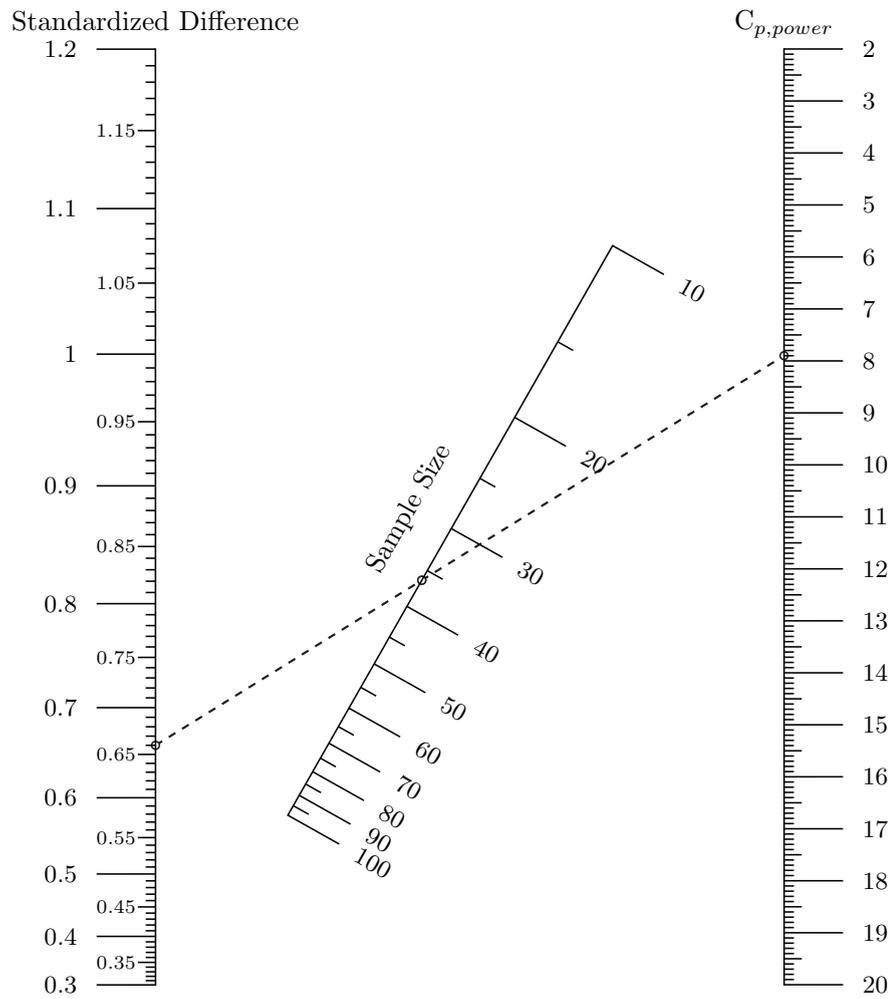
Figure 3.41: Nomogram showing selection of sample size from standardized difference and critical value dependent on p-value and power.

Chapter 4

Main study

4.1   Introduction

Following the sample size calculation in the last chapter (section 3.4), 40 oesopha-
geal case were simulated with the RTGrid. These were the next 40 cases identified
following the approach describe in section 3.2.6. This chapter will describe the
measures calculated for these 40 cases, and the analysis of these measures, to
determine if differences between doses calculated with the CCE algorithm and
MC simulations are statistically significant. Finally, situations where MC is of
most of benefit will be identified.

4.2   Materials and methods

As for the cases in the Pilot Study, after simulation and calibration, the doses
calculated by MC were held within CERR format files for analysis. The methods
of analysis are described below.

4.2.1   Dose Volume Histogram parameters

The DVH parameters described in section 3.2.1 were calculated and analysed for
the four dose calculation methods for all cases.

4.2.2 Conformity indices

DVH parameters such as PTV V95% are commonly used metrics to describe either how the dose distribution covers a target volume, or how the dose distribution avoids organs at risk. In the production of the DVH, the 3D dose distribution is reduced to a single line. A dose volume parameter is a single point along this line. Compressing the information from a 3D distribution to a single value allows different plans to be compared easily, but a lot of information is lost. Two plans may have the same PTV V95%, but the shapes of the V95% may be quite different. One plan may have the 95% isodose tightly conforming to the PTV, in the the other the 95% isodose may include a lot of tissue in addition to the PTV.

To address the shortcomings of DVH parameters, while preserving the convenience of summarising a 3D distribution as a single value, a conformity index may be used. A conformity index is a value that describes how well two volumes agree. Hanna et al. (2010) describes the use of conformity indices for comparing target volumes outlined on the same set of images by multiple observers, as well as comparing the volume defined by a given isodose with that of a target volume. From the pilot study, it is known that the PTV V95% varies enough between different calculation methods in oesophageal radiotherapy to be worth investigating. It follows that conformity indices describing the coverage of the PTV by the 95% isodose may also produce interesting results.

4.2.2.1 RTOG Index

The RTOG Conformance Index (RTOG-CI) was introduced by the RTOG in Shaw et al. (1993), and is given by Equation 4.1, where PIV is the Prescription Isodose Volume and TV is the Target Volume. This was used in Shaw et al. (2000) for the assessment of plan quality for radiosurgery. The advantage of the RTOG-CI is its simplicity; there is little ambiguity in the ratio of two volumes. This is

particularly important when trying to implement the calculation in a computer program, to avoid details of the implementation affecting the calculated value. The main disadvantage is that no account is taken of the overlap of the two volumes being assessed. In principle, the isodose volume of interest could wholly miss the target volume, but the RTOG Index could still exceed unity if the isodose volume is larger than the target volume.

$$\text{RTOG-CI} = \frac{PIV}{TV} \tag{4.1}$$

4.2.2.2 Jaccard Index

An example of an index that does account for the overlap between two volumes is the Jaccard Concordance Index (Jaccard-CI), defined in Hanna et al. (2010) as Equation 4.2, where A and B are the volumes of two structures being compared. According to Gwynne et al. (2012), the Jaccard-CI was originally defined in 1901, and has been known by various terms including the conformity index and the conformity index. In this work, it will be referred to as the Jaccard-CI.

$$\text{Jaccard-CI} = \frac{A \cap B}{A \cup B} \tag{4.2}$$

4.2.2.3 Dice Similarity Coefficient

A similar coefficient to the Jaccard-CI is Dice Similarity Coefficient (Dice-CI), defined in Hanna et al. (2010) by Equation 4.3.

$$\text{Dice-CI} = \frac{2(A \cap B)}{A + B} \tag{4.3}$$

4.2.2.4 van't Reit Conformation Number

Another measure is the van't Reit Conformation Number (Reit-CI), defined in van't Riet et al. (1997) as Equation 4.4, where $V_{T,ref}$ is the volume of the target receiving a dose greater than or equal to a reference dose, V_T is the volume of the target and V_{ref} is the volume of tissue receiving a dose greater than or equal to a reference dose. This may be re-written as Equation 4.5, where the terms A and B have the same meanings as in Equation 4.2 and Equation 4.3. The first term in the Reit-CI describes how much of the target volume is covered by the reference isodose. The second term describes how much tissue (target or otherwise) is covered by the target isodose. It can be seen that a low value may be due to the reference isodose covering little of the target, or the reference isodose covering a large volume of tissue outside the target. The same is true of the Jaccard-CI or Dice-CI, although this is less apparent from the equations.

$$\text{Reit-CI} \ = \frac{V_{T,ref}}{V_T} * \frac{V_{T,ref}}{V_{ref}} \tag{4.4}$$

$$\text{Reit-CI} \ = \frac{A \cap B}{A} * \frac{A \cap B}{B} \tag{4.5}$$

4.2.2.5 Discordance Index

Kepka et al. (2007) described a Discordance Index (DI-CI), which they applied to analysing the variation in delineated volumes for nodal volumes in lung cancer radiotherapy. DI-CI was defined according to Equation 4.6, where VI is the Volume of Intersection of two volumes, and DV is the Delineated Volume, i.e. the volume of the reference outline. DI-CI was intended to supplement a concordance index, (defined differently to the Jaccard Concordance Index), but Kepka et al. (2007)

found that there was no statistical difference in the two endpoints. Equation 4.7 shows the equation expressed in terms of the intersection between volumes.

$$\text{DI-CI } = (1 - \frac{VI}{DV}) * 100\% \tag{4.6}$$

$$\text{DI-CI } = (1 - \frac{A \cap B}{DV}) * 100\% \tag{4.7}$$

4.2.2.6 Geometric Miss Index

Another measure, which unlike most other indices increases as the agreement between two volumes decreases, is the Geometric Miss Index (GMI). Muijs et al. (2009) defined the GMI according to Equation 4.8. A GMI of 0 indicates that the one volume is completely enclosed by the other; a GMI of 1 indicates that the volumes have no overlap.

$$\text{GMI } = \frac{A - A \cap B}{A} \tag{4.8}$$

4.2.3 Mean Distance to Conformity

Jena et al. (2010) introduced the concept of a Mean Distance to Conformity (MDC), defined as the average distance that points on one surface must be moved to achieve perfect conformity with another surface. It can be computed as a single statistic, describing the average magnitude of the movement a point to achieve conformity. Alternatively, two statistics may be calculated, the first for points that must be moved outwards to the reference surface, and a second statistic for points that must be moved inwards towards the reference surface. As Jena et al. (2010) discussed the MDC for comparing two outlined volumes, they dubbed

this over or under-contouring. Carrington et al. (2016) have used the MDC for comparison of target volumes and isodose surfaces.

Jena et al. (2010) compared the MDC values with a conformity index (defined in the same way as the Jaccard Conformity Index, Equation 4.2), and found the MDC to be more sensitive to errors in conformity that was the conformity index.

4.2.4 Tumour Control Probability

The Tumour Control Probability (TCP) is the probability that the dose delivered to a tumour will either eradicate the tumour cells, or kill enough of the cells that the tumour will no longer threaten the life of the patient. It is not generally possible to calculate the TCP for a given patient, but models can be created from studies of groups of patients, and these models can be used to compare different treatment approaches. It this work, a published model for calculation of TCP for oesophageal cancer will be used to compare the effect of the different dose distributions from the different calculation methods.

4.2.5 Generation of measures

Having outlined the measures that will be compared for the 40 cases in the study, the methods used to generate these measures will be discussed.

4.2.5.1 Generation of Dose Volume Histogram parameters

The DVH parameters reported for oesophageal plans in Velindre Cancer Centre were calculated for all 40 cases, following the method described in section 3.2.4.11.

4.2.5.2   Generation of conformity indices

Calculation of the various conformity indices involved a two step process. Firstly, Matlab scripts were used to open the CERR (Deasy et al., 2003) files created previously (section 3.2.2), call existing functions from CERR to create a volume from the 95% isodose, and save this to the CERR file. Secondly, functions within CERR Trial Addition (Gwynne et al., 2012) were used to calculate RTOG-CI, Jaccard-CI, Dice-CI, Reit-CI, GMI and DI-CI conformity indices.

4.2.5.3   Generation of Mean Distance to Conformity

The method used to calculate the MDC values follows that for the Conformity Indices, again using functions within CERR Trial Addition.

4.2.5.4   Generation of Tumour Control Probability

Geh et al. (2006) developed an equation to calculate TCP for oesophagus cancer, based on outcome data, demographic data (age, gender) and treatment data, including radiotherapy dose. In their work, the dose was taken to be the prescription dose, so no account was taken of how the prescription isodose conformed to the target. Carrington et al. (2015) used the model from Geh et al. (2006), but used the DVH data for different treatment techniques and prescriptions to predict TCP and gastric toxicity. In this work, DVH data from CERR were used to allow calculation of the TCP value for each volume element. In order to assess the effect on the final analysis, in addition to calculating the TCP with the median values of the parameters and the recommended values of the coefficients from Geh et al. (2006), each of the parameters were replaced in turn with the minimum or maximum value, and the coefficients replaced in turn by either 90% or 110% of the

published values, and the highest and lowest TCP values produced were recorded. This produced three measures, TCP-Median, TCP-Min and TCP-Max.

4.3 Results

4.3.1 Correlations and multiple testing

In section 4.2, several measures to compare the different dose calculation methods have been proposed. When several tests are performed on the same data, the statistical results, notably the *p*-values, cannot be presented in isolation from each other. Put simply, if a sufficiently large number of tests are performed on a dataset, one of the tests will appear to be significant purely by chance. This is a common issue in fields such as radiomics (Yip and Aerts, 2016) and genomics, which frequently perform many tests on the same dataset. A number of authors have described approaches to correct *p*-values in multiple testing scenarios in these fields. Authors including Mu et al. (2015) and Orlhac et al. (2014) recommend grouping measures that are strongly correlated with one another, (i.e. Pearson correlation co-efficient greater that 0.8), and only including one measure per group in the analysis. Once the correlated measures have been identified, and all but one measure per 'group' have been excluded, Yip and Aerts (2016) describe two approaches to handling multiple testing. These approaches are a family-wise error rate controlling method such as the Holm-Bonferroni approach, and a false discovery rate controlling method such as the Benjamini-Hochberg approach. Both of these methods to correct the *p*-values require that the measures involved are independent.

4.3.1.1 Correlations between measures

For each of the 17 measures (six CI, six DVH parameters, two MDC values and three TCP values), the correlation co-efficient with every other measure was calculated (Table 4.1). This shows that the Dice-CI, GMI, Jaccard-CI, Reit-CI and RTOG conformity indices are strongly correlated (0.8 was used as the cut-off, but the results are relatively insensitive to the cut-off value; using 0.75 or 0.85 would have lead to the same conclusions). This correlation implies that the measures may provide redundant information, so only one should be used in the final analysis.

It is interesting that the RTOG is strongly correlated with other CIs (except DI-CI), given that RTOG is a relatively crude measure; it does not account for the overlap of the volumes compared, only their size. This may be because the PTV and V95% have a high degree of overlap, so accounting for this does not truly provide extra information. It is also interesting that PTV V95% is not strongly correlated with any of the CIs, or the MDC values, as all describe coverage of PTV by 95% isodose. This shows that the CIs and MDC values do contain different information to the PTV V95% values. The DI-CI is weakly correlated with the other conformity indices, probably because, as stated by Muijs et al. (2009), it is designed to be a complementary measure to a Jaccard-CI like conformity index.

It is not surprising that the three related approaches to calculating TCP are correlated with one another.

From Table 4.1 , the correlated groups are: -

- Dice-CI, GMI, Jaccard-CI, Reit-CI and RTOG. All except **Jaccard-CI** will be excluded from further analysis;

- TCP_{MEDIAN}, TCP_{MIN}, TCP_{MAX}. All except **TCP_{MEDIAN}** will be excluded from further analysis.

	PTV D99%	PTV V95%	GTV D100%	Heart V80%	Cord PRV D1cc	Lungs V20 Gy	DICE	DI	GMI	JACCARD	REIT	RTOG	MDC Over	MDC Under	TCP$_{median}$	TCP$_{min}$	TCP$_{max}$
PTV D99%																	
PTV V95%	0.4																
GTV D100%	0.0	0.0															
Heart V80%	0.2	0.4	0.0														
Cord PRV D1cc	0.1	0.3	0.1	0.2													
Lungs V20 Gy	0.1	0.1	0.1	0.4	0.0												
DICE	0.2	0.6	0.1	0.1	0.1	0.0											
DI	0.3	0.3	-0.1	0.0	0.1	0.2	0.2										
GMI	-0.3	-0.6	-0.1	-0.1	-0.1	-0.1	-1.0	-0.5									
JACCARD	0.2	0.6	0.1	0.1	0.1	0.0	1.0	0.2	-1.0								
REIT	0.2	0.6	0.1	0.1	0.1	0.0	1.0	0.3	-1.0	1.0							
RTOG	0.3	0.6	0.0	0.1	0.1	0.1	0.9	0.7	-1.0	0.9	0.9						
MDC Over	0.3	0.3	0.0	0.0	0.3	0.3	0.3	0.7	-0.6	0.4	0.4	0.7					
MDC Under	0.2	0.0	0.0	0.1	0.1	0.2	0.1	0.2	-0.1	0.1	0.1	0.2	0.3				
TCP$_{median}$	0.3	0.3	0.0	0.1	0.0	0.2	0.6	0.6	-0.7	0.6	0.6	0.8	0.5	0.4			
TCP$_{min}$	0.3	0.3	0.0	0.1	0.0	0.2	0.6	0.5	-0.7	0.6	0.6	0.8	0.5	0.4	1.0		
TCP$_{max}$	0.3	0.3	0.0	0.1	0.0	0.2	0.6	0.5	-0.7	0.6	0.6	0.8	0.5	0.4	1.0	1.0	

Table 4.1: Correlation coefficients between different measures, derived from CCE and MC D$_{Mat}$ calculated doses.

4.3.1.2 Holm-Bonferroni method

In a family-wise error rate controlling method, the critical *p*-value is modified, such that the probability of there being one (or more) false positive results in the family of tests is less than the critical *p*-value. McDonald (2014) states that the most common family-wise error rate controlling method is the Bonferroni method.

In the Bonferroni method, the *p*-value of interest, e.g. 0.05, is divided by the number of tests. Assuming ten tests, the critical value would be $0.05/10 = 0.005$, and only individual tests with a *p*-value less than 0.005 are considered significant at the 0.05 level. The Bonferroni method is simple to apply, but for large numbers of tests the resulting critical value may be so low as to be impossible to meet.

A less stringent method is the Holm-Bonferroni method (Holm, 1979), also known as the Bonferroni Step Down method. In this approach, the *p*-values for individual tests are calculated, and the tests ranked by *p*-value, smallest to largest. The first *p*-value is multiplied by the number of tests N. If the result is less than the critical *p*-value, e.g. 0.05 the test is significant at that level. For the second test, the *p*-value is multiplied by $N-1$, and compared with the critical value. This continues until a non-significant test is found. This is based around the principle that if a single test is performed and the *p*-value shows it to be significant, performing a subsequent test should not retrospectively alter the previous finding. However, as each new test is performed, the *p*-value calculated must take account of the preceding tests.

4.3.1.3 Benjamini-Hochberg method

An alternative approach to the family-wise error rate controlling approach is the false-discovery rate controlling approach. An example of this is the Benjamini-Hochberg method, proposed by Benjamini and Hochberg (1995). In this method an acceptable false positive rate, is set, e.g. 10%. Tests are ranked from smallest

to largest p-value, and the p-value is compared to the Benjamini-Hochberg critical value, $\frac{rank}{N} * M$, where N is the number of tests and M is the acceptable false positive rate. The largest p-value that is less than the Benjamini-Hochberg critical value is significant, as are all smaller p-values – even if they were larger than the critical value for their rank. This approach is suitable when performing an initial analysis, to decide which measures will be investigated in more detail by subsequent experiments, so false positives will be dealt with later. If a single false positive would be unacceptable, a family-wise error rate controlling method would be more appropriate.

The Holm-Bonferroni method will be used in this work to correct the p-values calculated for the various measures.

4.3.2   Testing for Gaussian distribution

Having calculated a given measure for each case, for different dose calculation methods, an important question to ask is whether the mean or median of the two datasets are compatible. Since the same plans have been calculated with two methods, this is a matched pair problem. There are three approaches for finding differences in a paired sample for interval data (Bland, 2000):

- normal distribution, for samples with more than 100 items

- paired t-test, for samples with less than 100 items, where the differences follow a Gaussian distribution

- Wilcoxon matched pairs test, for samples with less than 100 items, where the differences do not follow a Gaussian distribution

Therefore, three questions must be answered for each dataset being compared:

1. Do the differences in a measure form an interval scale?

2. Are there enough items for the normal distribution to be used?

3. Do the differences follow a Gaussian distribution?

An interval scale is a scale in which differences between two points has a precise meaning, which does not change with the point on the scale. A difference of 1 m/s is the same at 10 m/s or 50 m/s, so speed is an interval scale. Compare this with an ordinal scale, a numerical scale in which differences between two points may not be the same at different points on the scale. An example of an ordinal scale is the American Association of Anaesthesiologists (ASA) physical status classification system, a system for describing the fitness of a patient prior to surgery. In this scale, while a performance score of two is better than a score of one, the difference between a score of one and two is not necessarily equivalent to the difference between a score of five and six. Considering these examples, the differences in Dose Volume Histogram parameters, and other measures used in this chapter, lie on an interval scale.

The second question, does the sample have more than 100 items, is trivial; there are 40 items, so use of the normal distribution for analysing the differences in the sample is not suitable.

The third question, do the differences follow a Gaussian distribution can be answered in a number of ways: -

1. By visualising the data, to see if it follows a Gaussian curve;

2. By calculating the skewness and kurtosis of the distribution, and see if they are compatible with that of a Gaussian distribution;

3. By using a more formal test for Gaussian distribution, such as the Shapiro-Wilk test.

Taking PTV V95% for CCE versus MC as an example, a histogram of the difference data was plotted, and a Gaussian with the same mean and standard deviation overlaid on this (Figure 4.1). It can be seen that the data do not conform well to a Gaussian; the distribution is asymmetric about the mean, and has some values more the three times the standard deviation from the mean. It is worth noting that the distribution is not symmetric around zero. A second graph was plotted, known as a quantile-quantile, or QQ, plot (Figure 4.2). In a QQ plot, the quantiles of the sample distribution are plotted against the theoretical quantiles, in this case for a Gaussian distribution. If the sample data follows a Gaussian, the data in a QQ plot will lie along the red line (this is analogous to plotting logarithmic data on a logarithmic scale, producing a straight line). From Figures 4.1 and 4.2, the data do not appear to follow a Gaussian.
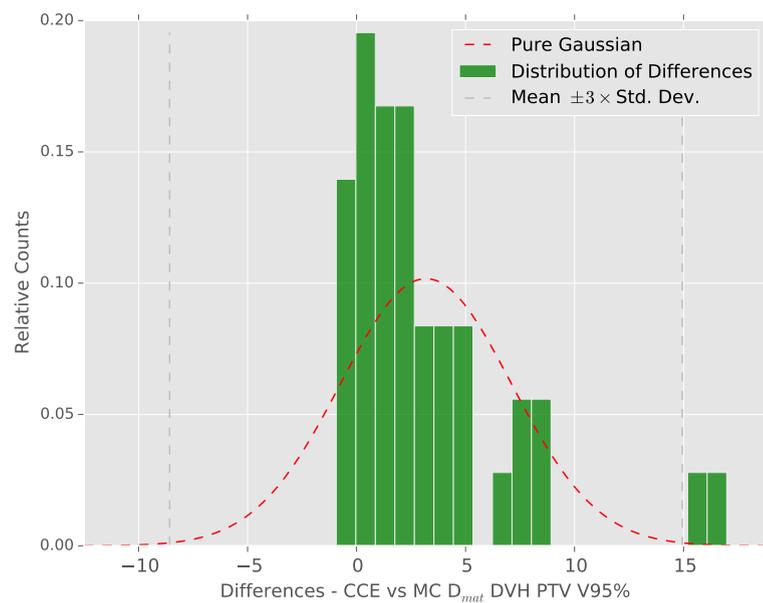
Figure 4.1: Comparison of distribution for differences in PTV V95% calculated from CCE and MC D_{Mat} dose with Gaussian distribution.
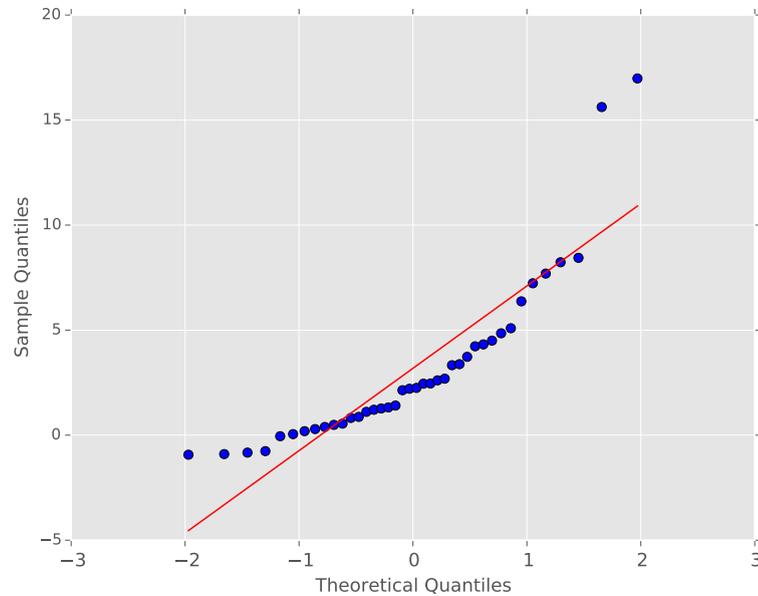
Figure 4.2: Quantile-quantile plot for differences in PTV V95% calculated from CCE and MC D_{Mat} doses.

To confirm this, the skewness and kurtosis of the differences were calculated. The skewness was calculated as 4.2 by the 'stats' module of the Scipy package[1], whereas a Gaussian distribution would have a Skew of 0. The p-value calculated was much less than 0.01, implying very strong evidence of a difference between this distribution and a Gaussian. The kurtosis was calculated as 3.2 by the 'stats' module of the Scipy package, whereas a Gaussian distribution would have a kurtosis of 0. The p-value calculated was much less than 0.01, implying very strong evidence of a difference between this distribution and a Gaussian.

The Shapiro-Wilk test (Shapiro and Wilk, 1956) calculates a statistic, often named W, which tests whether a random sample comes from a Gaussian distribution. The Shapiro-Wilk statistic and associated p-value were calculated for each measure, using the 'stats' module of the Scipy package for Python. Taking

---

[1]Jones E, Oliphant E, Peterson P, et al. SciPy: Open Source Scientific Tools for Python, 2001-, Available at: http://www.scipy.org/ (Accessed 2016-12-16)

PTV V95% as an example, the Shapiro-Wilk statistic calculated was 0.8, the *p*-value calculated was ≪0.001, so the null hypothesis (that the data is Gaussian) can be rejected, i.e. the data is non-Gaussian.

This excludes the paired *t*-test as an analysis method for the PTV V95%, CCE versus MC comparison, leaving the Wilcoxon matched pairs test, the non parametric analogue of the paired *t*-test, as the preferred method.

The same approach was applied to all measures (aside from those eliminated in section 4.3.1 due to being correlated with other measures). The results of the tests for Gaussian are shown in Table 4.2. For most measures, the kurtosis, skewness and Shapiro-Wilk all pointed to the same outcome. However, for PTV V95% and MDC_{Over}, the results of the different tests for Gaussian distribution were inconsistent. For both of these measures, the histogram and QQ plots of the data were used to determine if the data was Gaussian or not.

4.3.3 Wilcoxon matched pairs test

For each measure where the differences between the CCE and MC data were shown to be non-Gaussian, the Wilcoxon matched pairs test was performed using the Scipy package for Python. The results are shown in Table 4.3.

4.3.4 Paired T-test

For each measure where the differences between the CCE and MC data were shown to follow a Gaussian distribution, the paired *t*-test was performed using the Scipy package for Python. The results are shown in Table 4.3.

Measures	Kurtosis	Kurtosis p-value	Result	Skew	Skew p-value	Result	Shapiro-Wilk	Shapiro-Wilk p-value	Result	Overall
PTV D99%	0.9	0.4	Gaussian	1.5	0.1	Gaussian	1.0	0.4	Gaussian	Gaussian
PTV V95%	3.1	0.002	Gaussian	4.1	«0.001	Non-Gaussian	0.8	«0.001	Non-Gaussian	From Graphs: Non-Gaussian
GTV D100%	0.2	0.8	Gaussian	-1.0	0.3	Gaussian	0.9	0.006	Gaussian	Gaussian
Heart V80%	4.6	«0.001	Non-Gaussian	5.5	«0.001	Non-Gaussian	0.7	«0.001	Non-Gaussian	Non-Gaussian
Cord PRV D_{1cc}	1.1	0.3	Gaussian	1.1	0.3	Gaussian	1.0	0.2	Gaussian	Gaussian
Combined Lungs V20 Gy	0.2	0.8	Gaussian	-1.2	0.2	Gaussian	1.0	0.2	Gaussian	Gaussian
Discord. Index	2.4	0.02	Gaussian	2.6	0.008	Gaussian	0.9	0.04	Gaussian	Gaussian
Jaccard	3.8	«0.001	Non-Gaussian	4.6	«0.001	Non-Gaussian	0.8	«0.001	Non-Gaussian	Non-Gaussian
MDC_{Over}	4.0	«0.001	Non-Gaussian	3.2	0.001	Gaussian	0.8	«0.001	Non-Gaussian	From Graphs: Non-Gaussian
MDC_{Under}	3.7	«0.001	Non-Gaussian	-4.0	«0.001	Non-Gaussian	0.9	«0.001	Non-Gaussian	Non-Gaussian
TCP_{Median}	0.9	0.4	Gaussian	1.1	0.3	Gaussian	1.0	0.7	Gaussian	Gaussian

Table 4.2: Tests for whether distribution of differences between measures derived from CCE and MC D_{Mat} dose calculations follow a Gaussian distribution.

Test Name	Test Applied	Test Statistic	p-value	Critical p-value	Modified p-value	Significant
TCP_{Median}	Paired t-test	13.5	$\ll 1.0 \times 10^{-3}$	1.0×10^{-3}	$\ll 1.0 \times 10^{-3}$	True
MDC_{Under}	Wilcoxon	17.0	$\ll 1.0 \times 10^{-3}$	1.0×10^{-3}	$\ll 1.0 \times 10^{-3}$	True
Combined Lungs V20 Gy	Paired t-test	-6.0	$\ll 1.0 \times 10^{-3}$	1.0×10^{-3}	$\ll 1.0 \times 10^{-3}$	True
Heart V80%	Wilcoxon	41.0	$\ll 1.0 \times 10^{-3}$	1.0×10^{-3}	$\ll 1.0 \times 10^{-3}$	True
PTV V95%	Wilcoxon	42.5	$\ll 1.0 \times 10^{-3}$	1.0×10^{-3}	$\ll 1.0 \times 10^{-3}$	True
Jaccard	Wilcoxon	103.0	$\ll 1.0 \times 10^{-3}$	1.0×10^{-3}	$\ll 1.0 \times 10^{-3}$	True
Discord. Index	Paired t-test	3.5	1.2×10^{-3}	1.0×10^{-3}	6.0×10^{-3}	False
PTV D99%	Paired t-test	3.5	1.3×10^{-3}	1.0×10^{-3}	5.2×10^{-3}	False
Cord PRV D_{1cc}	Paired t-test	2.7	1.1×10^{-2}	1.0×10^{-3}	3.3×10^{-2}	False
MDC_{Over}	Wilcoxon	261.5	4.6×10^{-2}	1.0×10^{-3}	9.2×10^{-1}	False
GTV D100%	Paired t-test	1.0	3.1×10^{-1}	1.0×10^{-3}	3.1×10^{-1}	False

Table 4.3: Results for whether CCE and MC D_{Mat} give significantly different values, tested by either T-test or Wilcoxon Matched Pairs test.

4.3.5 Significant results

Once the test statistics and *p*-values were calculated as described above, the Benjamini-Hochberg method was used to correct the calculated *p*-values. The original and corrected *p*-values are shown in Table 4.3, which shows that the differences in the measures DI-CI, PTV D99%, Cord PRV D_{1cc}, MDC_{Over} and GTV D100% between CCE and MC D_{Mat} are not significant. The absence of a statistically significant difference between the doses calculated by CCE and MC for these parameters, based on investigation of 40 cases, gives confidence that CCE is accurate for these parameters. These parameters will not be discussed further.

4.3.6 Boxplots

Boxplots, or box and whisker plots, are a way of visualising the distributions of data. The boxplot shows various statistical measures in the following way:

- The First Quartile, or 25th percentile, is represented by the bottom of the

box;

- The Third Quartile, or 75th percentile, is represented by the top of the box;

- The Mean of the distribution is represented by a red square;

- The Median, or Second Quartile, or 50th percentile, of the distribution is represented by a horizontal red line;

The whiskers of the boxplot may be defined in various ways, but in this work the following approach is used:

- The Upper cap, or line at the end of the whisker, is the highest datum within 1.5 times the Inter-Quartile Range (IQR);

- The Lower cap is the lowest datum within 1.5 times the IQR;

- Points outside 1.5 times the IQR, i.e. outliers, are plotted as a '+' symbol;

Boxplots are useful alternative to histograms when comparing multiple datasets, as several boxes can be plotted side by side. This makes boxplots a useful tool for comparing the distribution of measures, such as PTV V95%, for the four different dose calculation methods PBE, CCE, MC D$_{Mat}$ and MC D$_{Wtr}$. As for the table-graphics of section 3.3, two graphs are produced for each measure: the first compares PBE and MC D$_{Wtr}$ and the second compares PBE, CCE and MC D$_{Mat}$.

Figure 4.3(a) compares the PTV V95% for PBE and MC D$_{Wtr}$. It can be seen that PTV V95% for PBE has a very small spread of values, with a mean value of nearly 100%. The boxplot for the MC D$_{Wtr}$ simulations shows a drop in the mean and median, and a marked increase in the IQR of values. Figure 4.3(b) compares the PTV V95% for PBE, CCE and MC D$_{Mat}$. CCE also shows a drop in the mean and median compared to the PBE results, and a marked increase

in the IQR of values. The MC D$_{Mat}$ results show a reduction in the mean and median values compared with the PBE data, of the order of the IQR.
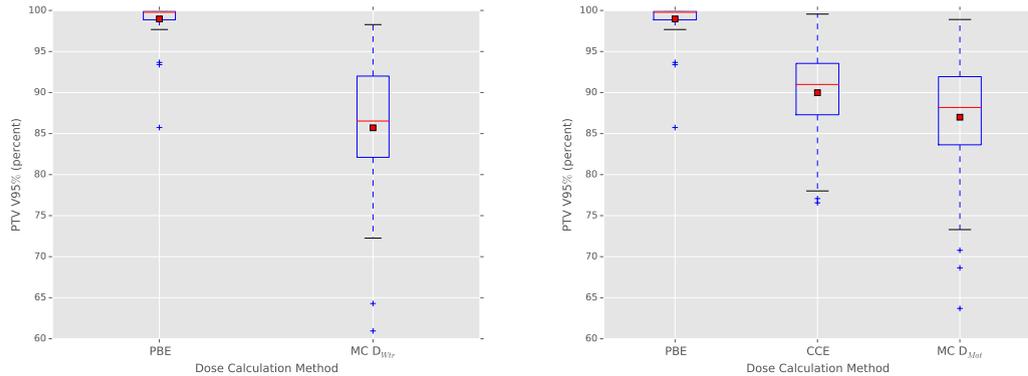
(a) PTV V95%: PBE and MC D$_{Wtr}$ **(b)** PTV V95%: PBE, CCE and MC D$_{Mat}$

Figure 4.3: Boxplots comparing PTV V95% for PBE, CCE, MC D$_{Mat}$ and MC D$_{Wtr}$.

Figures 4.4(a) and (b) compare the values for the heart V80% for the different calculation methods, and shows that the mean, median and IQR values are comparable across the different dose calculation methods. The Wilcoxon test shows that the difference in heart V80% between CCE and MC D$_{Mat}$ is significant (Table 4.3). This is due to a consistent trend downward trend in the majority of cases, as shown in the table-graphic of Chapter 3 (Figure 3.34).

Figures 4.5(a) and (b) compare the values for the lung V20 Gy for the different calculation methods. Figure 4.5(a) shows a reduction in the mean and median values from the PBE to the MC D$_{Wtr}$, just as Figure 4.5(b) shows a reduction in the mean and median values from the PBE to the CCE dose calculation methods. There is an increase in the mean and median values from the CCE to the MC D$_{Mat}$, which may be due to the beam penumbra broadening further in low density materials in the MC simulations compared to the analytical algorithms.

Figure 4.6(a) shows that the mean and median values for the Jaccard-CI decreases between the PBE and MC D$_{Wtr}$, but that the decrease is small compared to the

(a) heart V80% : PBE and MC D$_{Wtr}$ **(b)** heart V80% : PBE, CCE and MC D$_{Mat}$
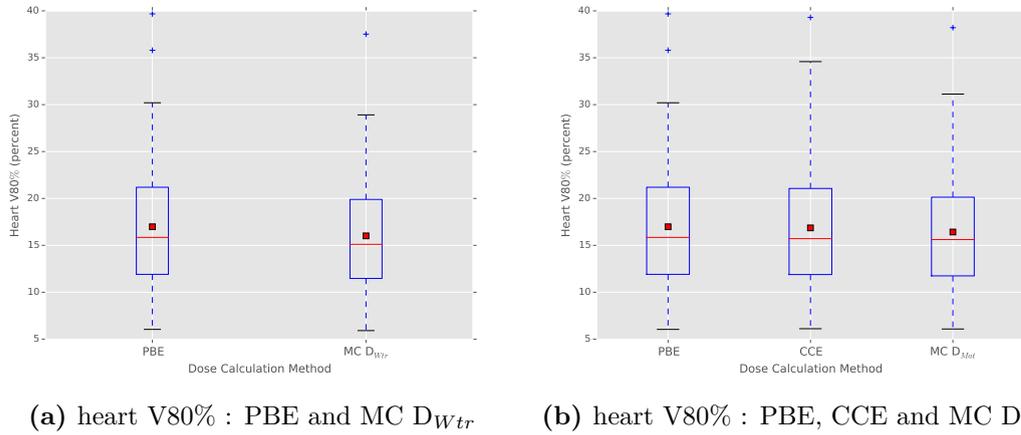
Figure 4.4: Boxplots comparing heart V80% for PBE, CCE, MC D$_{Mat}$ and MC D$_{Wtr}$.

IQR. The mean and median values for the PBE and CCE are comparable, as shown in Figure 4.6(b), but there is a decrease in the mean and median values of the Jaccard-CI between the CCE and MC D$_{Mat}$ dose calculation methods.

Figure 4.7(a) shows that a clear reduction in the mean and median values of the MDC-Under between PBE and MC D$_{Wtr}$. Figure 4.7(b) shows that a large reduction in the mean and median values of the MDC-Under between the PBE and CCE algorithms, but an increase between the CCE and MC D$_{Mat}$ dose calculation methods.

Figure 4.8(a) shows that a large reduction in the mean and median values of the TCP-Median between PBE and MC D$_{Wtr}$. Figure 4.8(b) shows that a large reduction in the mean and median values of the TCP-Median between PBE and CCE, and a further reduction between CCE and MC D$_{Mat}$. This behaviour is consistent with the equivalent graphs for TCP-Min and TCP-Max, so this result is not dependent on the values of the parameters used in the TCP model.

(a) PBE and MC D_{Wtr} **(b)** PBE, CCE and MC D_{Mat}
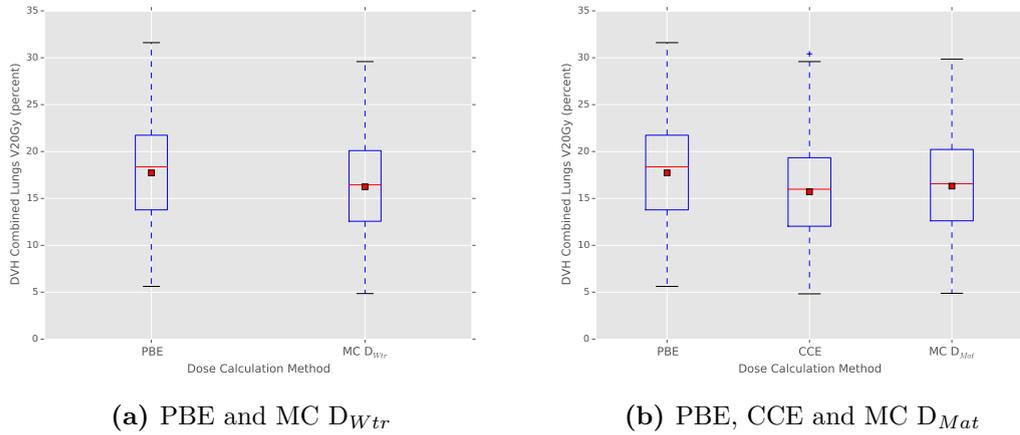
Figure 4.5: Boxplots comparing Combined lung V20 Gy for PBE, CCE, MC D_{Mat} and MC D_{Wtr}.

4.4 Discussion

4.4.1 Discussion of selected cases

Figure 4.9 shows a comparison of the DVHs for CCE and MC D_{Mat}, for case #9 (i.e. the ninth case in the group of 40 cases used in the main body of the study). For the PTV, the plotted line for MC D_{Mat} calculated dose is slightly higher at the 95% level, but lower elsewhere, compared with the CCE calculated dose. For the cord PRV, there are differences where there are steps in the plotted lines, as seen in section 3.4.1. For the heart, the dose from MC D_{Mat} is generally lower than the dose calculated by the CCE algorithm. For the combined lungs, the there is very little difference between the two dose calculation methods.

Figures 4.11 (a) and (b) show the 95% isodose (47.5 Gy, legend within Figures) for the CCE and MC D_{Mat} dose calculation, for case #9. It can be seen in Figure 4.11(b) that the coverage of the PTV by the 95% isodose is better in the region overlapping with the lungs, compared with Figure 4.11(a). As in the cases

(a) Jaccard-CI: PBE and MC D$_{Wtr}$ **(b)** Jaccard-CI: PBE, CCE and MC D$_{Mat}$
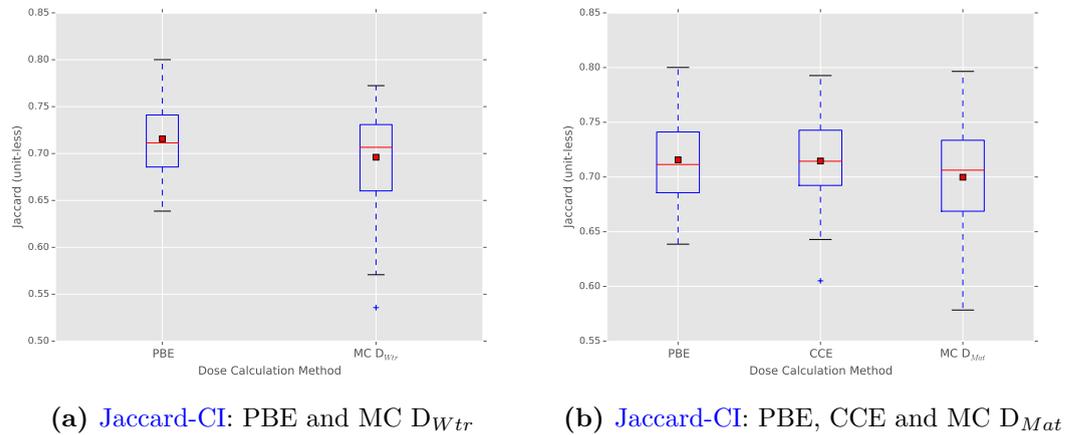
Figure 4.6: Boxplots comparing Jaccard-CI for PBE, CCE, MC D$_{Mat}$ and MC D$_{Wtr}$.

discussed in section 3.4.1, the superior modelling of scatter in the MC simulations, compared with the CCE algorithm results in the differences seen in these regions.

However, the coverage where the PTV overlaps with the bone of the vertebra is better of the CCE case Figure 4.11(a) than the MC D$_{Mat}$ case Figure 4.11(b). There are also 'holes' in the 95% isodose throughout the PTV. These 'holes' are directly anterior of the lateral edges of the vertebra, so it is likely that the differences observed are due to differences in the modelling of attenuation in high density tissues between the CCE and MC D$_{Mat}$ dose calculation methods.

Figure 4.10 shows a comparison of the DVHs for CCE and MC D$_{Mat}$, for case #37. For the PTV, the plotted line for MC D$_{Mat}$ demonstrates lower coverage than that for the CCE algorithm. For the cord PRV, the DVH plots deviate below 90% of the volume, while there is god agreement for the high dose region. The DVH for the heart and combined lungs show good agreement, although the dose calculated by MC D$_{Mat}$ is generally higher that that from CCE for the heart, while the reverse is true for the combined lungs.

Figures 4.12 (a) and (b) show the 95% isodose (47.5 Gy isodose line, legend within figures) for the CCE and MC D$_{Mat}$ dose calculation, for case #37, which showed

(a) Under MDC: PBE and MC D$_{Wtr}$ **(b)** Under MDC: PBE, CCE and MC D$_{Mat}$
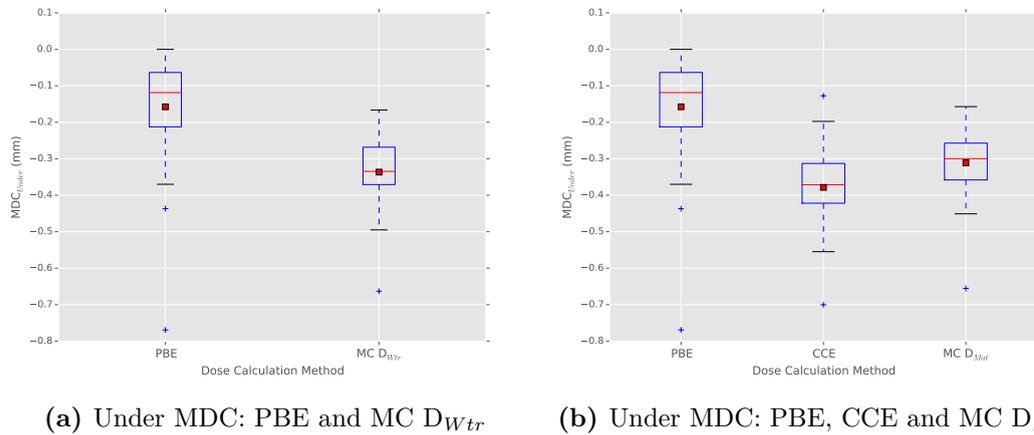
Figure 4.7: Boxplots comparing Under Mean Distance to Conformity for PBE, CCE, MC D$_{Mat}$ and MC D$_{Wtr}$.

a large difference in PTV V95%, produced by CERR. Figures 4.12 (a) and (b) both show poor coverage where the PTV overlaps the right lung (left side of the image). Figure 4.12 (b) shows that the coverage of the PTV for the MC D$_{Mat}$ is poorer than for the CCE algorithm, particularly where the PTV overlaps with the high density vertebra. This suggests poorer agreement between the CCE and MC dose calculations is related to non-soft tissue regions in the PTV. Visually, the differences are not large, but as this behaviours is repeated over several slices, the combined effect produces one of the largest changes in PTV V95%.

4.4.2 Correlation with PTV material

It has been shown that re-calculating oesophageal radiotherapy plans with MC gives a value for PTV V95% that is significantly different to that obtained with the original CCE calculation. For sections 3.4.1 and 4.4.1, differences occur in regions where the PTV overlaps lung or bone. It would be useful if a prediction could be made regarding which cases would benefit the most from MC simulation. From examining isodoses for outliers, the composition of the PTV appears to be
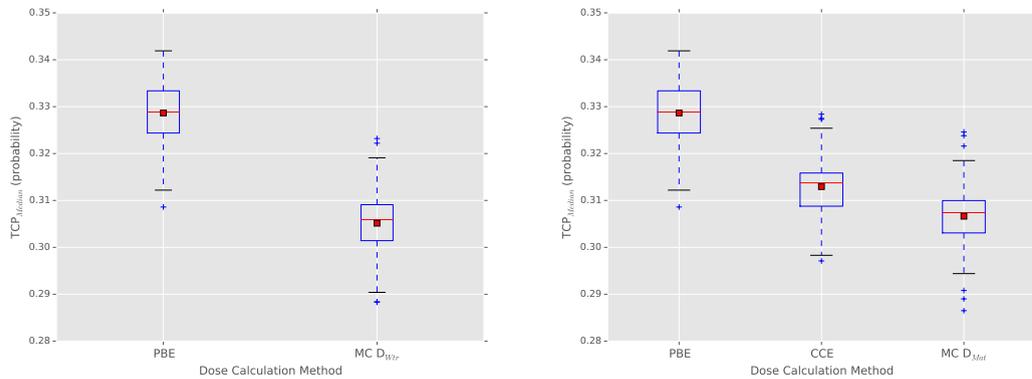
(a) TCP Median: PBE and MC D$_{Wtr}$ **(b)** TCP Median: PBE, CCE and MC D$_{Mat}$

Figure 4.8: Boxplots comparing Tumour Control Probability Median for PBE, CCE, MC D$_{Mat}$ and MC D$_{Wtr}$.

a factor. Could the percentage of the PTV that is lung, for example, predict whether MC should be used?

A common tool for measuring the effectiveness of the predictive power of a test is the Receiver Operator Characteristic (ROC) curve. A predictive test that aims to answer a true or false question may have four outcomes: -

1. The test gives a positive result, when the result should be positive (a **true positive**);

2. The test gives a positive result, when the result should be negative (a **false positive**);

3. The test gives a negative result, when the result should be negative (a **true negative**);

4. The test gives a negative result, when the result should be positive (a **false negative**).
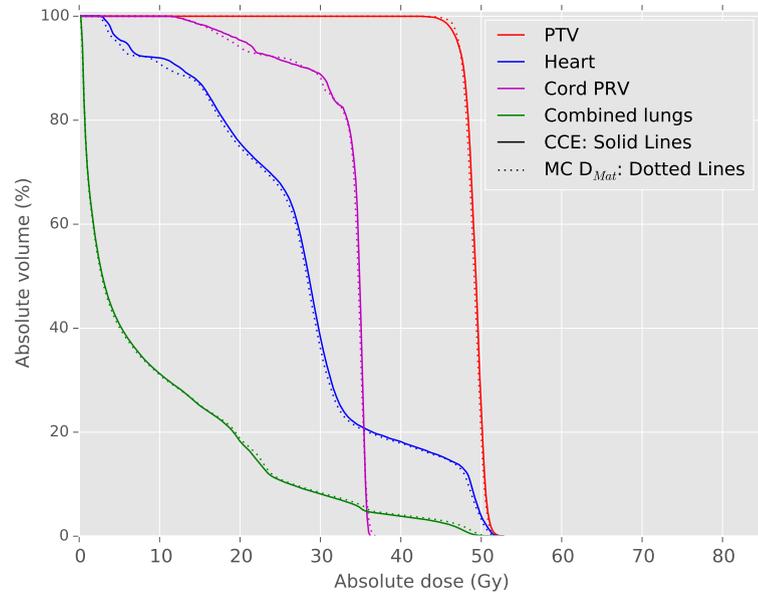
Figure 4.9: Comparison of DVH parameters for CCE and MC D_{Mat}, for case #9 of 40.

The ratio of true positive test results to the 'real' number of positive outcomes is known as the sensitivity. A related quantity is the specificity, which the the ratio of the true negative test results to the 'real' number of negative outcomes. Ideally a test will have 100% sensitivity, so correctly identifies all true outcomes, and 100% specificity, so will not incorrectly identify any false outcomes as true. However, many tests in which a threshold value is used to make a true or false prediction will not behave in this ideal way; as the threshold is varied to improve the sensitivity, the specificity may be compromised. By plotting the sensitivity (or True Positive Rate) against 1-specificity (also known as the False Positive Rate) as the threshold value is varied, a ROC curve is produced. A ROC curve may be used to determine the effectiveness of a test, to compare different tests, and choose a suitable threshold. Figure 4.13 shows three ROC curves used to predict if the change in PTV V95%, is greater than 5%. The first curve uses the percentage of air or lung in the PTV. The second curve uses the percentage of
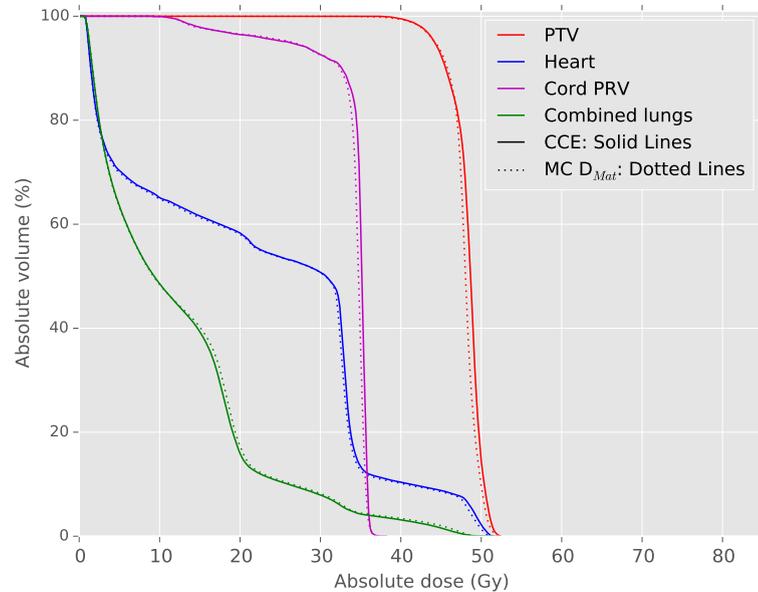
Figure 4.10: Comparison of DVH parameters for CCE and MC D$_{Mat}$, for case #37 of 40.

bone in the PTV. The third curve uses the percentage of air, lung or bone in the PTV; this is equivalent to the percentage of the PTV that is not soft tissue.

An ideal test will have a point at the top left extreme of the plot, sensitivity = 1, and 1 - specificity = 0. A test which is no better that random will have points lying along the x = y line. It can be seen than none of these three parameters predicts a 5% change in PTV ideally, but all three are better than random. A way to compare these tests is to calculate the Area Under the Curve (AUC). An ideal test will have an AUC of 1, a test that is no better than random will have a AUC of 0.5. A test with an AUC of less than 0.5 has higher false positive and false negative results than would be expected by a random guess. In this case the logic of the test can be reversed to give an AUC greater than 0.5, so the test has predictive ability.

Comparing the three AUC values, the percentage of PTV that is either air, lung or bone is the best test. However, the point on the ROC curve closest to the

top left corner is not necessarily the best value. Consider a screening test for cancer. One approach to setting the threshold used in the test would be strike the best balance between sensitivity and specificity. However, choosing a threshold value such that the false negative rate is minimised would mean that number of patients with cancer that is 'missed' by the test is minimised. Even if this comes at the expense of a lower sensitivity, i.e. more false positive results meaning more patients without cancer are referred for further tests, it is likely that this will be the preferred approach.

Consider the three points in Figure 4.13, P1, P2 and P3. Using a threshold value of 31% of PTV that is bone, will correctly identify 53% of the cases that will see a 5% change in PTV V95%, without classifying any that do not need MC as needing it. It will however 'miss' 47% of those that would benefit from MC.

P3 relates to applying a threshold of 39% to the PTV that is either Air, Lung or Bones. This threshold identifies all cases that would have a greater than 5% reduction in PTV V95%, but also includes 78% of those that would not. This is not a particularity useful test, if that many cases are to be simulated with MC, all cases may as well be.

P2 picks of 87.5% of those needing MC, and only 41% of those not needing MC. This is still not ideal, but is a useful test, and could be used to aid prioritisation of MC resources. It also suggests modification of this parameter (i.e. composition of PTV plus a margin) may yield a more effective prediction.

Assuming that 10% reduction in PTV is the criteria used to determine if MC is required, then P4 in Figure 4.14 picks up 68% of those requiring MC, and none where CCE is sufficient. P5 identifies all cases requiring MC, and only 37% of those that do not require MC, using the percentage of the PTV that is air, lung or bone greater than 62% as the threshold. It is therefore recommended that if a PTV has greater than 62% of the volume consisting of air, lung or bone, then MC simulation should be used where available.

4.4.3 Comparison of results with other published work

As discussed in section 2.6.3, there are few studies comparing different calculation methods for oesophageal RT treatments. Wills et al. (2009) found a reduction in PTV V95% of 9% when comparing the PB and CC algorithms for oesophageal plans. This agrees with the results presented in Figure 4.3(b). Wills et al. (2009) found that the doses to the OARs changed by less than 2% between the two calculation algorithms, which agrees with the comparisons of the PBE and CCE DVH parameters for the heart V80% in Figure 3.34, cord PRV D$_{1cc}$ in Figure 3.35, and lung V20 Gy in Figure 3.36. As no MC simulations were included in the work of Wills et al. (2009), no comparison can be made with the results obtained in the current work for MC calculations.

Irvine et al. (2004) found that the minimum dose to the PTV reduced by up to 23% when the PB plan was re-calculated with the CC algorithm. Although there was substantial variation between the 10 cases examined, the PTV minimum dose always decreased when re-calculating with the CC algorithm. The minimum dose to the PTV was not examined in the current work, as it is not a parameter required by the SCOPE1 trial. The most comparable parameter in the current work is the PTV D99% (the minimum dose to the PTV could be called PTV D100%, the minimum dose received by 100% of the PTV volume). Figure 3.31(b) shows qualitative agreement with the results of Irvine et al. (2004); the PTV D99% always decreases when re-calculating with CCE, and there is substantial variation in the magnitude of the reduction. No reductions in PTV D99% of 20% were obtained in this work; the largest reduction shown in Figure 3.31(b) is approximately 15%. D99% and D100% are different parameters, so these differences in the results are not contradictory.

As in the work of Wills et al. (2009), Irvine et al. (2004) compared the 'classic' PB and CC algorithms, rather than the enhanced versions used in this work

(section 2.4.6). As the work of Irvine et al. (2004) did not involve comparisons with MC calculated doses, no comparison can be made with the results obtained in the current work for MC calculations.

Haga et al. (2014) found differences in the PTV of less than 2% between the analytical algorithm used and the MC simulation, for the single oesophageal case examined. This agrees with the mean changes in PTV found in the current work between CCE and MC D_{Mat} calculations, as presented in Table 3.6 and Figure 4.3(b), although there is substantial variation between cases in the current work. Haga et al. (2014) found differences in the mean lung dose of less than 2%, while the results in the current work (Table 3.6 and Figure 4.5(b)) show differences within 1% for heart V80% . This may be due to the differences in the measures used for lung, V20 Gy versus mean dose.

Haga et al. (2014) found differences in the maximum dose to the spinal cord of 5%. The most comparable parameter in this work is the Cord PRV D1*cc*. The mean change in Cord PRV D1*cc* between the CCE and MC D_{Mat} calculations presented in Table 3.6 was 0.05 Gy, or approximately 1.3%, and the analysis in Chapter 4 revealed no significant change in the Cord PRV D1*cc* when comparing CCE and MC D_{Mat} calculations. This may reflect that the absolute maximum dose to an OAR is a less stable parameter than D1*cc*. The maximum dose for a MC simulation is likely to be affected by the statistical fluctuations in the voxels. Haga et al. (2014) only examined a single oesophageal case, so these results cannot be extrapolated to other patients.

Jang et al. (2006) compared doses at discrete points rather than comparing DVH parameters, so direct comparison of results cannot be performed. Despite finding point differences of up to 10%, the DVHs presented for the lung cases show good agreement between Pinnacle and MC simulations. The DVHs presented in sections 3.4.1 and 4.4.1 also show the DVH graphs appear very similar despite differences being found for individual DVH parameters, and differences being

apparent for isodose distributions.

No comparisons can be made with the results presented in Table 4.3, showing which of the PTV parameters were found to change significantly between the CCE and MC D$_{Mat}$ calculations, as no know studies have taken the approach used in the current work, for oesophageal or other treatment sites.

4.5   Conclusion

40 oesophageal cases have been simulated on the RTGrid. For all cases a number of measures (DVHs parameters, Conformity Indices, Mean Distance to Conformity values and TCP values) have been calculated for the four different sets of doses produced, PBE, CCE, MC D$_{Mat}$ and MC D$_{Wtr}$. For measures that were correlated with other measures, all but one from each 'group' were excluded. For each remaining measure, the differences in the values calculated from the CCE and MC D$_{Mat}$ dose distributions produced a distribution of values; each distribution was tested to determine if it was a normal distribution. For normally distributed data, the paired *t*-test was used to test if the differences were significant or not; for non-normal distributions, the Wilcoxon paired test was used.

For each of the *t*-tests and Wilcoxon tests, the *p*-value was calculated. Due to the number of measures analysed, the Holm-Bonferroni method was used to adjust the critical *p*-values. It was found that differences in TCP-Median, MDC-Under, combined lungs V20Gy, heart V80% PTV V95% and Jaccard-CI were statistically significant. The measures DI-CI, PTV D99%, Cord PRV D$_{1cc}$, MDC$_{Over}$ and GTV D100% between CCE and MC D$_{Mat}$ were not found to be significant.

Boxplots were used to visualise the differences in the measures for different dose calculation methods, and shown that the PTV V95% fell between PBE and CCE, and then fell further for the MC D$_{Mat}$ dose calculation.

Isodose plots show qualitatively that the agreement is poorer between different dose calculation methods in regions of lung, bone, or other non-soft tissue regions. It was found that the composition of the PTV shows some utility for predicting those cases that would benefit the most from MC. It is recommended that if the percentage of non-soft tissue voxels in the PTV exceeds 62%, MC should be used where available (P5 in Figure 4.14).
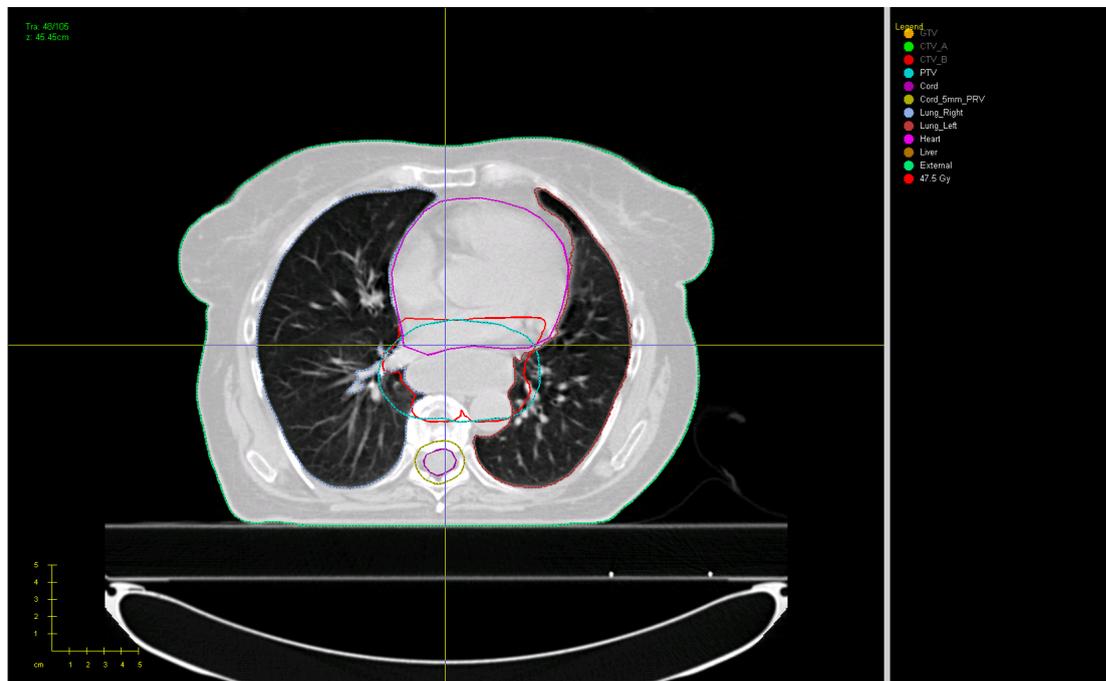
(a) 95% isodose calculated from CCE dose distribution.
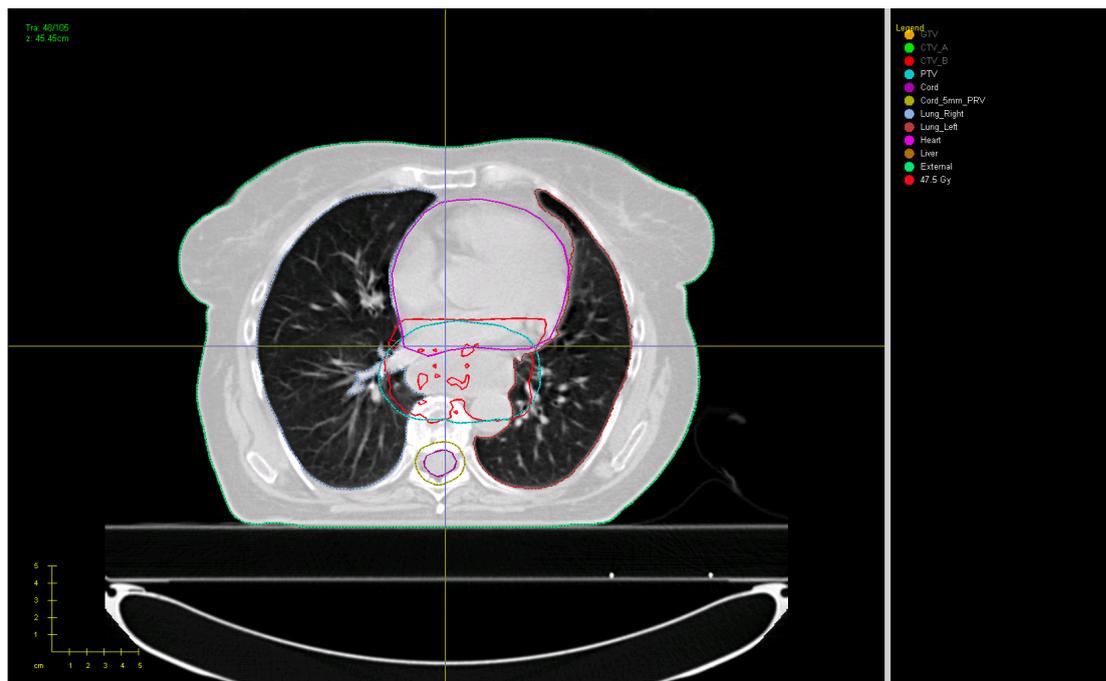
(b) 95% isodose calculated from MC D_{Mat} dose distribution.

Figure 4.11: Screen-shots comparing 95% isodose coverage of PTV in CT slice, for case #9 of 40.
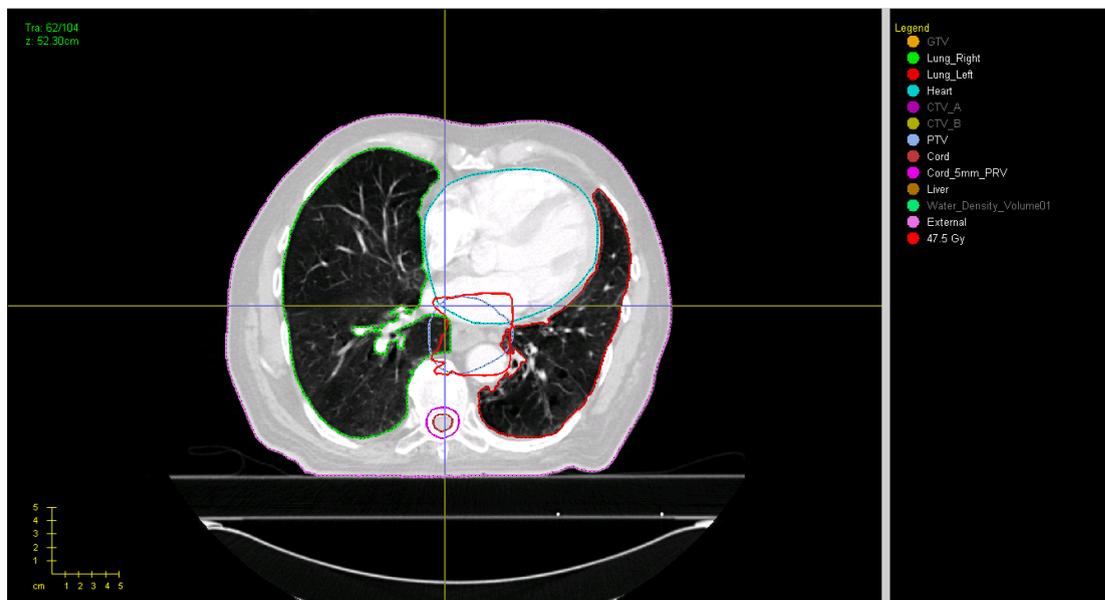
(a) 95% isodose calculated from CCE dose distribution.
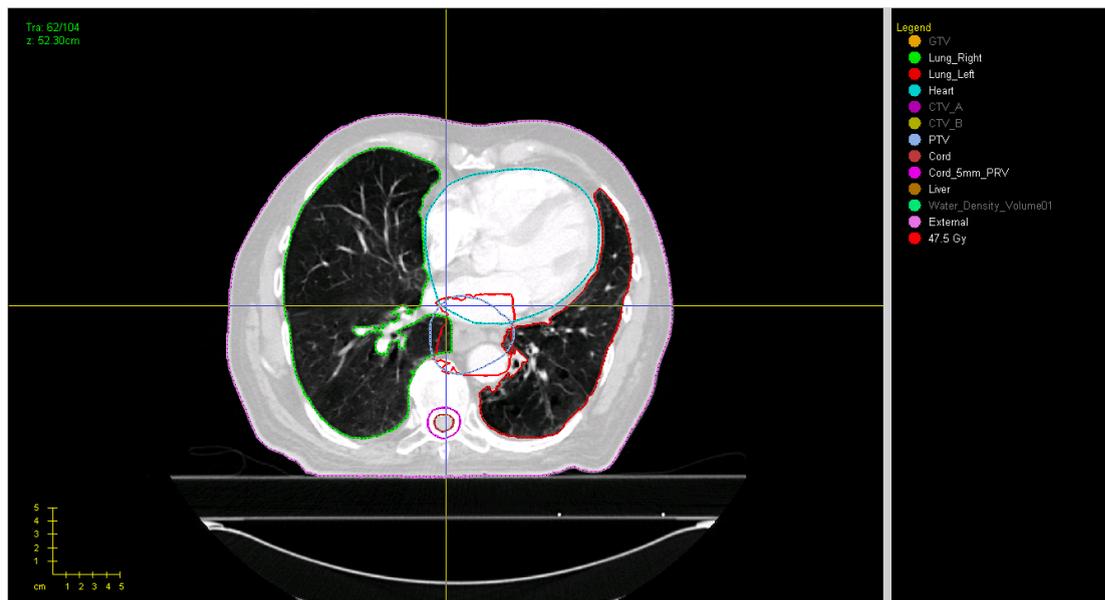
(b) 95% isodose calculated from MC D_{Mat} dose distribution.

Figure 4.12: Screen-shots comparing 95% isodose coverage of PTV in CT slice, for case #37 of 40.
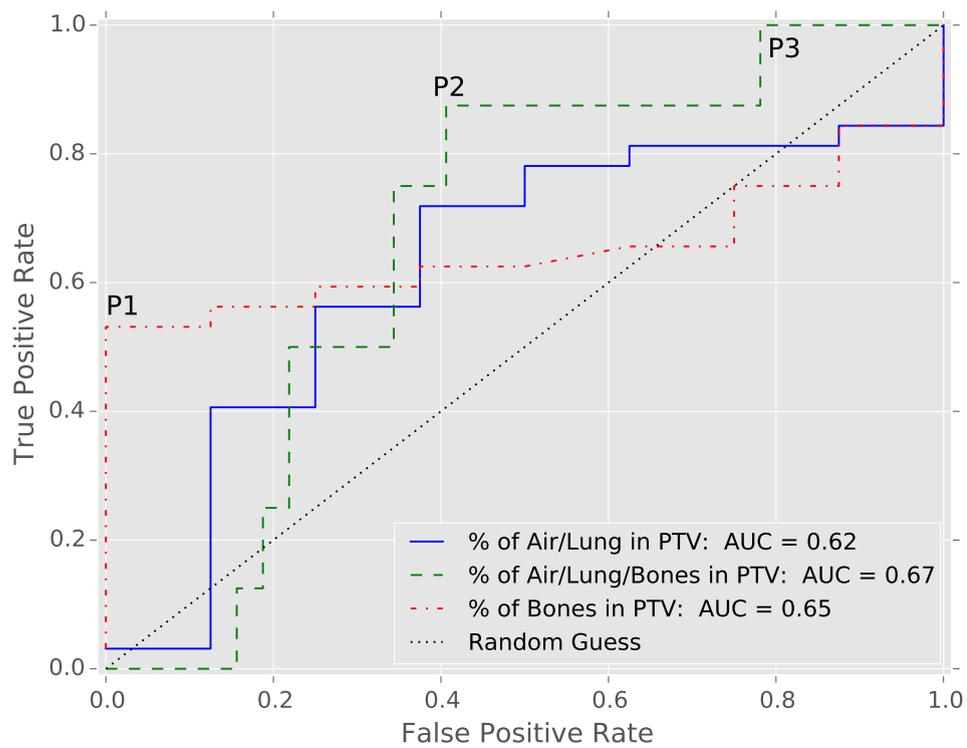
Figure 4.13: Receiver Operator Characteristic curves for the use of PTV composition to predict a change in V95% greater than 5%.
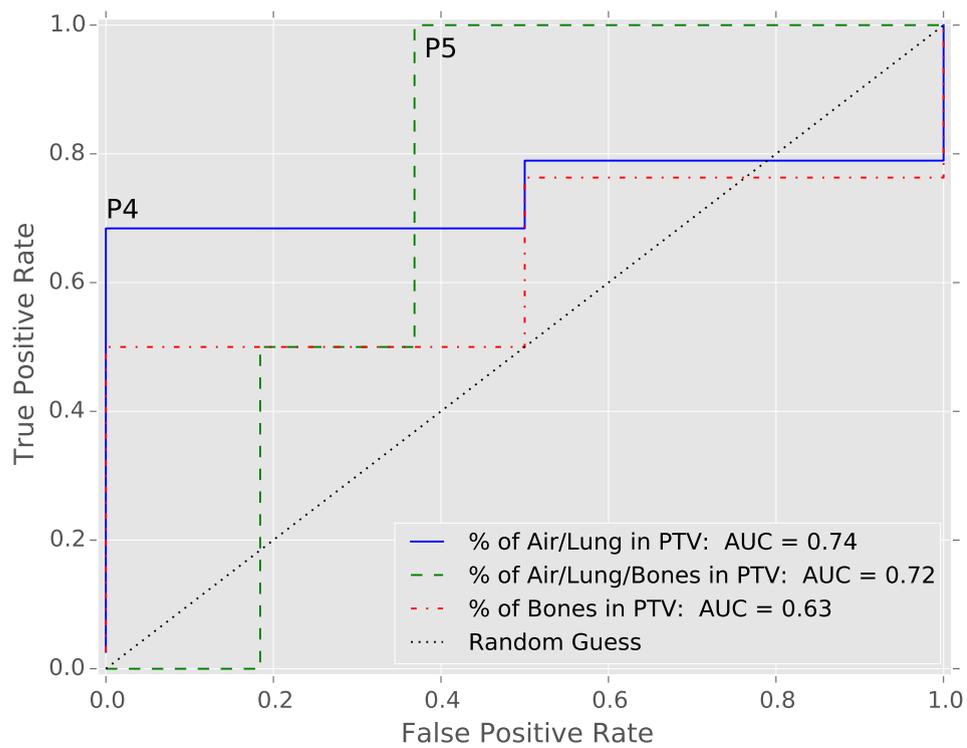
Figure 4.14: Receiver Operator Characteristic curves for the use of PTV composition to predict a change in V95% greater than 10%.

Chapter 5

Conclusions

5.1   Conclusions

This work has examined the effect of the method of dose calculation on RT plans
for oesophageal cancer. Oesophageal RT treatment plans created at Velindre
Cancer Centre following the SCOPE protocol were used for this study. The plans
had been created using the Oncentra MasterPlan TPS, and calculated using the
PBE algorithm. These plans were then re-calculated using the CCE algorithm,
and MC simulations performed. For these MC simulations the BEAMnrc and
DOSXYZnrc codes were chosen due to their accurate physical models for the en-
ergy ranges and materials of interest to photon beam RT, and their active user
base. The RTGrid was used to execute BEAMnrc and DOSXYZnrc simulations
on an HTCondor pool. This allowed existing PCs in the Medical Physics depart-
ment to be used overnight when they were (for the most part) unused. This was
a cost-effective approach to accessing a computing platform capable of executing
up to 200 'jobs' in parallel.

Before oesophageal RT treatment plans could be simulated on the RTGrid, a
number of changes needed to be made to the profile used to create the jobs
executed on the HTCondor pool. A CTRamp, OesRAMP, was produced from
scans of a CIRS-062M phantom. The EDW component module was incorporated
into the RTGrid profile, to allow the simulation of plans including Varian EDW

fields. Code was developed to remove the high density intra-venous contrast from CT scans and to ignore data on the CT scan outside the external contour. Further code was developed to create artificial volumes at the superior and inferior extent of the CT scan, so that the MC simulation followed the approach of the TPS calculations.

Initially, 12 oesophageal treatment plans were simulated, and DVH parameters calculated from the MC dose distributions. Parameters commonly reported from the DVHs, such as PTV V95%, were compared for the CCE and MC, and sample size calculations were performed. From this analysis, it was decided to simulate 40 clinical cases for the main body of the work, primarily to investigate the coverage of the PTV with the 95% isodose.

For these 40 oesophageal cases, four different dose calculation methods, PBE, CCE, MC D_{Mat} and MC D_{Wtr} were considered. For each dose calculation method, a number of measures (DVHs parameters, Conformance Indices, MDC values and TCP values) were calculated. For those measures where the differences between the CCE and MC D_{Mat} datasets followed a Gaussian distribution, the paired *t*-test was used to test if the differences were significant. Where the differences between the datasets were non-Gaussian, the Wilcoxon matched pairs test was used to test if the differences were significant. Some of these differences may have appeared to be significant by chance, due to the large number of comparisons made. To correct for this effect, firstly any measures that were correlated with other measures were excluded. Then, the Holm-Bonferroni method was used to adjust the critical *p*-values for the statistical tests performed. It was found that differences in TCP-Median, MDC-Under, combined lungs V20 Gy, heart V80%, PTV V95% and Jaccard-CI were statistically significant.

Analysis of isodose plots showed that the agreement is poorer between different dose calculation methods in regions of lung, bone, or other non-soft tissue regions. It was found that if the percentage of non-soft tissue voxels in the PTV

exceeds 62% this predicts those cases in which the PTV V95% would decrease by 10% or more with a sensitivity of 68% and a specificity of 100%. It is therefore recommended that MC should be used where available in such cases (P5 in Figure 4.14).

Comparing the results in this work for DVH parameters for the PTV for the changes between the PBE and CCE algorithms, against other published results found good agreement with Wills et al. (2009) and Irvine et al. (2004). The differences observer for DVH parameters for the lungs and heart, comparing the CCE and MC D_{Mat} dose calculations agreed with the results presented by Haga et al. (2014), although the difference observed by Haga et al. (2014) for the maximum dose to the spinal cord between analytical calculation and MC simulation (5%) exceeds any differences in this work for cord PRV D1*cc*. As no known work has determined the statistical significance of a range of measures calculated for analytical calculations and MC simulations, no comparisons of this aspect of the work can be made with the literature.

In summary, the main findings of this work are as follows: -

- MC produces statistically significantly different PTV V95% compared to CCE when planning CFRT for oesophageal cancer;

- It is necessary to consider using MC when performing outcome modelling, as variation in TCP arising from different dose calculation methods is statistically significant. This suggests that the differences in coverage of the PTV by the 95% isodose will lead to clinically significant differences;

- The use of MC is indicated if non-soft tissue voxels make up $> 60\%$ of the PTV.

5.2 Further work

Possible areas for further study are presented below.

5.2.1 Extension to other treatment sites

For this work oesophageal treatment plans were chosen due to the significant degree of heterogeneous material in the calculation region, and the paucity of studies of this treatment site. Treatment plans for lung cancer are also performed in regions of significant heterogeneities. While there have been a large number of studies published comparing analytical and MC dose calculation methods for lung RT, the sample size has often not been considered in the comparisons. Of particular interest for future work would be trials involving dose escalation. In such trials, the dose to the target is increased as high as possible (up to a limit) while keeping the doses to OARs below set targets. Any inaccuracies in the dose calculation for OARs could have a marked impact on the prescribed dose.

5.2.2 Automation of CTRamp creation and validation

During this work a CTRamp, OesRAMP, was created for a specific CT scanner protocol, the ctcreate executable re-compiled with the material and density boundaries that define the ramp, and the resulting egsphant files examined to ensure the materials used matched the anatomical information. While aspects of this task were automated, notably the generation of the fitting parameters in Equation 3.3, other aspects were time consuming. It would be possible to automate the whole process, for example: -

- For a CT scan of the CIRS-062 phantom, use the existing code to generate the relation between HU and density for the CT scan of interest;

- Compile a modified version of ctcreate.exe with a new CTRamp, relating to the CT protocol of interest;

- Generate graphs plotting the CTRamp, and measured density versus calculated density, to validate the CTRamp;

- Run the ctcreate executable for a reference CT scan, to generate an egsphant file;

- Perform automated analysis on the egsphant file to ensure the boundaries of the material bins have been set correctly, for example by testing the material assigned to voxels within ROIs drawn on the CT;

- If material assignment is not satisfactory, adjust the boundaries used, recompile the ctcreate executable, re-create the egsphant file and repeat the analysis of material assignment;

- Repeat this iterative process until material assignment is satisfactory.

This would streamline the process of producing CTRamps, and would facilitate the study of RT plans for different treatment regions with MC.

5.2.3   Newer source models

As described in section 3.2.2, the RTGrid profile used in the current work performs simulations with enough histories to achieve the required statistical uncertainty for each beam. Therefore, for plans with six beams, the statistical uncertainty at the isocentre will be better than that for a four beam plan. If the four beam plan is simulated with sufficient histories, this implies that a six beam plan is simulated with too many histories. Also, the feature of the RTGrid to end a simulation once, for example, 95% of the jobs have completed was not used, in case the 5% of jobs not performed all related to the same beam. (If a PC in

the HTCondor pool is used intermittently, the execution of jobs assigned to this PC will be repeatedly suspended, and restarted. If all jobs must finish before a simulation is completed, the completion of a simulation may be delayed by a single job, which is delayed by repeated suspension to execution.)

Using source model SOURCE21, by Lobo and Popescu (2010) would allow the simulation of all the beams in a plan to be combined into a single simulation. While this simulation would still be split into multiple jobs, the number of histories used could be based on the statistical uncertainty to be set for the plan as a whole. This would also allow the simulations to be ended after 95% (or other suitable percentage) without fear of compromising the statistics for one beam. As SOURCE21 is designed for dynamic treatments, adopting SOURCE21 would allow the same profile to be used for CFRT as well as IMRT and VMAT.

5.2.4 Better predictive model for plans requiring Monte Carlo

In section 4.4.2, methods to predict which oesophageal plans would benefit most from MC were discussed. The percentage of voxels in the PTV which are not soft-tissue showed some utility for predicting plans that would have large changes in PTV V95%. It may be expected that the composition of voxels just outside the PTV will in turn affect the dose calculated in the PTV. This is particularly relevant for the dose just inside the PTV, which will affect the PTV V95%. Therefore, investigation of the composition of the volume including the PTV plus a margin as a predictor of the change in PTV V95% should be explored.

Appendix A

Dissemination

A.1   Presentations and posters

Monte Carlo Simulation of Oesophageal Radiotherapy Plans, Johns, D.; presented at MCNEG (Monte Carlo Users Group), Cardiff, UK, June 2012

Comparison of Monte Carlo and Convolution Based Treatment Planning Calculations for Oesophageal Radiotherapy, Johns, D., Downes, P., Spezi, E. and Lewis, D. G.; presented at Third Meeting of the European Working Group on Monte Carlo Treatment Planning, Seville, Spain, May 2012

Distributed Monte Carlo Simulations of Oesophageal Treatment Plans, Johns, D., Downes, P., Spezi, E, Wells, P. and Lewis, D. G.; presented at 18th International Conference on the use of Computers in Radiation Therapy, London, UK, June 2016

Comparison of Treatment Planning Algorithms and Monte Carlo Simulations in Oesophageal Radiotherapy, Johns, D., Downes, P., Spezi, E and Lewis, D. G.; Accepted for the 36th Congress of the European Society of Therapeutic Radiation Oncology, Vienna, Austria, May 2017

A.2   Published abstracts

Johns, D., Downes, P., Spezi, E. and Lewis, D. G. (2013) 'EP-1188: Comparison of treatment planning system algorithms and Monte Carlo simulations for oesophageal radiotherapy', *Radiotherapy and Oncology*, 106, p. S449.

Appendix B

EDW profiles

As discussed in Section 3.2.5.4, in this work beam profiles were produced for 6 MV and 10 MV for three field sizes in varying size phantoms, for the four EDW wedge angles, 15°, 30°, 45°, and 60°. The results for 5 cm by 5 cm and 15 cm by 15 cm fields are presented below.

B.1   5 cm by 5 cm fields

Figure B.1 compares the measurement and MC simulation for a 5 cm by 5 cm, 6 MV beam with a 15° EDW wedge. For the profile in the X-direction, the agreement between simulation and measurement is within 1% for all points. For the Y-direction, some points in the beam 'shoulders' show a 1-2% discrepancy between simulation and measurement, while all other points agree to within 1%. The agreement of the shape of the beam shoulders and penumbra is visually very good.

The measurement and MC simulation for a 5 cm by 5 cm, 6 MV beam with a 30° EDW wedge are compared in Figure B.2. For the profile in the X-direction, the agreement between simulation and measurement is within 1%, aside from a few points in the shoulders of the beam; the agreement for these points is within 2%. For the Y-direction all points agree to within 2%.

Figure B.3 compares the measurement and MC simulation for a 5 cm by 5 cm,

6 MV beam with a 45° EDW wedge. For the profile in the X-direction, the agreement between simulation and measurement is within 1%, aside for a few points in the shoulders of the beam that are within 2% of measurement. For the Y-direction all other points agree to within 3%; although agreement is better away from the toe end on the wedge.

The measurement and MC simulation for a 5 cm by 5 cm, 6 MV beam with a 60° EDW wedge are compared in Figure B.4. For the profile in the X-direction, again agreement is within 1%, aside from a few points in the shoulders of the beam, which are within 2% of measurement. For the Y-direction the agreement away from the toe end of the wedge is within 2%, although differences in the toe of the wedge approach 4%.

Figure B.5 compares the measurement and MC simulation for a 5 cm by 5 cm, 10 MV beam with a 15° EDW wedge. The profiles in both the X-direction and Y-direction show agreement within 1%, aside from a few points in the shoulders of the beam. This behaviour is repeated in Figure B.6, which compares the measurement and MC simulation for a 5 cm by 5 cm, 10 MV beam with a 30° EDW wedge.

Figure B.7 compares the measurement and MC simulation for a 5 cm by 5 cm, 10 MV beam with a 45° EDW wedge. For the profile in the X-direction, the agreement between simulation and measurement is within 1%, aside for a few points in the shoulders of the beam that are within 2% of measurement. For the Y-direction all other agree to within 3%; although agreement is better away from the toe end on the wedge.

The measurement and MC simulation for a 5 cm by 5 cm, 10 MV beam with a 60° EDW wedge are compared in Figure B.8. For the profile in the X-direction, the agreement between simulation and measurement is within 1%. For the profile in the Y-direction, the agreement between simulation and measurement is within 3%, with the majority of points being within 2%.

B.2   15 cm by 15 cm fields

Figure B.9 compares the measurement and MC simulation for a 15 cm by 15 cm, 6 MV beam with a 15° EDW wedge. For the profile in the X-direction, the agreement between simulation and measurement is within 2%, aside from a few points in the shoulders of the beam; the agreement for these points is within 3%. For the Y-direction, the agreement away from the toe end of the wedge is within 2%, although differences in the toe of the wedge and the shoulders of the beam approach 4%.

The measurement and MC simulation for a 15 cm by 15 cm, 6 MV beam with a 30° EDW wedge are compared in Figure B.10. For the profile in the X-direction, the agreement between simulation and measurement is again within 2%, aside from a few points in the shoulders of the beam where the agreement is within 3%. For the Y-direction, the agreement away from the toe end of the wedge is within 2%, although differences in the toe of the wedge and the shoulders of the beam approach 5%.

Figure B.11 compares the measurement and MC simulation for a 15 cm by 15 cm, 6 MV beam with a 45° EDW wedge. As for previous 15 cm by 15 cm EDW fields, the agreement in the X-direction is within 2%, aside from a few points in the shoulders of the beam where the agreement is within 3%. For the Y-direction, the agreement away from the toe end of the wedge is within 2%. In the toe of the wedge, the agreement is the worst of all fields compared, with differences between simulation and measurement up to 6%.

The measurement and MC simulation for a 15 cm by 15 cm, 6 MV beam with a 60° EDW wedge are compared in Figure B.12. The profiles in the X-direction show the largest differences in the X-direction for all fields compared, with differences of up to 4%. In the Y-direction, the differences are with 6%, with differences in both the toe and heel ends of the wedge.

Figure B.13 compares the measurement and MC simulation for a 15 cm by 15 cm, 10 MV beam with a 15° EDW wedge. For the profile in the X-direction, the agreement between simulation and measurement is within 2%. For the Y-direction the agreement away from the toe end of the wedge is within 2%, although there are differences of up to 4% in the toe of the wedge.

The measurement and MC simulation for a 15 cm by 15 cm, 10 MV beam with a 30° EDW wedge are compared in Figure B.14. For the profile in the X-direction, the agreement between simulation and measurement is within 2%, aside from some points in the shoulders of the beams, which are within 3%. For the Y-direction the agreement away from the toe end of the wedge is within 2%, although there are differences of up to 5% in the toe of the wedge.

Figure B.15 compares the measurement and MC simulation for a 15 cm by 15 cm, 10 MV beam with a 45° EDW wedge. For the profile in the X-direction, the agreement between simulation and measurement is again within 2%, aside from some points in the shoulders of the beams, which are within 3%. For the Y-direction the agreement away from the toe end of the wedge is within 2%, although there are differences of up to 5% in the toe of the wedge.

The measurement and MC simulation for a 15 cm by 15 cm, 10 MV beam with a 60° EDW wedge are compared in Figure B.16. Once again, the profile in the X-direction shows that the agreement between simulation and measurement is within 2%, aside from some points in the shoulders of the beams, which are within 3%. For the Y-direction, agreement is within 3%, although the differences are of a similar magnitude in the heel and toe. This is in contrast to most fields where the heel end of the wedge is exhibits significantly better agreement than the toe end.

Overall, agreement is very good, with most fields agreeing to within 3%, aside from some larger discrepancies in the toe end of the wedge. The worst cases occur for the 15 cm by 15 cm fields, with 45° or 60° wedges, for both 6 MV and

10 MV. These fields exhibit discrepancies of 5-6%. The dose varies by a factor of 2-2.5 between the heel and toe of the wedge in these beams. Wedges are used to counteract dose gradients due to obliquities in the patient surface, or due to beam arrangements. A variation in dose for factor of 2-2.5 would be an extremely large variation to have arisen due to these effects, so it is not anticipated that such wedges will be used in clinical oesophageal plans.
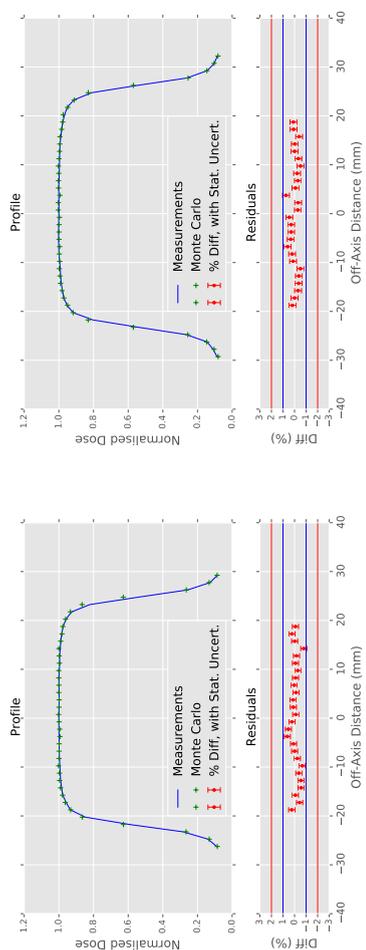
(a) 5 cm deep, X-direction
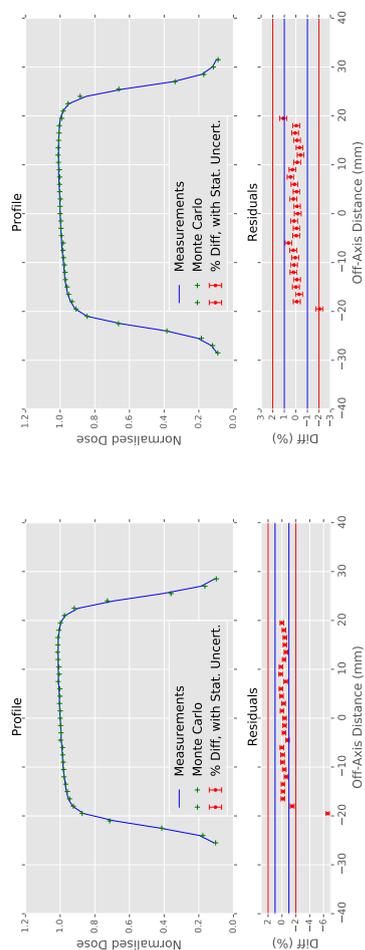
(b) 10 cm deep, X-direction
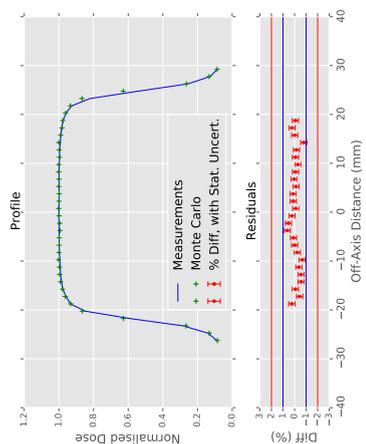
(c) 5 cm deep, Y-direction
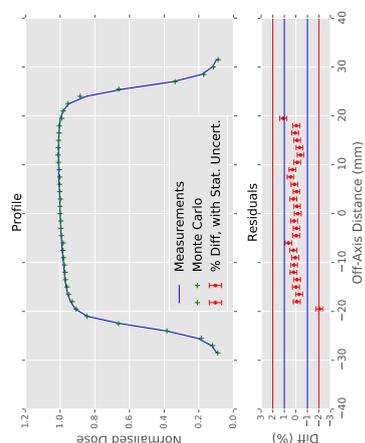
(d) 10 cm deep, Y-direction

Figure B.1: Comparison of MC calculated and commissioning profiles for 6 MV, 5 cm by 5 cm 15° EDW field.
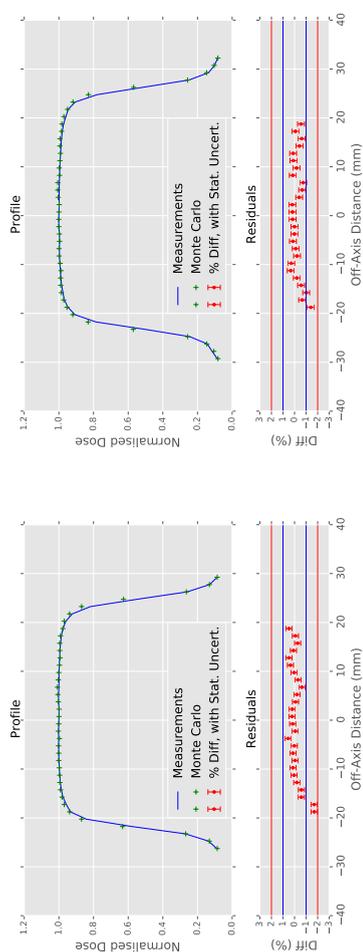
(a) 5 cm deep, X-direction
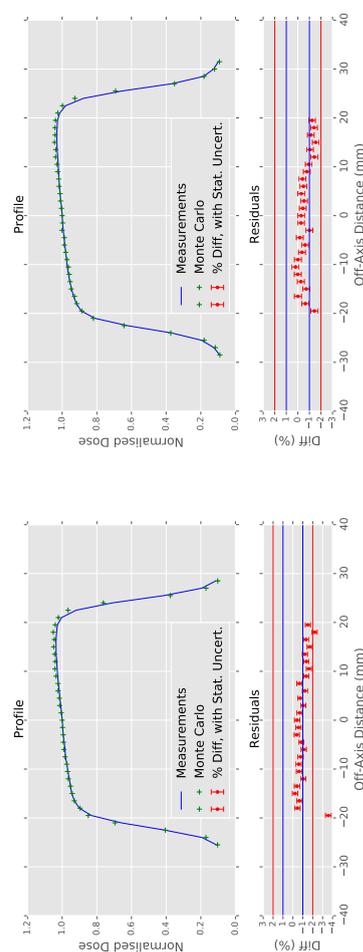
(b) 10 cm deep, X-direction
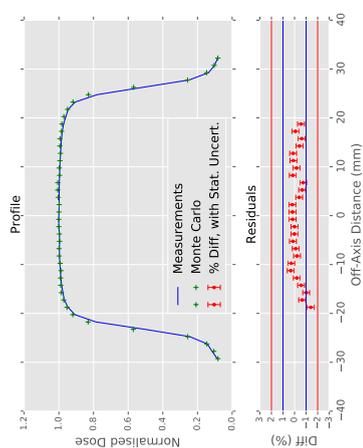
(c) 5 cm deep, Y-direction
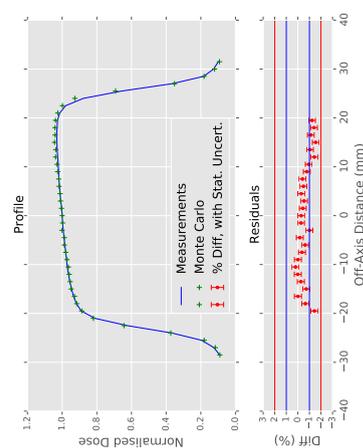
(d) 10 cm deep, Y-direction

Figure B.2: Comparison of MC calculated and commissioning profiles for 6 MV, 5 cm by 5 cm 30° EDW field.
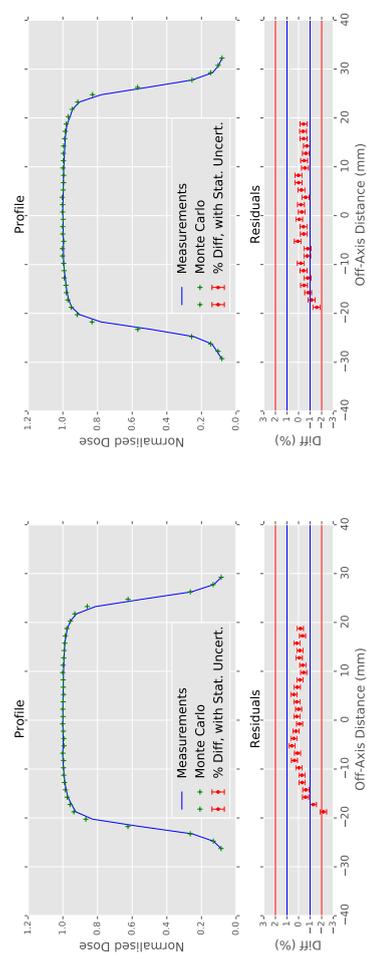
(a) 5 cm deep, X-direction
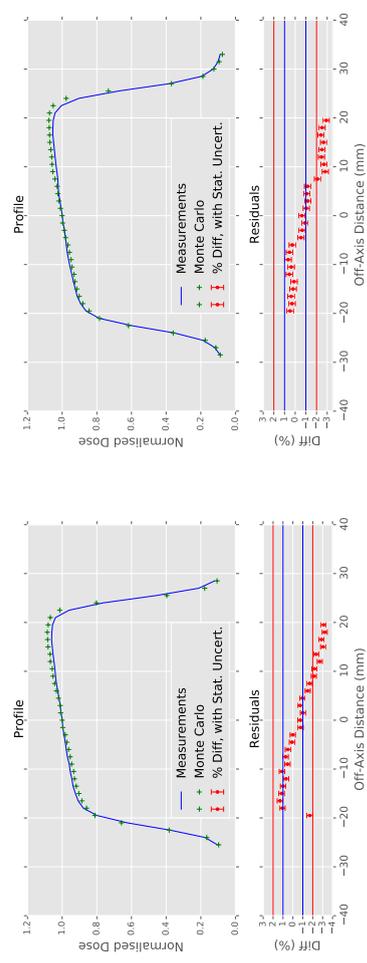
(b) 10 cm deep, X-direction
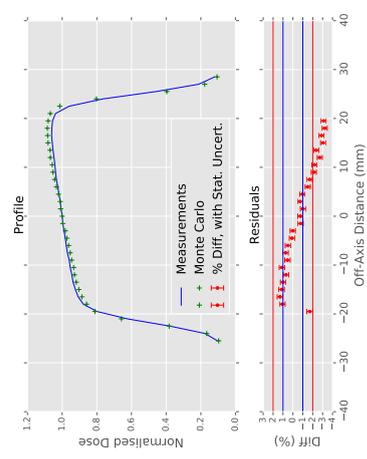
(c) 5 cm deep, Y-direction
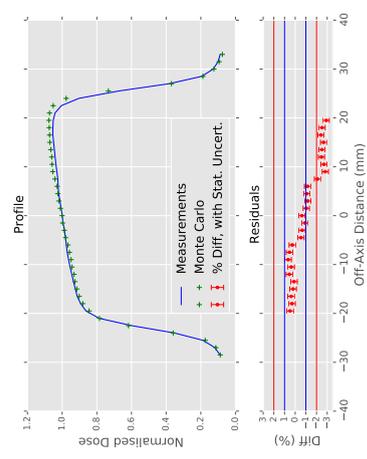
(d) 10 cm deep, Y-direction

Figure B.3: Comparison of MC calculated and commissioning profiles for 6 MV, 5 cm by 5 cm 45° EDW field.
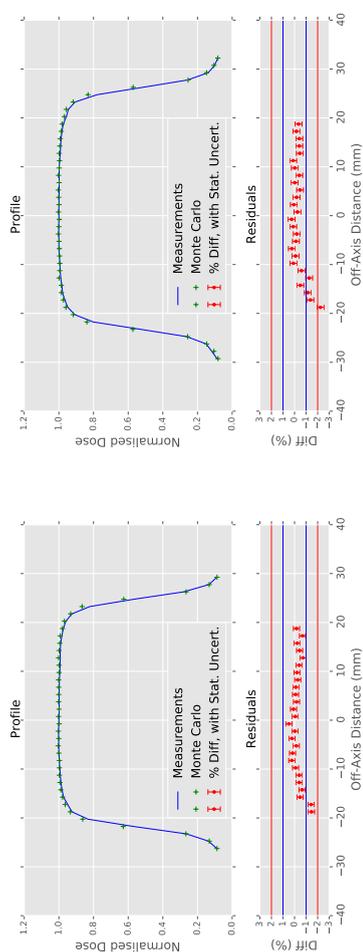
(a) 5 cm deep, X-direction
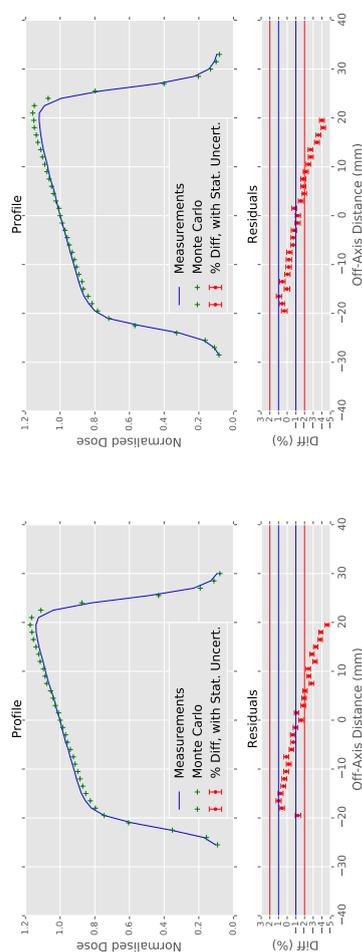
(b) 10 cm deep, X-direction
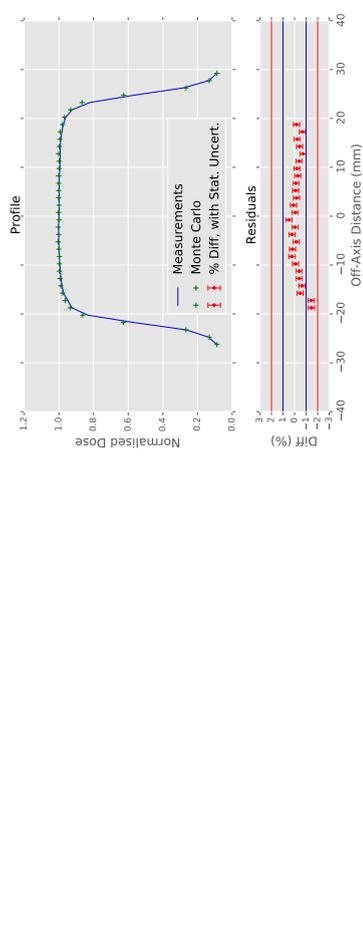
(c) 5 cm deep, Y-direction
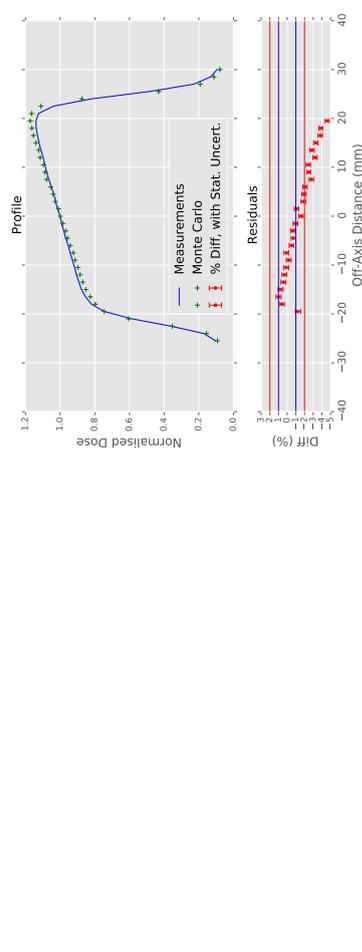
(d) 10 cm deep, Y-direction

Figure B.4: Comparison of MC calculated and commissioning profiles for 6 MV, 5 cm by 5 cm 60° EDW field.
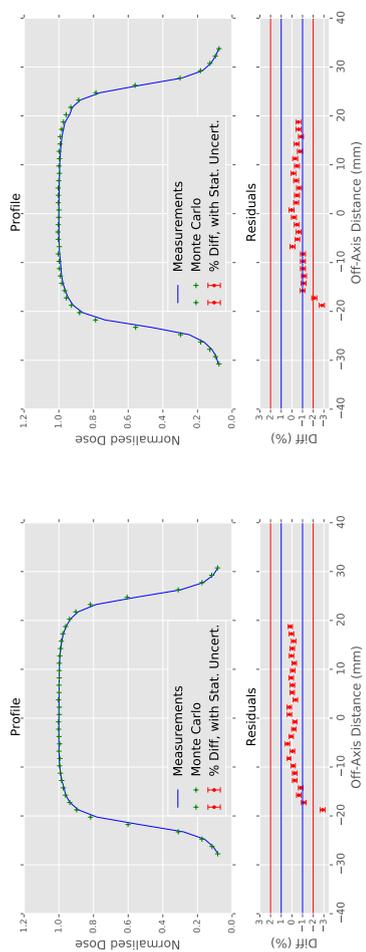
(a) 5 cm deep, X-direction
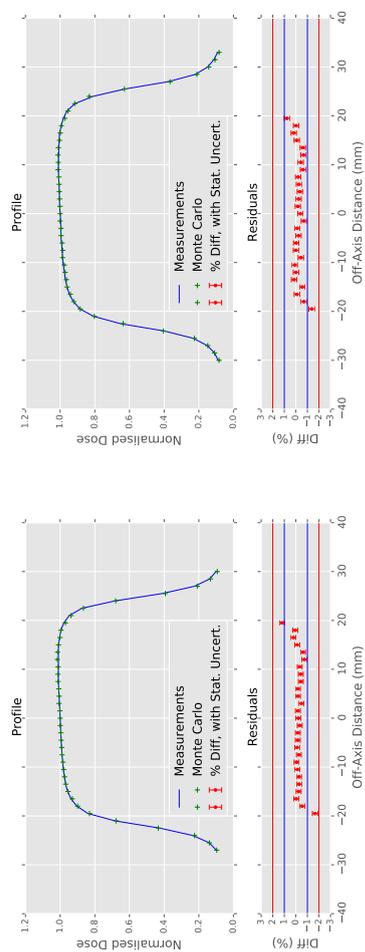
(b) 10 cm deep, X-direction
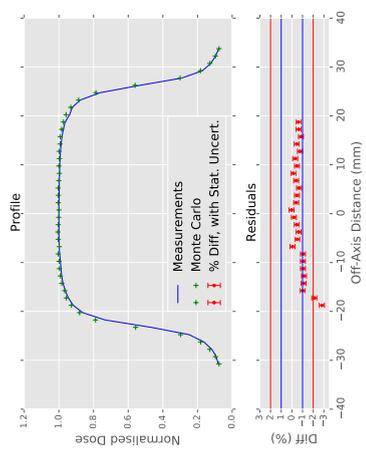
(c) 5 cm deep, Y-direction
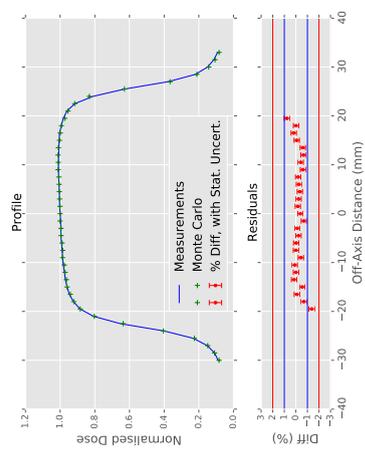
(d) 10 cm deep, Y-direction

Figure B.5: Comparison of MC calculated and commissioning profiles for 10 MV, 5 cm by 5 cm 15° EDW field.
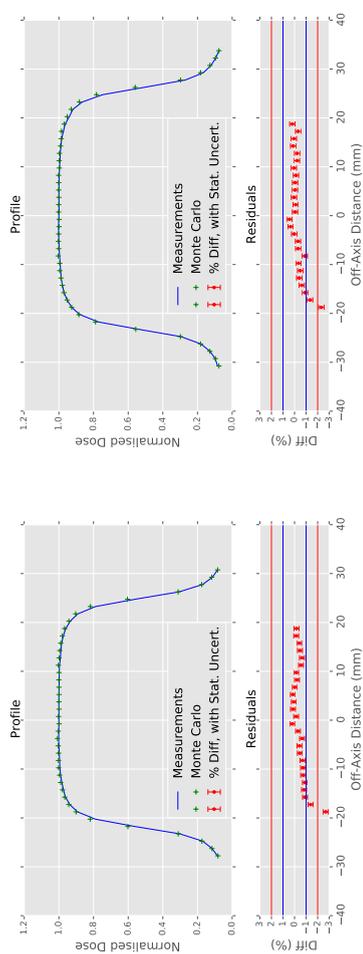
(a) 5 cm deep, X-direction
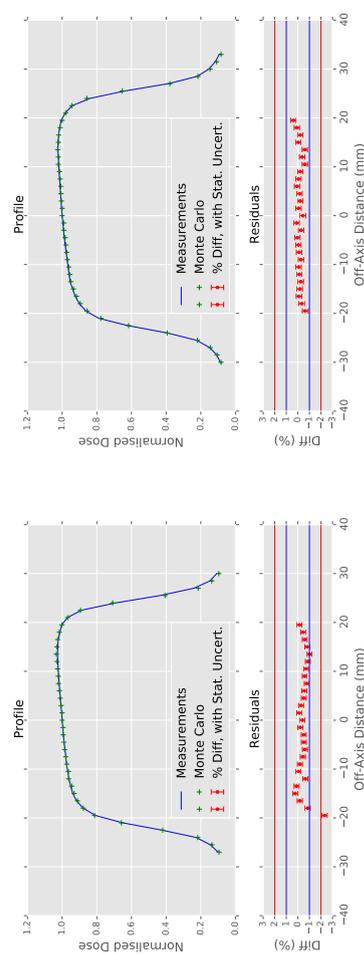
(b) 10 cm deep, X-direction
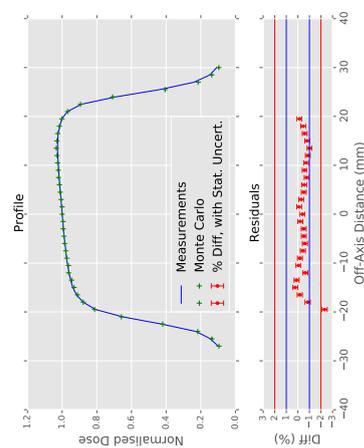
(c) 5 cm deep, Y-direction
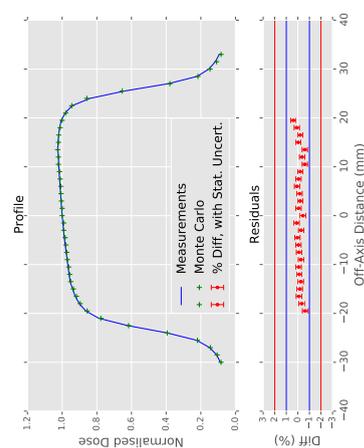
(d) 10 cm deep, Y-direction

Figure B.6: Comparison of MC calculated and commissioning profiles for 10 MV, 5 cm by 5 cm 30° EDW field.
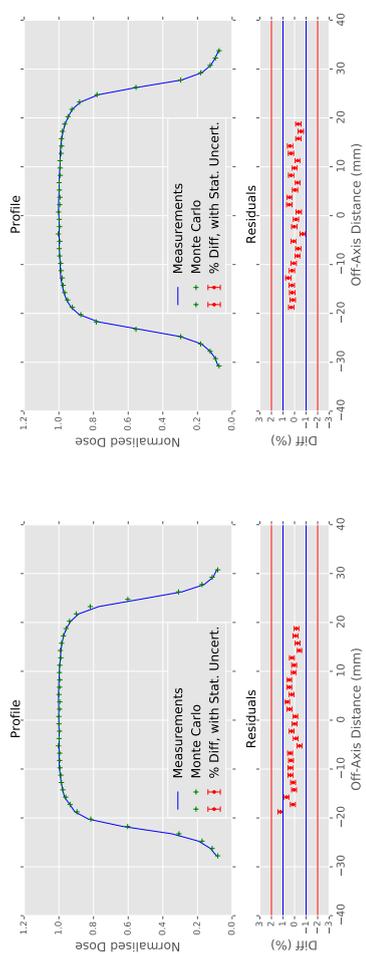
(a) 5 cm deep, X-direction
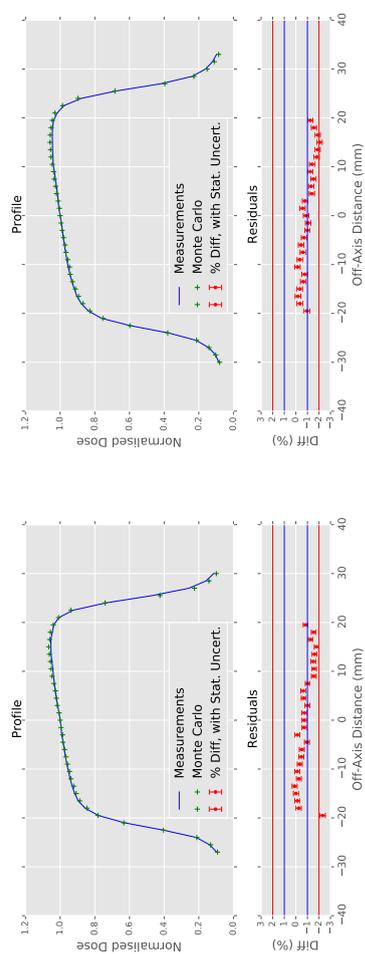
(b) 10 cm deep, X-direction
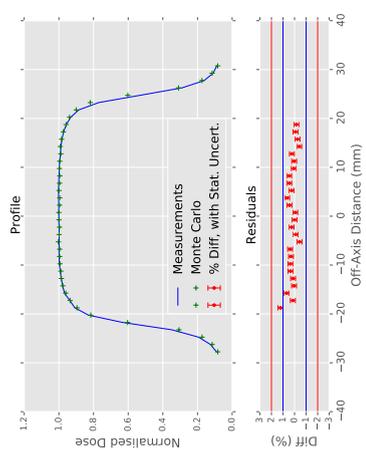
(c) 5 cm deep, Y-direction
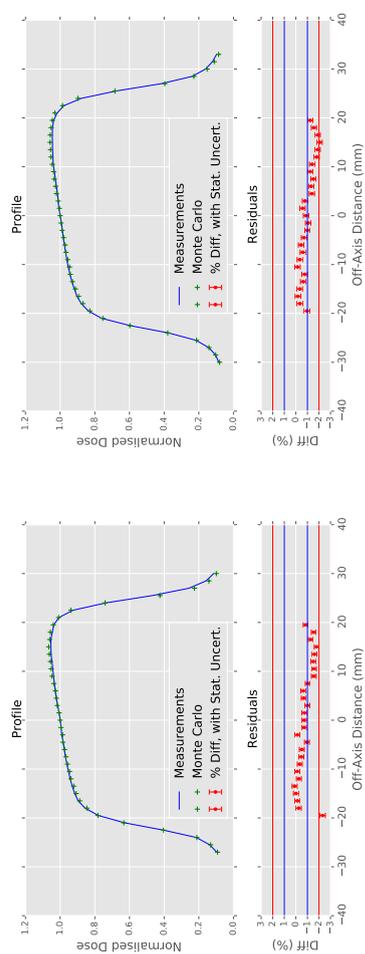
(d) 10 cm deep, Y-direction

Figure B.7: Comparison of MC calculated and commissioning profiles for 10 MV, 5 cm by 5 cm 45° EDW field.
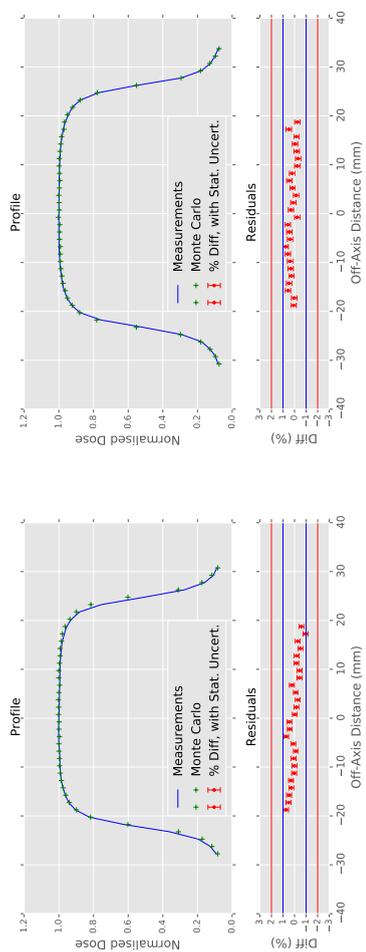
(a) 5 cm deep, X-direction
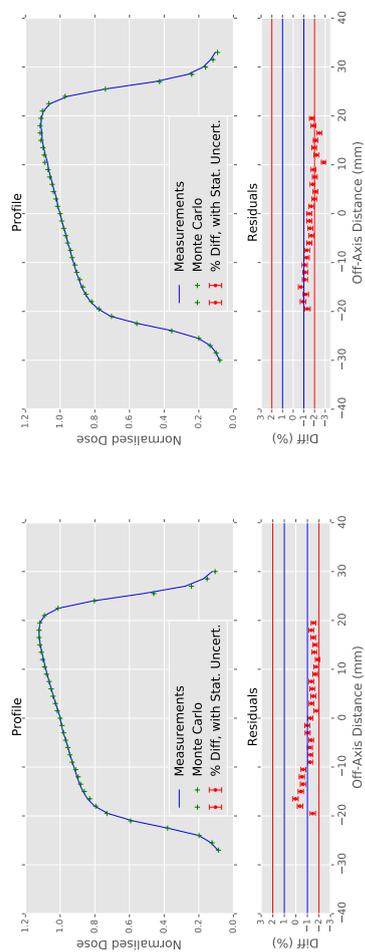
(b) 10 cm deep, X-direction
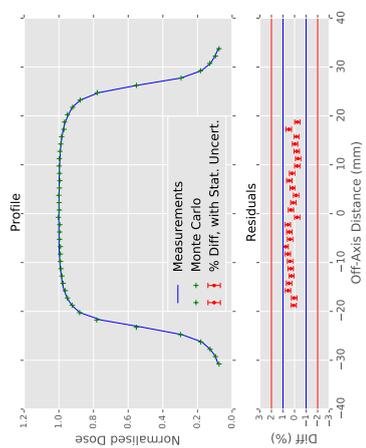
(c) 5 cm deep, Y-direction
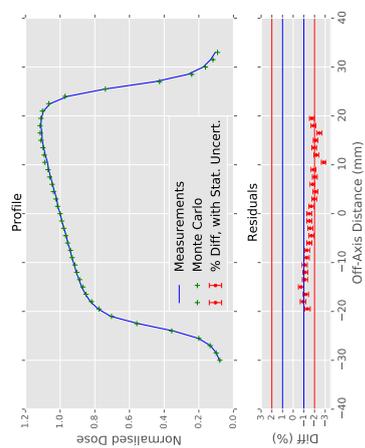
(d) 10 cm deep, Y-direction

Figure B.8: Comparison of MC calculated and commissioning profiles for 10 MV, 5 cm by 5 cm 60° EDW field.
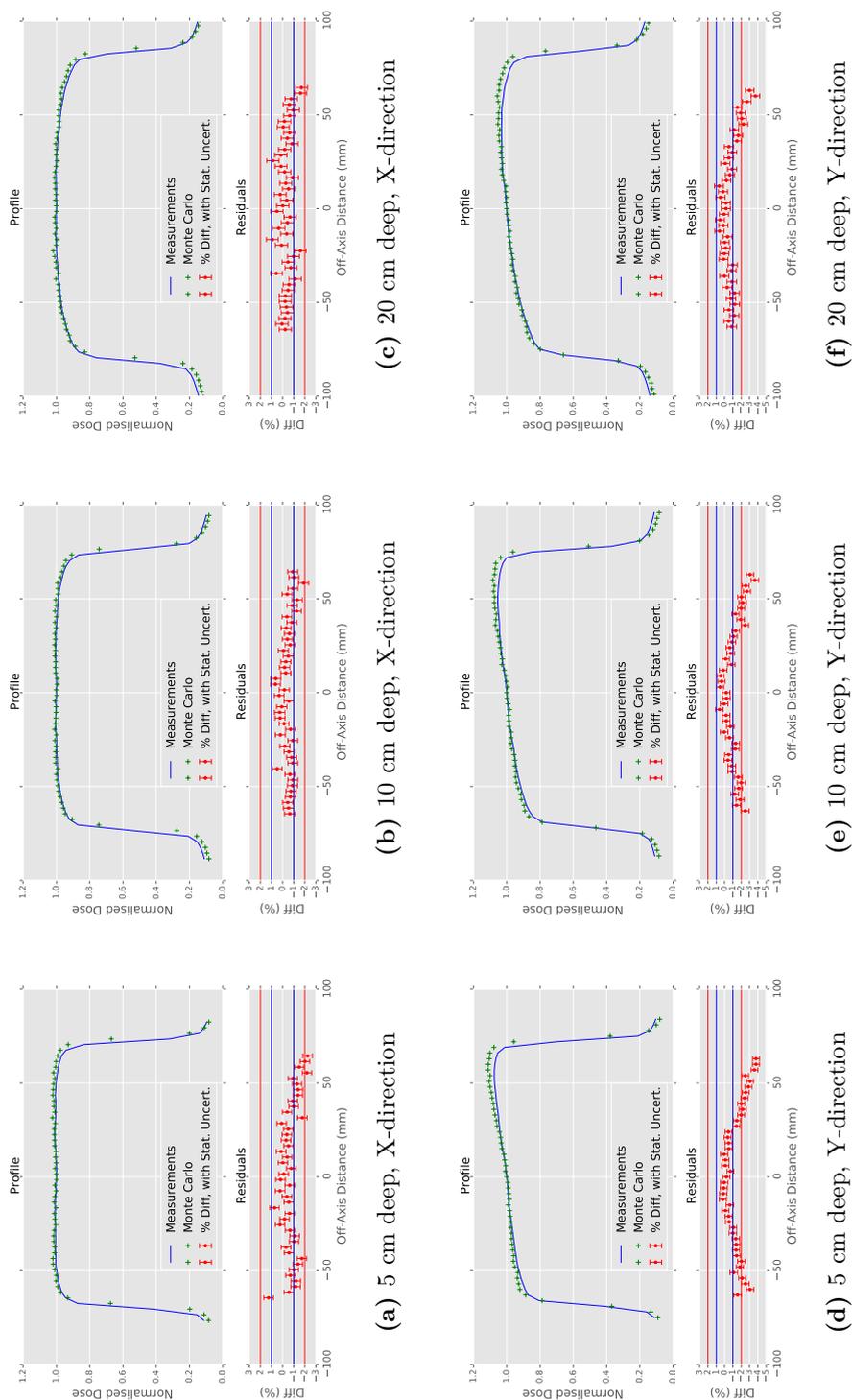
Figure B.9: Comparison of MC calculated and commissioning profiles for 6 MV, 15 cm by 15 cm 15° EDW field.

Figure B.10: Comparison of MC calculated and commissioning profiles for 6 MV, 15 cm by 15 cm 30° EDW field.
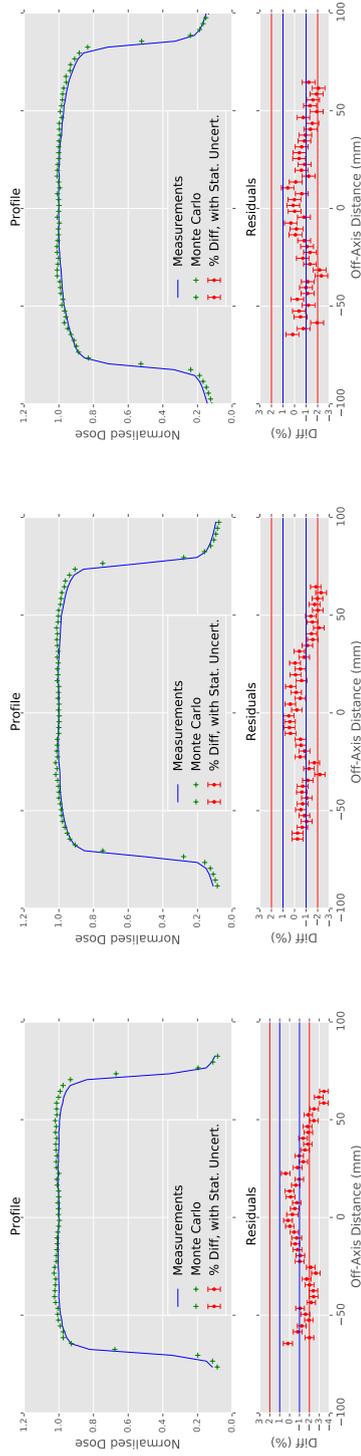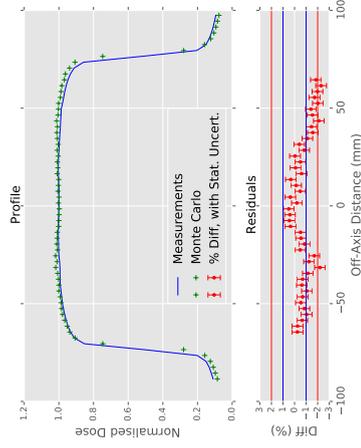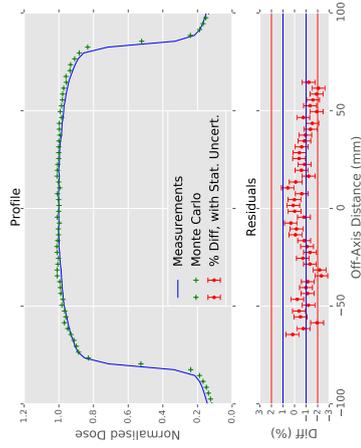
(a) 5 cm deep, X-direction

(b) 10 cm deep, X-direction

(c) 20 cm deep, X-direction
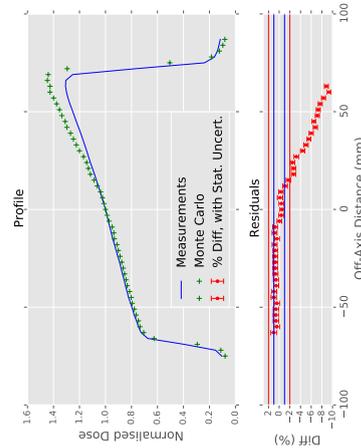
(d) 5 cm deep, Y-direction
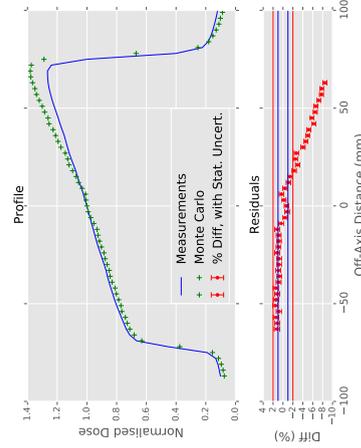
(e) 10 cm deep, Y-direction
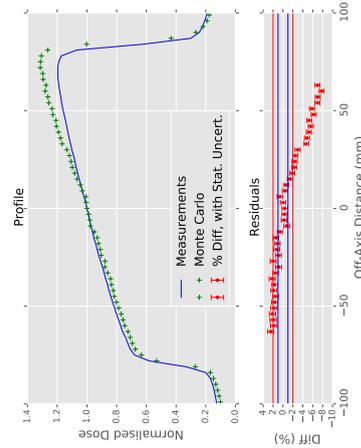
(f) 20 cm deep, Y-direction

Figure B.11: Comparison of MC calculated and commissioning profiles for 6 MV, 15 cm by 15 cm 45° EDW field.
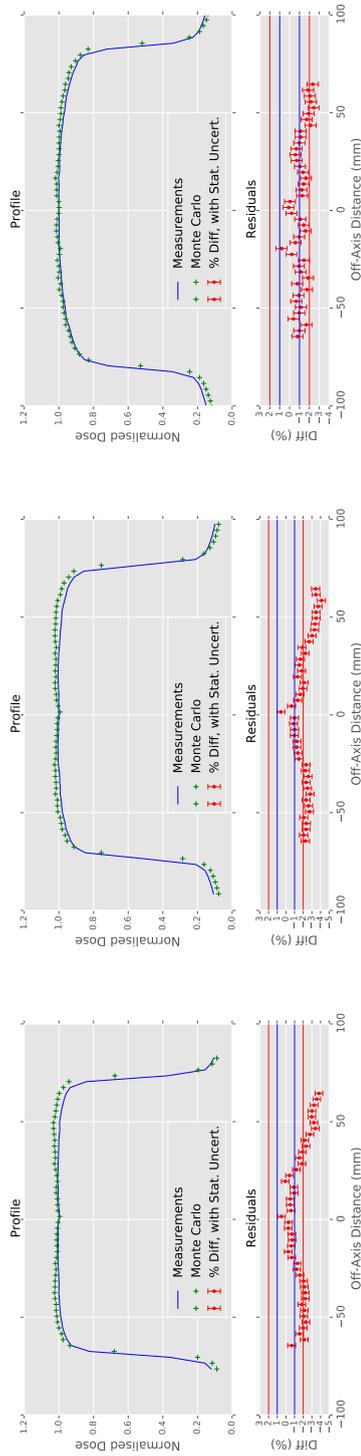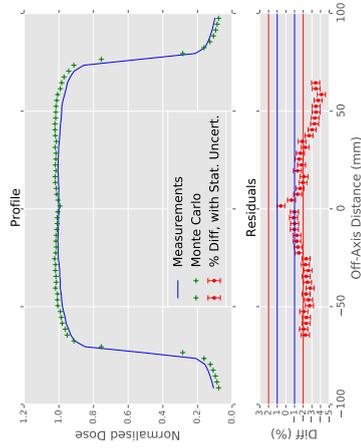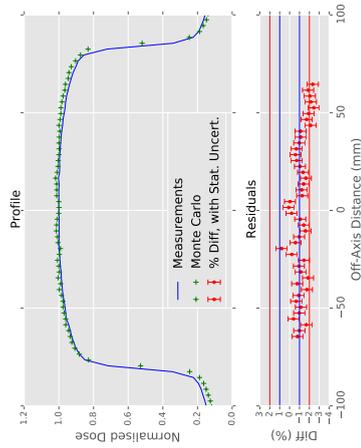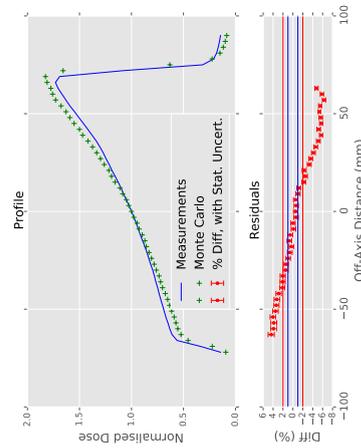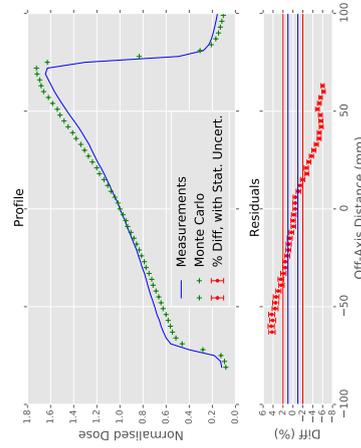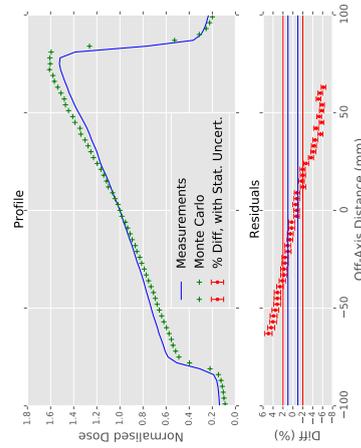
Figure B.12: Comparison of MC calculated and commissioning profiles for 6 MV, 15 cm by 15 cm 60° EDW field.
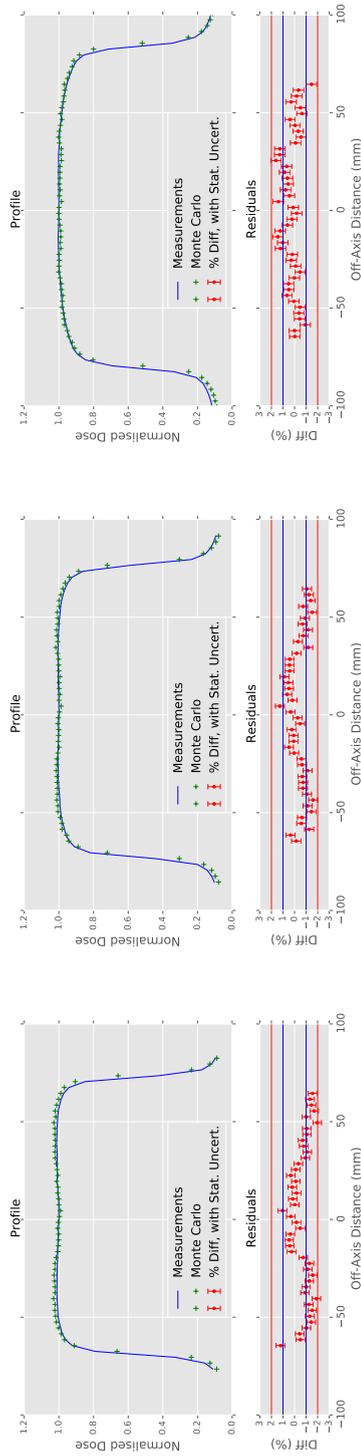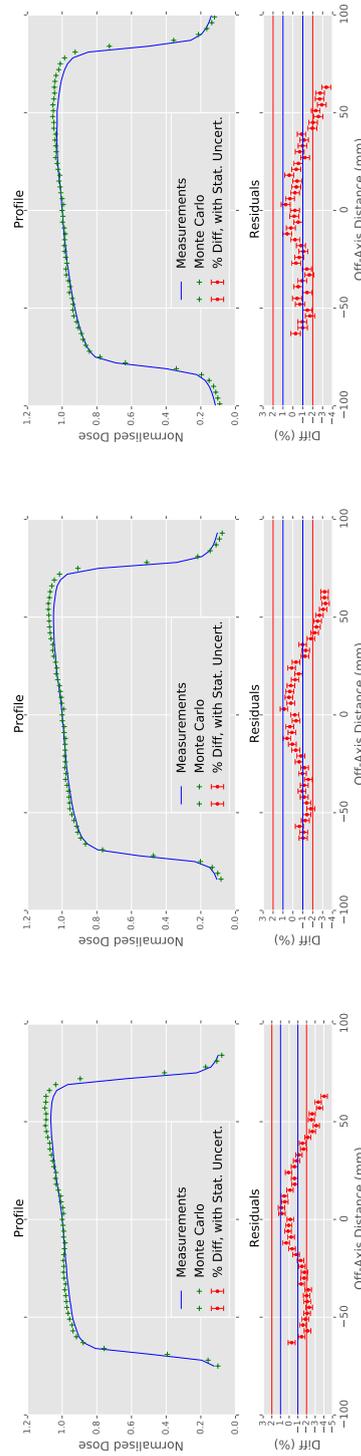
Figure B.13: Comparison of MC calculated and commissioning profiles for 10 MV, 15 cm by 15 cm 15° EDW field.
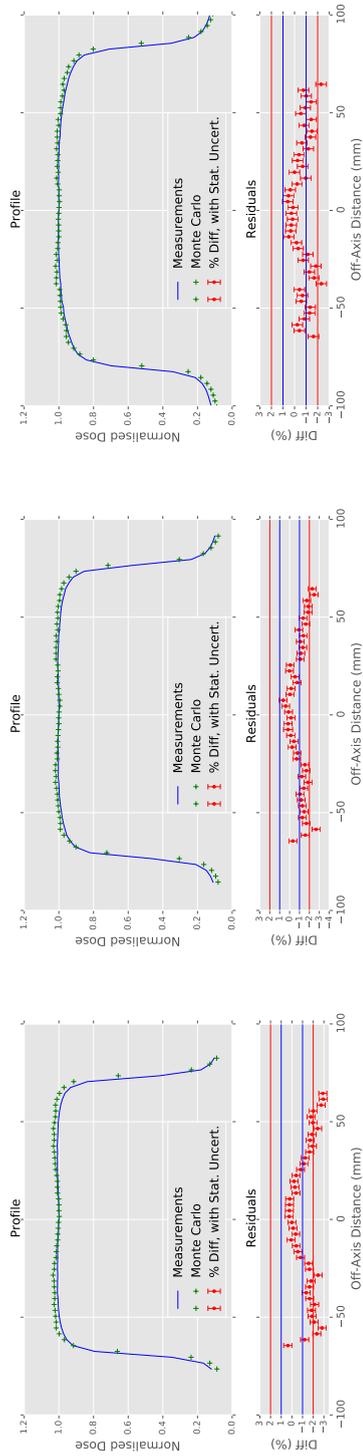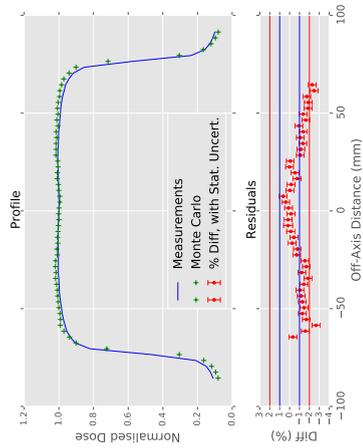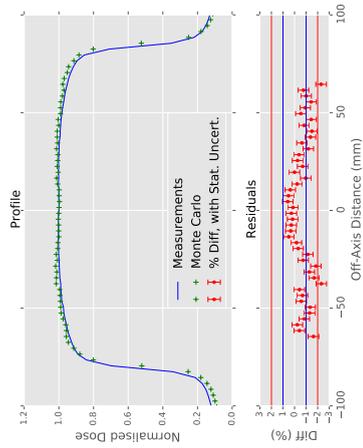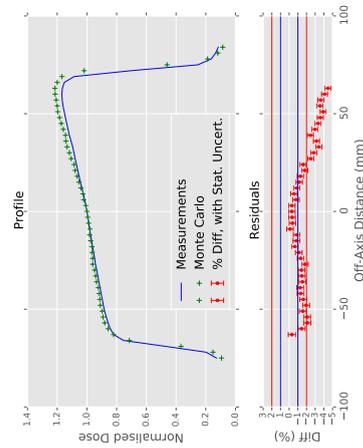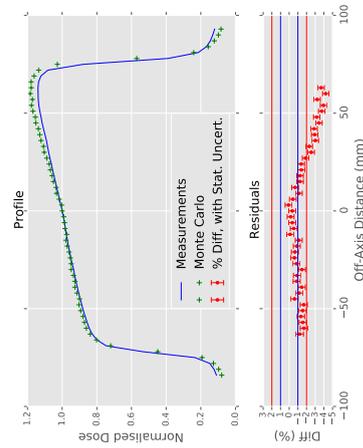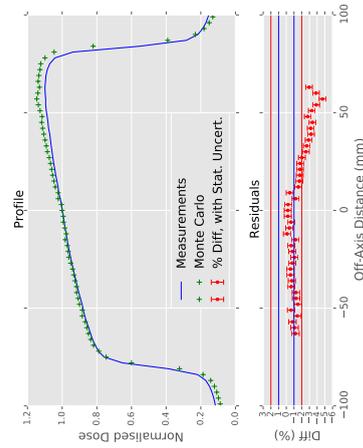
Figure B.14: Comparison of MC calculated and commissioning profiles for 10 MV, 15 cm by 15 cm 30° EDW field.
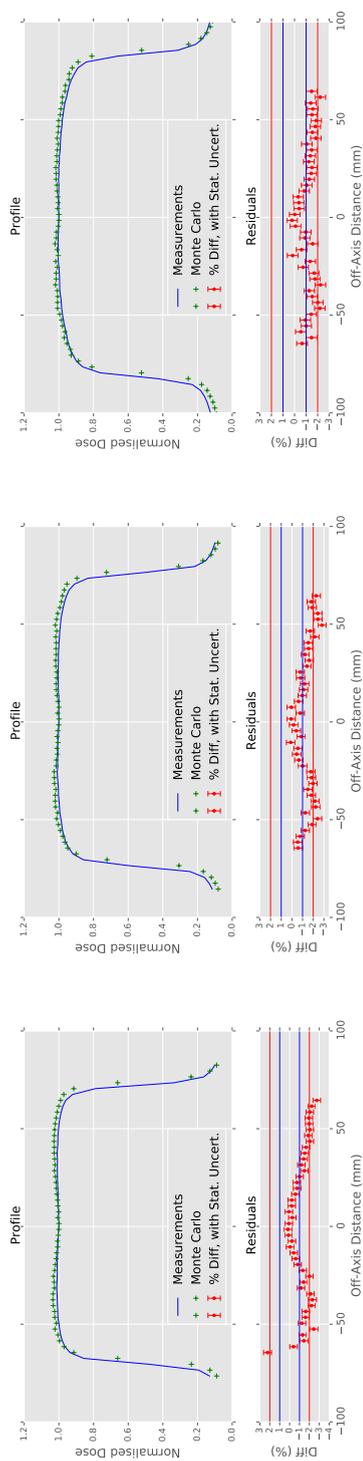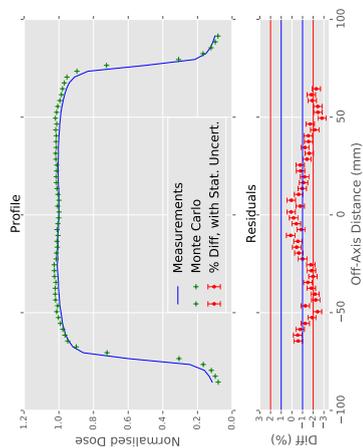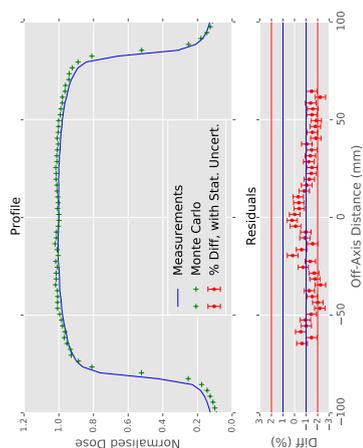
(a) 5 cm deep, X-direction

(b) 10 cm deep, X-direction

(c) 20 cm deep, X-direction
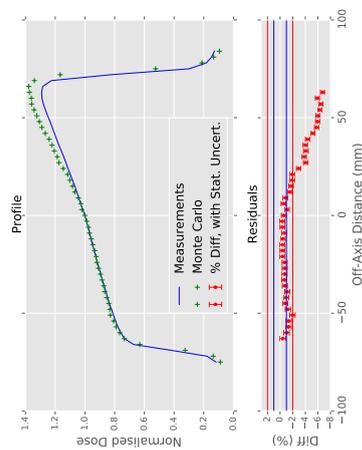
(d) 5 cm deep, Y-direction
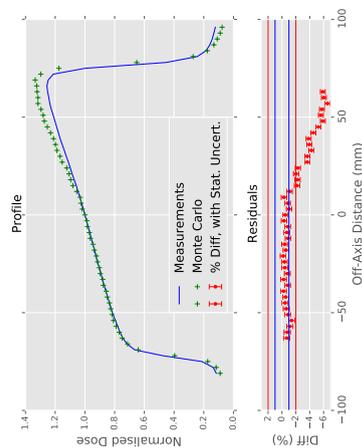
(e) 10 cm deep, Y-direction
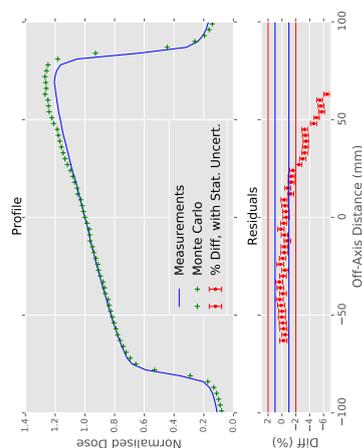
(f) 20 cm deep, Y-direction

Figure B.15: Comparison of MC calculated and commissioning profiles for 10 MV, 15 cm by 15 cm 45° EDW field.
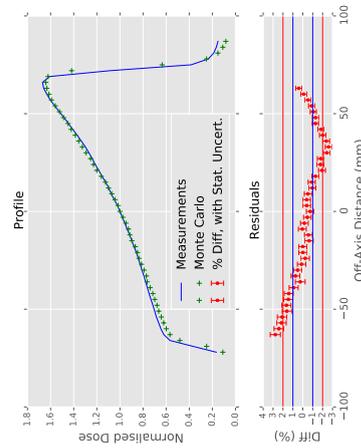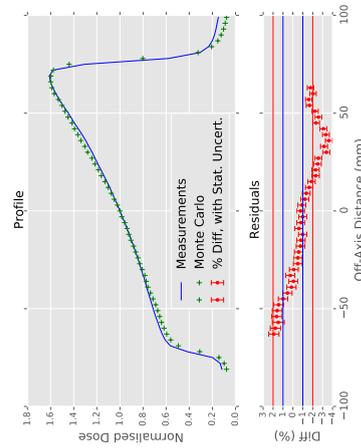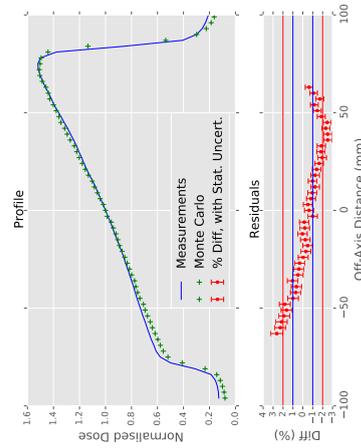
Figure B.16: Comparison of MC calculated and commissioning profiles for 10 MV, 15 cm by 15 cm 60° EDW field.

Glossary

American Association of Anaesthesiologists (ASA) Association of physicians organized to maintain and raise the standards of the medical practice of anaesthesiology. Produced the ASA Physical Status Classification System, a system for describing the fitness of a patient prior to surgery. 172

Apoptosis Programmed cell death, the mechanism by which damaged cells are killed. 2

Bremsstrahlung Braking Radiation, photons produced by the deceleration of a charged particle. 17, 36, 37, 39, 48, 55, 57, 104

carbon ion Form of particle therapy, in which beams of carbon nuclei are used to irradiate malignant tumours. The deposition of energy in matter for particles is described by the Bragg curve, which has a pronounced peak, known as the Bragg peak, just before the particle comes to rest. 10

Carcinogenesis The formation of cancer, in which normal cells transform into cancerous cells. 10

Cell Cycle Series of changes in a cell in which the DNA is duplicated, and the cell divides into two. 10–12

Central Processing Unit (CPU) The main piece of circuitry in a computer that performs the instructions provided by computer programs. Often known as the 'chip'. 50, 64, 71, 81

Charged Particle Equilibrium (CPE) For radiation beam passing through a material, the situation where the energy imparted to the electrons by the radiation beam at a point is balanced by the energy deposited by electrons

in the medium. When charged particle equilibrium is established, absorbed dose is equal to collision Kinetic Energy Released per unit Mass. 38, 44, 116

Chemo-radiotherapy (CRT) Combination of chemotherapy and radiotherapy. 7, 8

Clinical Target Volume (CTV) Gross Tumour Volume plus a margin for subclinical spread of disease. 79, 139

Collapsed Cone Algorithm (CC) Type B Dose Calculation algorithm used in Oncentra MasterPlan. 15, 16, 24, 25, 28, 44, 46, 50–52, 72–75, 188

Collapsed Cone Enhanced (CCE) Enhanced version of the Collapsed Cone Algorithm, which is faster than standard due to efficiencies in the calculation, while using a higher resolution head scatter grid. Throughout this work, the Enhanced version of the Collapsed Cone Algorithm has been used. iv, xvi–xviii, xxi, 51–53, 76, 77, 86, 113, 140–154, 156–158, 160, 169, 173–190, 192, 193, 196–198

Component Module (CM) A 'building block' used to create a model of the linear accelerator for the BEAMnrc Monte Carlo code. 57, 58, 62, 100

Computed Tomography (CT) Form of x-ray imaging that produces a series of cross-sectional images through the body, in which regions of high electron density appear white, and regions of low electron density appear darker . xiv, xvii, xviii, 8, 14, 15, 22, 48, 61, 67, 78, 79, 81, 88–91, 93–95, 99, 106–108, 110–114, 133, 139, 149, 156–158, 192, 193, 197, 199, 200

Condensed History (CH) In Monte Carlo, an approach to transporting lower energy electrons in which multiple scattering events are modelled as a single event, to save computation time. 54, 55, 57

Conformal Radiotherapy (CFRT) Type of radiotherapy in which radiation beams are shaped to match the shape of the target, usually by Multi-Leaf Collimators. 14, 16, 24, 25, 80, 139, 198, 201

deoxyribonucleic acid (DNA) Molecule that holds the genetic instructions necessary for the growth, development, function and reproduction of liv-

ing organisms. Has a double-helical structure, in which two strands wind around each other, like a twisted ladder. 9, 11

Dice Similarity Coefficient (Dice-CI) Conformance Index for measuring how one region of interest conforms to another, calculated as the ratio of twice the intersection between the volumes to the sum of the volumes. 162, 163, 166, 168

Digital Imaging and COmmunications in Medicine (DICOM) A standard for transferring, viewing and storing digital images for medical use. 22, 23, 69, 83–85, 88, 108, 113, 114

Digital Imaging and COmmunications in Medicine, Radiotherapy Supplement (DICOM-RT)

Extension to DICOM standard for Radiotherapy objects, such as structures, radiotherapy plans and dose matrices. 22, 23, 64, 65, 67, 69, 84–88, 108, 111–113, 115

Discordance Index (DI-CI) Conformance Index for measuring how little one region of interest conforms to another, calculated as the ratio of the intersection of the volumes to the volume of the reference object, subtracted from 1. As the overlap volume increases, this index decreases. 163, 166, 168, 177, 190

Dose Volume Histogram (DVH) Graphical tool used in Radiotherapy to describe amount of dose being delivered to a volume. Histograms are usually plotted cumulatively, so a point on the graph shows the volume receiving up to or including the specified dose. iv, xvii, xviii, 15, 59, 66–68, 76, 80, 81, 87, 113, 140, 147, 149–152, 156, 160, 161, 165, 166, 168, 181, 182, 185, 186, 188–190, 197, 198

Double Strand Break (DSB) In radiobiology, a double strand break is a break in both of the two helices in the structure of the DNA molecule, which is much harder to repair than a single strand break. 9

electron-gamma-shower (EGS) A general purpose Monte Carlo code, used extensively in Medical Physics, as well as other fields. 56–58

Electronic Portal Imaging Device (EPID) Imaging device used on linear accelerators, to produce an image from radiation passing through the patient. 65, 69

Energy fluence Amount of radiant energy per unit area, at a given position, traversing a plane perpendicular to the beam. 46–49, 51, 52

Enhanced Dynamic Wedge (EDW) Method of producing a wedged isodose distribution from certain Varian linacs, in which a collimator is driven across the radiation beam. xv, xvi, xviii, xix, 21, 78, 100, 101, 103, 127–136, 156, 196, 204–207, 209–224

External Beam Radiotherapy (EBRT) Type of radiotherapy in which radiation beams originating outside of the body are incident on the body. 4, 16, 28, 29, 32, 34

Field Size Factor (FSF) Absorbed dose at depth of maximum dose on the beam axis for a 10 cm × 10 cm field, divided by dose at the same point for a given field size. Inverse of the Output Factor. 47, 103

Floating Point Operations Per Second (FLOPS) Metric used for comparing performance of supercomputers. 62

Focus to Source Distance (FSD) Distance from a reference plan in the linear accelerator head containing the focus of the electron beam, to the surface of the patient or phantom. 102

Geometric Miss Index (GMI) Measure describing how little two volumes overlap; GMI=0 implies one volume is completely enclosed in the other, GMI=1 implies no overlap between volumes. 164, 166, 168

Graphical User Interface (GUI) A type of user interface that allows users to interact with electronic devices, particularly using a mouse to click on icons, buttons and menus, as opposed to typing commands into a text based interface. 66, 67

Graphics Processing Unit (GPU) Computing component designed to accelerate the creation and manipulation of images, and performing similar calculations. May be dedicated to calculations, and not output an image to a display device. 50, 70–72, 77

Gross Tumour Volume (GTV) Volume identified as disease, normally by 3D imaging. 79, 106, 139, 143, 154

High Performance Computing (HPC) A Computing paradigm focussing on executing tightly coupled computing tasks that require high number of Floating Point Operations per second. 62–64, 77

High Throughput Computing (HTC) A Computing paradigm focussing on efficient execution of loosely coupled tasks. 62, 63, 77, 81

Hounsfield Units (HU) Value representing electron density in CT scan, relative to that of water. xiv, 88–91, 93–95, 108, 112, 199

Hydro-peroxy Radical A molecule that has unpaired valence electrons, and a Oxygen-Oxygen bond, such as HO_2^-. 10

Hypoxic State in which a region of the body is deprived of Oxygen. 10

Intensity Modulated Radiotherapy (IMRT) Type of Radiotherapy in which Multi-Leaf Collimators move across the radiation beam during irradiation, to produce dose distributions that cannot be achieved with Conformal Radiotherapy. 46, 52, 68, 76, 139, 201

International Commission on Radiation Units and Measurement (ICRU) Body that establishes international standards for radiation units and measurement. 57, 99

Inter-Quartile Range (IQR) Measure of variability of data, based on dividing rank-ordered data into four groups of equal size, and finding the difference between the highest value in the third quartile and the highest value in the first quartile . 178–180

Isodose Line of equal dose, tool used to display radiotherapy dose overlaid on a Computed Tomography scan, or similar. xvii, 15, 20, 42, 59, 66, 149, 150, 156–158, 183, 190

Jaccard Concordance Index (Jaccard-CI) Conformance Index for measuring how one region of interest conforms to another, calculated as the ratio of intersection between the volumes to the union of the volumes. xvii, 162, 163, 166, 168, 179, 180, 182, 190, 197

Job Submission Description Language (JSDL) An XML file describing the programs to be executed, the parameters passed to the executables, the resources they can be run on etc. in the RTGrid. 83, 85

kerma Kinetic energy released per unit mass. 38, 45

kilo-electron volts (keV) A unit of energy, usually used for to the energy of a charged particle. An Electron will gain 1 eV, equal to 1.602×10^{-19} joule, when accelerated through a potential of 1 volt. 1 keV is 10^3 eV. 30, 35, 57, 60

Linear Accelerator (linac) Treatment Machine used to deliver radiotherapy. 15, 16, 18, 20–22, 24, 37, 46, 55, 57–62, 65, 66, 71, 77, 78, 87, 99–101, 103–105, 113, 114, 139

Linear Energy Transfer (LET) Quantity describing the energy transferred by an ionizing particle to the material traversed, per unit distance. 10

Mean Distance to Conformity (MDC) Method of comparing two volumes, equal to the average distance that a point on one surface must be moved to achieve perfect conformance with the other surface. 164–166, 168, 197

Medical Research Council (MRC) One of seven research councils in the United Kingdom, the Medical Research Council co-ordinates and funds research aimed at improving human health. 7

mega-electron volts (MeV) A unit of energy, usually used for to the energy of a charged particle. An Electron will gain 1 eV, equal to 1.602×10^{-19} joule, when accelerated through a potential of 1 volt. 1 MeV is 10^6 eV. xiii, 29, 30, 32, 34, 57, 61, 73

Monitor Units (MU) A linac is calibrated such that a known signal from the Monitor Chamber gives a known dose under reference conditions. Typically 100 MU will deliver 1 Gy at a reference point, under reference conditions. 21, 22, 41, 46, 66, 75, 101–103, 112

Monte Carlo (MC) Numerical Technique involving use of random numbers. In this work, refers to calculation of radiotherapy dose distributions by simulating the passage of photons and electrons through materials, where

random number generation is used to account of probabilistic nature of interaction between radiation and matter. iv, xv, xvi, xviii–xxi, 16, 24, 25, 28, 53, 54, 56–64, 66–78, 81–83, 86–88, 100, 104, 105, 108, 111–115, 117–132, 134–138, 143, 150, 156, 160, 173, 175, 177, 179, 182–184, 187–189, 191, 196–201, 204–207, 209–224

Monte Carlo, Dose To Material (MC D_{Mat}) Absorbed dose calculated by Monte Carlo simulation, calculated using the interaction cross-sections for the material in the voxel or region of interest . xvi–xviii, xxi, 86, 113, 140–154, 157, 158, 169, 173, 174, 176–186, 189, 190, 192, 193, 197, 198

Monte Carlo, Dose To Water (MC D_{Wtr}) Absorbed dose calculated by Monte Carlo simulation, scaled to give the dose as if the voxel were comprised of water. xvi, xvii, 86, 113, 140–148, 178–184, 190, 197

Multi-Disciplinary Team (MDT) A group of health care workers who are members of different disciplines. 8

Multi-Leaf Collimator (MLC) Series of thin collimators, which can be used to shape the radiation beam. xii, 19, 21, 22, 27, 46, 47, 58, 62, 65, 67, 76, 103, 139

National Institute of Standards and Technology (NIST) Measurement standards laboratory in the United States of America. 57

Off Axis Factor (OAF) Ratio of the absorbed dose at a point away from the beam axis, to the absorbed dose at a point on the beam axis at the same depth. Depends on the energy of the beam, depth, and the field size. 41

Organ at Risk (OAR) Organs to which the radiation dose must be taken into account during the treatment planning process. 14, 15, 22, 25, 75, 79, 80, 110, 149, 188, 189, 199

Output Factor (OPF) Ratio of absorbed dose at the depth of maximum dose on the beam axis for a given field size, to the absorbed dose at the same depth on the beam axis for a 10 cm by 10 cm field. 103

Oxygen Enhancement Ratio (OER) The ratio of doses needed to produce a given effect in the presence and absence of oxygen. 10

Pencil Beam Algorithm (PB) Type A Dose Calculation algorithm used in Oncentra MasterPlan. 15, 16, 24, 25, 28, 44, 46, 49–53, 72–75, 188

Pencil Beam Enhanced Algorithm (PBE) Enhanced version of the Pencil Beam Algorithm, which is faster than standard due to efficiencies in the calculation, while using a higher resolution head scatter grid. Throughout this work, the Enhanced version of the Pencil Beam Algorithm has been used. iv, xvi, xvii, 50, 52, 76, 77, 86, 113, 140–148, 178–184, 188, 190, 196–198

Percentage Depth Dose (PDD) The PDD at a point in a water phantom is the absorbed dose at that point divided by the absorbed dose at the depth of maximum dose for the same field size, expressed as a percentage. PDD applies only to the points along the beam axis. xv, 37–39, 41, 47, 114–120, 127

Personally Identifiable Information (PII) Any data that could potentially identify a specific individual. 68

peta-electron volts (PeV) A unit of energy, usually used for to the energy of a charged particle. An Electron will gain 1 eV, equal to 1.602×10^{-19} joule, when accelerated through a potential of 1 volt. 1 PeV is 10^{15} eV. 60

Photo-electric effect (PE) Process by which a photon interacts will an electron bound to an atom, and the electron is ejected. 29, 31, 38

Pixel A portmanteau of Picture Element; one value on a regular grid in two dimensional space; may be smallest addressable element in a display device. 48, 107, 108, 110, 112

Planning Reference Volume (PRV) A margin added to certain Organs at Risk, to account for set-up uncertainties. 79, 80, 143, 147, 149, 154, 177, 181, 182, 188–190, 198

Planning Target Volume (PTV) Clinical Target Volume plus a margin for set-up uncertainties. iv, xvi–xviii, 16, 25, 74–76, 79, 110, 139–143, 147, 149, 150, 152, 154, 156–158, 161, 168, 173–175, 177, 178, 181–195, 197, 198, 201

Positron Emission Tomography (PET) Functional imaging technique, used to examine metabolic processes in the body. 8, 22, 79

Prostate Specific Antigen (PSA) Screening test for Prostate Cancer. 3

Radiation Therapy Oncology Group (RTOG) Clinical cooperative group for the purpose of conducting radiation therapy research and cooperative clinical investigations in the USA. 68, 161, 162, 168

Radiotherapy (RT) Use of ionising radiation to treat disease, primarily cancer. iv, 4, 8, 10, 13, 14, 22, 24, 74, 76–79, 110, 133, 154, 188, 196, 199, 200

Radiotherapy Trials Quality Assurance (RTTQA) Group funded by the National Cancer Research Institute (NCRI), whom ensure that patients in all NCRI radiotherapy trials are treated in accordance with the trial protocol, and are treated according to nationally accepted standards. 8, 9, 23

Receiver Operator Characteristic (ROC) Plot of sensitivity of a test against 1-specificity, as a threshold used in the test is varied. Used to assess the effectiveness of a test. 184–186

Region Of Interest (ROI) Organ, region, or group of tissues outlined prior to Radiotherapy Treatment Plan production. Includes targets such as the Planning Target Volume, and Organs At Risk. 87, 88, 112, 149, 200

RTOG Conformance Index (RTOG-CI) Conformance Index proposed by RTOG for measuring how one region of interest conforms to another, calculated as the the union of the two volumes divided by the volume of the first volume, multiplies by the union of the two volumes divided by the volume of the second volume. 161, 166

Secure SHell (SSH) Network protocol using cryptography to provide a secure communications channel over an unsecured network. 67, 86, 113

Single Strand Break (SSB) In radiobiology, a single strand break is a break in one of the two helices in the structure of the DNA molecule. 9

Stoichiometric Of or pertaining to the calculation of relative quantities of chemicals in a compound or mixture. 89

Stopping Power Ratio (SPR) In Bragg-Gray cavity theory, the quantity that relates the absorbed dose to a medium to the absorbed dose to water. xiv, 99–102, 113

Structured Query Language (SQL) Computer language used for the retrieval and management of data stored in relational database management systems. 133

Tissue Air Ratio (TAR) The TAR at a point in a water phantom is the ratio of the total absorbed dose at that point, divided by the total absorbed dose at a point with the same distance from the source, but measured with the phantom shifted so that this point is at the depth of maximum dose. Since the dose at depth of maximum dose is wholly from primary photons, TAR is a way of expressing the dose normalised in terms of primary dose. Despite the name, the current definition does not involve measurement in air. 42

total energy released per unit mass (terma) Concept analogous to kerma, but includes all energy released, not just kinetic energy, i.e. radiative energy is included. 45

Treatment Planning System (TPS) Specialised type of Computer Aided Design software, used to design Radiotherapy Treatment Plans. iv, 14, 15, 22, 24, 25, 37, 44, 61, 64–68, 75, 105, 111, 156, 196, 197

Tumour Control Probability (TCP) In Radiotherapy, the probability of killing enough of the cancerous cells to control the tumour. iv, 165–168, 190, 197, 198

van't Reit Conformation Number (Reit-CI) Conformance Index for measuring how one region of interest conforms to another, calculated as the ratio of the volumes of the two objects of interest. 163, 166, 168

Volumetric Modulated Arc Radiotherapy (VMAT) Type of Radiotherapy in which the gantry rotates, the dose rate is varied, and Multi-Leaf Collimators move across the radiation field during delivery, to produce dose distributions that cannot be achieved with Conformal Radiotherapy. 46, 67, 75, 139, 201

Voxel A portmanteau of Volume Element; one value on a regular grid in three dimensional space. iv, 23, 48, 50, 51, 64, 88, 94, 95, 102, 108, 111, 114, 120, 149, 189, 191, 197, 198, 200, 201

Wedge Factor (WF) Ratio of dose at a reference depth, often taken to be depth of maximum dose, with and without a wedge in the beam. Depends on energy of the beam, angle of wedge, and for dynamic beams the field size. 41

Z Atomic number of an atom, or medium. 29, 30, 32, 34, 37, 53

Bibliography

Agostinelli, S., Allison, J., Amako, K., et al. GEANT4 - a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

Ahmad, A. S., Ormiston-Smith, N., and Sasieni, P. D. Trends in the lifetime risk of developing cancer in Great Britain: comparison of risk for those born from 1930 to 1960. *British Journal of Cancer*, 112(5):943–947, 2015.

Ahnesjö, A. Collapsed cone convolution of radiant energy for photon dose calculation in heterogeneous media. *Medical Physics*, 16(4):577–592, 1989.

Ahnesjö, A. and Aspradakis, M. M. Dose calculations for external photon beams in radiotherapy. *Physics in Medicine and Biology*, 44(11):R99–R155, 1999.

Ahnesjö, A., Saxner, M., and Trepp, A. A pencil beam model for photon dose calculation. *Medical Physics*, 19(2):263–273, 1992.

Alder, B., Fernbach, S., and Rotenberg, M., editors. *Methods in Computational Physics*. Acedemic Press, New York, 1963.

Alexander, A., DeBlois, F., Stroian, G., et al. MMCTP: a radiotherapy research environment for Monte Carlo and patient-specific treatment planning. *Physics in Medicine and Biology*, 52(13):N297–N308, 2007.

Allum, W. H., Stenning, S. P., Bancewicz, J., et al. Long-Term Results of a Randomized Trial of Surgery With or Without Preoperative Chemotherapy in Esophageal Cancer. *Journal of Clinical Oncology*, 27(30):5062–5067, 2009.

Altman, D. G. *Practical Statistics for Medical Research*. Chapman & Hall/CRC, Boca Raton, 1999. ISBN 0-412-27630-5.

Arnfield, M. R., Hartmann Siantar, C., Siebers, J., et al. The impact of electron transport on the accuracy of computed dose. *Medical Physics*, 27(6):1266–1274, 2000.

Baró, J. PENELOPE: An algorithm for Monte Carlo simulation of the penetration and energy loss of electrons and positrons in matter. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 100(1):31–46, 1995.

Benjamini, Y. and Hochberg, Y. Controlling the False Discovery Rate : A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300, 1995.

Bentzen, S. M., Heeren, G., Cottier, B., et al. Towards evidence-based guidelines for radiotherapy infrastructure and staffing needs in Europe: the ESTRO QUARTS project. *Radiotherapy and Oncology*, 75(3):355–365, 2005.

Bland, M. *An Introduction to Medical Statistics*. OUP Oxford, third edition, 2000. ISBN 0-19-263269-8.

British Institute of Radiology. Central axis depth dose data for use in radiotherapy. *British Journal of Radiology*, Supplement 25, 1996. Joint Working Party of the British Institute of Radiology and the Institute of Physics and Engineering in Medicine and Biology.

Brown, F. *MCNP - A General N-Particle Transport Code, Version 5*. Los Alamos National Laboratory, Los Alamos, NM, 2003.

Bush, K., Townson, R., and Zavgorodni, S. Monte Carlo simulation of RapidArc radiotherapy delivery. *Physics in Medicine and Biology*, 53(19):N359–N370, 2008.

Carrington, R., Staffurth, J., Warren, S., et al. The effect of dose escalation on gastric toxicity when treating lower oesophageal tumours: a radiobiological investigation. *Radiation Oncology*, 10(1):236 1–9, 2015.

Carrington, R., Spezi, E., Gwynne, S., et al. The influence of dose distribution on treatment outcome in the SCOPE 1 oesophageal cancer trial. *Radiation Oncology*, 11(1):19 1–8, 2016.

Chetty, I. J., Curran, B., Cygler, J. E., et al. Report of the AAPM Task Group No. 105: Issues associated with clinical implementation of Monte Carlo-based photon and electron external beam treatment planning. *Medical Physics*, 34 (12):4818–4853, 2007.

Cominos, M., Mosleh-Shirazi, M. A., Tait, D., et al. Quantification and reduction of cardiac dose in radical radiotherapy for oesophageal cancer. *British Journal of Radiology*, 78(936):1069–1074, 2005.

Conneely, E., Alexander, A., Stroian, G., et al. An investigation into the use of MMCTP to tune accelerator source parameters and testing its clinical application. *Journal of Applied Clinical Medical Physics*, 14(2):3–14, 2013.

Conneely, E., Alexander, A., Ruo, R., et al. Monte Carlo investigation of collapsed versus rotated IMRT plan verification. *Journal of Applied Clinical Medical Physics*, 15(3):133–147, 2014.

Crosby, T., Hurt, C., Falk, S., et al. Chemoradiotherapy with or without cetuximab in patients with oesophageal cancer (SCOPE1): a multicentre, phase 2/3 randomised trial. *The Lancet Oncology*, 14(7):627–637, 2013.

Cufflin, R. S., Spezi, E., Millin, A. E., and Lewis, D. G. An investigation of the accuracy of Monte Carlo portal dosimetry for verification of IMRT with extended fields. *Physics in Medicine and Biology*, 55(16):4589–4600, 2010.

Deasy, J. O., Blanco, A. I., and Clark, V. H. CERR: A computational environment for radiotherapy research. *Medical Physics*, 30(5):979–985, 2003.

Downes, P. *Monte Carlo Dose Verification of Radiotherapy Treatment Planning.* PhD thesis, National University of Ireland, Galway, 2010.

Downes, P., Yaikhom, G., Giddy, J., et al. High-performance computing for Monte Carlo radiotherapy calculations. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367(1897):2607–2617, 2009.

Duzenli, C., McClean, B., and Field, C. Backscatter into the beam monitor chamber: implications for dosimetry of asymmetric collimators. *Medical physics*, 20 (2 Pt 1):363–7, 1993.

Eyges, L. Multiple Scattering with Energy Loss. *Physical Review*, 74(10):1534–1535, 1948.

Fassó, A., Ferrari, A., and Sala, P. R. Electron-Photon Transport in FLUKA: Status. In A Kling, F Barao, M Nakagawa, L. T. and Vaz, P., editors, *Advanced Monte Carlo for Radiation Physics, Particle Transport Simulation and Applications:Proc. Monte Carlo 2000 Meeting (Lisbon)*, pages 159–164, Berlin, 2000. Springer.

Ferrari, A., Sala, P. R., Fassó, A., and Ranft, J. *Fluka: A Multi-particle Transport Code : (program Version 2005)*. CERN, 2005. ISBN 9789290832607.

Fippel, M. Fast Monte Carlo dose calculation for photon beams based on the VMC electron algorithm. *Medical Physics*, 26(8):1466–1475, 1999.

Fix, M. K., Manser, P., Frei, D., et al. An efficient framework for photon Monte Carlo treatment planning. *Physics in Medicine and Biology*, 52(19):N425–N437, 2007.

Francescon, P., Cavedon, C., Reccanello, S., and Cora, S. Photon dose calculation of a three-dimensional treatment planning system compared to the Monte Carlo code BEAM. *Medical Physics*, 27(7):1579–1587, 2000.

Gatenby, P. A. C., Hainsworth, A., Caygill, C., et al. Projections for oesophageal cancer incidence in England to 2033. *European Journal of Cancer Prevention*, 20(4):283–286, 2011.

Geh, J. I., Bond, S. J., Bentzen, S. M., and Glynne-Jones, R. Systematic overview of preoperative (neoadjuvant) chemoradiotherapy trials in oesophageal cancer: Evidence of a radiation and chemotherapy dose response. *Radiotherapy and Oncology*, 78(3):236–244, 2006.

Gwynne, S., Hurt, C., Evans, M., et al. Definitive Chemoradiation for Oesophageal Cancer - a Standard of Care in Patients with Non-metastatic Oesophageal Cancer. *Clinical Oncology*, 23(3):182–188, 2011.

Gwynne, S., Spezi, E., Wills, L., et al. Toward Semi-automated Assessment of Target Volume Delineation in Radiotherapy Trials: The SCOPE 1 Pretrial Test Case. *International Journal of Radiation Oncology Biology Physics*, 84 (4):1037–1042, 2012.

Gwynne, S., Falk, S., Gollins, S., et al. Oesophageal Chemoradiotherapy in the UK-Current Practice and Future Directions. *Clinical Oncology*, 25(6):368–377, 2013.

Haga, A., Magome, T., Takenaka, S., et al. Independent absorbed-dose calculation using the Monte Carlo algorithm in volumetric modulated arc therapy. *Radiation Oncology*, 9(1):75 1–9, 2014.

Hanahan, D. and Weinberg, R. A. The Hallmarks of Cancer. *Cell*, 100:57–70, 2000.

Hanna, G. G., Hounsell, A. R., and O'Sullivan, J. M. Geometrical Analysis of Radiotherapy Target Volume Delineation: a Systematic Review of Reported Comparison Methods. *Clinical Oncology*, 22(7):515–525, 2010.

Hartmann Siantar, C. L., Walling, R. S., Daly, T. P., et al. Description and dosimetric verification of the PEREGRINE Monte Carlo dose calculation system for photon beams incident on a water phantom. *Medical physics*, 28(7): 1322–1337, 2001.

Heath, E. and Seuntjens, J. Development and validation of a BEAMnrc component module for accurate Monte Carlo modelling of the Varian dynamic Millennium multileaf collimator. *Physics in Medicine and Biology*, 48(24):4045–4063, 2003.

Henzen, D., Manser, P., Frei, D., et al. Monte Carlo based beam model using a photon MLC for modulated electron radiotherapy. *Medical Physics*, 41(2): 021714 1–10, 2014.

Hissoiny, S., Ozell, B., Bouchard, H., and Després, P. GPUMCD: A new GPU-oriented Monte Carlo dose calculation platform. *Medical Physics*, 38(2):754–764, 2011a.

Hissoiny, S., Ozell, B., Després, P., and Carrier, J.-F. Validation of GPUMCD for low-energy brachytherapy seed dosimetry. *Medical Physics*, 38(7):4101–4107, 2011b.

Hissoiny, S., Raaijmakers, A. J. E., Ozell, B., et al. Fast dose calculation in magnetic fields with GPUMCD. *Physics in Medicine and Biology*, 56(16):5119–5129, 2011c.

Holm, S. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979.

Hounsell, A. R. Monitor chamber backscatter for intensity modulated radiation therapy using multileaf collimators. *Physics in Medicine and Biology*, 43(2): 445–454, 1998.

Hughes, D. B., Karzmark, C. J., and Levy, R. M. Conventions for wedge filter specifications. *British Journal of Radiology*, 45:868, 1972.

Hurt, C. N., Nixon, L. S., Griffiths, G. O., et al. SCOPE1: a randomised phase II/III multicentre clinical trial of definitive chemoradiation, with or without cetuximab, in carcinoma of the oesophagus. *BMC Cancer*, 11(1):466 1–11, 2011.

IAEA. *Technical Reports Series No. 430: Commissioning and Quality Assurance of Computerized Planning Systems for Radiation Treatment of Cancer.* International Atomic Energy Agency, Vienna, 2013.

ICRU. *Determination of Absorbed Dose in a Patient Irradiated by Beams of X or Gamma Rays in Radiotherapy Procedures ICRU Report Report 24.* International Commission on Radiation Units and Measurements, Washington, 1976.

ICRU. *Stopping Powers for Electrons and Positions ICRU Report Report 37.* International Commission on Radiation Units and Measurements, Washington, 1984.

ICRU. *Photon, Electron, Proton and Neutron Interaction Data for Body Tissues. ICRU Report 46.* International Commission on Radiation Units and Measurements, Bethesda, 1992.

ICRU. *Prescribing, recording, and reporting photon beam therapy. ICRU Report 50.* International Commission on Radiation Units and Measurements, Bethesda, 1993.

ICRU. *Prescribing, Recording and Reporting Photon Beam Therapy (Supplement to ICRU Report 50). ICRU Report 62.* International Commission on Radiation Units and Measurements, Bethesda, 1999.

Irvine, C., Morgan, A., Crellin, A., et al. The Clinical Implications of the Collapsed Cone Planning Algorithm. *Clinical Oncology*, 16(2):148–154, 2004.

Jahnke, L., Fleckenstein, J., Wenz, F., and Hesser, J. GMC: a GPU implementation of a Monte Carlo dose calculation based on Geant4. *Physics in Medicine and Biology*, 57(5):1217–1229, 2012.

Jang, S. Y., Liu, H. H., Wang, X., et al. Dosimetric verification for intensity-modulated radiotherapy of thoracic cancers using experimental and Monte Carlo approaches. *International Journal of Radiation Oncology Biology Physics*, 66(3):939–48, 2006.

Jena, R., Kirkby, N. F., Burton, K. E., et al. A novel algorithm for the morphometric assessment of radiotherapy treatment planning volumes. *The British Journal of Radiology*, 83(985):44–51, 2010.

Jia, X., Gu, X., Sempau, J., et al. Development of a GPU-based Monte Carlo dose calculation code for coupled electron--photon transport. *Physics in Medicine and Biology*, 55(11):3077–3086, 2010.

Jia, X., Gu, X., Graves, Y. J., et al. GPU-based fast Monte Carlo simulation for radiotherapy dose calculation. *Physics in Medicine and Biology*, 56(22): 7017–7031, 2011.

Johns, H. E. and Cunningham, J. R. *The Physics of Radiology*. Charles C Thomas, Springfield, Illinois, fourth edition, 1983. ISBN 0-398-04669-7.

Kawrakow, I. and Rogers, D. W. O. *The EGSnrc code system: Monte Carlo simulation of electron and photon transport. Technical Report PIRS-701*. Ottawa, Ontario, 2000.

Kawrakow, I. and Fippel, M. VMC^{++}, a MC algorithm optimized for electron and photon beam dose calculations for RTP. In Enderle, J. D., editor, *Proceedings of the 22nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (Cat. No.00CH37143)*, volume 2, pages 1490–1493, 345 E 47TH ST, NEW YORK, NY 10017 USA, 2000. IEEE. ISBN 0-7803-6465-1.

Kawrakow, I., Fippel, M., and Friedrich, K. 3D electron dose calculation using a Voxel based Monte Carlo algorithm (VMC). *Medical Physics*, 23(4):445–457, 1996.

Keall, P. J., Siebers, J. V., Arnfield, M., et al. Monte Carlo dose calculations for dynamic IMRT treatments. *Physics in Medicine and Biology*, 46(4):929–941, 2001.

Kepka, L., Bujko, K., Garmol, D., et al. Delineation variation of lymph node stations for treatment planning in lung cancer radiotherapy. *Radiotherapy and Oncology*, 85(3):450–455, 2007.

Khan, F. M. *The Physics Of Radiation Therapy*. Lippincott Williams and Wilkins, Philadelphia, USA, third edition, 2003. ISBN 0-7817-3065-1.

Knöös, T., Ahnesjö, A., Nilsson, P., and Weber, L. Limitations of a pencil beam approach to photon dose calculations in lung tissue. *Physics in Medicine and Biology*, 40(9):1411–1420, 1995.

Kosunen, A. and Rogers, D. W. O. Beam quality specification for photon beam dosimetry. *Medical Physics*, 20(4):1181–1188, 1993.

Krieger, T. and Sauer, O. A. Monte Carlo- versus pencil-beam-/collapsed-cone-dose calculation in a heterogeneous multi-layer phantom. *Physics in Medicine and Biology*, 50(5):859–868, 2005.

Lam, K. L., Muthuswamy, M. S., and Ten Haken, R. K. Measurement of backscatter to the monitor chamber of medical accelerators using target charge. *Medical Physics*, 25(3):334–338, 1998.

Liu, H. H., Mackie, T. R., and McCullough, E. C. Modeling photon output caused by backscattered radiation into the monitor chamber from collimator jaws using a Monte Carlo technique. *Medical Physics*, 27(4):737–744, 2000.

Liu, H. H., Keall, P., and Hendee, W. R. D_m rather than D_w should be used in Monte Carlo treatment planning. *Medical Physics*, 29(5):922–924, 2002.

Lobo, J. and Popescu, I. A. Two new DOSXYZnrc sources for 4D Monte Carlo simulations of continuously variable beam configurations, with applications to RapidArc, VMAT, TomoTherapy and CyberKnife. *Physics in Medicine and Biology*, 55(16):4431–4443, 2010.

Ma, C.-M., Pawlicki, T., Jiang, S. B., et al. Monte Carlo verification of IMRT dose distributions from a commercial treatment planning optimization system. *Physics in Medicine and Biology*, 45(9):2483–2495, 2000.

Ma, C.-M., Li, J. S., Pawlicki, T., et al. A Monte Carlo dose calculation tool for radiotherapy treatment planning. *Physics in Medicine and Biology*, 47(10): 1671–1689, 2002.

Mackeprang, P.-H., Volken, W., Terribilini, D., et al. Assessing dose rate distributions in VMAT plans. *Physics in Medicine and Biology*, 61(8):3208–3221, 2016.

Mackie, T. R., Bielajew, A. F., Rogers, D. W. O., and Battista, J. J. Generation of photon energy deposition kernels using the EGS Monte Carlo code. *Physics in Medicine and Biology*, 33(1):1–20, 1988.

Maddams, J., Utley, M., and Møller, H. Projections of cancer prevalence in the United Kingdom, 2010-2040. *British Journal of Cancer*, 107(7):1195–1202, 2012.

McDonald, J. *Handbook of Biological Statistics*. Sparky House Publishing, Baltimore, Maryland, third edition, 2014.

Metropolis, N. and Ulam, S. The Monte Carlo Method. *Journal of the American Statistical Association*, 44(247):335–341, 1949.

Millin, A. E. and Lewis, G. Routine Dosimetric Verification of Stereotactic Treatments Using Monte Carlo Methods in a Distributed Computing Environment. In *Radiosurgery*, pages 28–43, Basel, 2010. Karger.

Mu, W., Chen, Z., Liang, Y., et al. Staging of cervical cancer based on tumor heterogeneity characterized by texture features on 18 F-FDG PET images. *Physics in Medicine and Biology*, 60(13):5123–5139, 2015.

Muijs, C. T., Schreurs, L. M., Busz, D. M., et al. Consequences of additional use of PET information for target volume delineation and radiotherapy dose distribution for esophageal cancer. *Radiotherapy and Oncology*, 93(3):447–453, 2009.

Mukherjee, S., Hurt, C. N., Gwynne, S., et al. NEOSCOPE: a randomised Phase II study of induction chemotherapy followed by either oxaliplatin/capecitabine or paclitaxel/carboplatin based chemoradiation as pre-operative regimen for resectable oesophageal adenocarcinoma. *BMC Cancer*, 15(1):48 1–9, 2015.

Olsen, A. H., Parkin, D. M., and Sasieni, P. Cancer mortality in the United Kingdom: projections to the year 2025. *British Journal of Cancer*, 99(9):1549–1554, 2008.

Orlhac, F., Soussan, M., Maisonobe, J.-A., et al. Tumor Texture Analysis in ^{18}F-FDG PET: Relationships Between Texture Parameters, Histogram Indices, Standardized Uptake Values, Metabolic Volumes, and Total Lesion Glycolysis. *Journal of Nuclear Medicine*, 55(3):414–422, 2014.

Parkin, D. M., Boyd, L., and Walker, L. C. 16. The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. *British Journal of Cancer*, 105(S2):S77–S81, 2011.

Peedell, C., editor. *Concise Clinical Oncology*. Elsevier, Philadelphia, USA, 2005. ISBN 0-7506-8836-X.

Popescu, I. A., Shaw, C. P., Zavgorodni, S. F., and Beckham, W. A. Absolute dose calculations for Monte Carlo simulations of radiotherapy beams. *Physics in Medicine and Biology*, 50(14):3375–3392, 2005.

Popescu, I. A., Atwal, P., Lobo, J., et al. Patient-specific QA using 4D Monte Carlo phase space predictions and EPID dosimetry. *Journal of Physics: Conference Series*, 573(1):012004 1–10, 2015.

Quaresma, M., Coleman, M. P., and Rachet, B. 40-year trends in an index of survival for all cancers combined and survival adjusted for age and sex for each cancer in England and Wales, 1971-2011: A population-based study. *The Lancet*, 385(9974):1206–1218, 2015.

Rodríguez, M., Sempau, J., and Brualla, L. PRIMO: A graphical environment for the Monte Carlo simulation of Varian and Elekta linacs. *Strahlentherapie und Onkologie*, 189(10):881–886, 2013.

Rodríguez, M. L. PENLINAC: extending the capabilities of the Monte Carlo code PENELOPE for the simulation of therapeutic beams. *Physics in Medicine and Biology*, 53(17):4573–4593, 2008.

Rogers, D. W. O. Fifty years of Monte Carlo simulations for medical physics. *Physics in Medicine and Biology*, 51(13):R287–R301, 2006.

Rogers, D. W. O., Faddegon, B. A., Ding, G. X., et al. BEAM: A Monte Carlo code to simulate radiotherapy treatment units. *Medical Physics*, 22(5):503–524, 1995.

Schneider, U., Pedroni, E., and Lomax, A. The calibration of CT Hounsfield units for radiotherapy treatment planning. *Physics in Medicine and Biology*, 41(1):111–24, 1996.

Schneider, W., Bortfeld, T., and Schlegel, W. Correlation between CT numbers and tissue parameters needed for Monte Carlo simulations of clinical dose distributions. *Physics in Medicine and Biology*, 45(2):459–78, 2000.

Seco, J. and Verhaegen, F., editors. *Monte Carlo Techniques in Radiation Therapy*. CRC Press, Boca Raton, 2013. ISBN 9781466507920.

Seeley, R. R., Stephens, T. D., and Tate, P. *Anatomy & Physiology*. McGraw-Hill, Boston, fifth edition, 2000.

Sempau, J., Wilderman, S. J., and Bielajew, A. F. DPM, a fast, accurate Monte Carlo code optimized for photon and electron radiotherapy treatment planning dose calculations. *Physics in Medicine and Biology*, 45(8):2263–2291, 2000.

Shapiro, A. S. S. and Wilk, M. B. An Analysis of Variance Test for Normality. *Biometrika*, 52(3):591–611, 1956.

Shaw, E., Kline, R., Gillin, M., et al. Radiation therapy oncology group: Radiosurgery quality assurance guidelines. *International Journal of Radiation Oncology Biology Physics*, 27(5):1231–1239, 1993.

Shaw, E., Scott, C., Souhami, L., et al. Single dose radiosurgical treatment of recurrent previously irradiated primary brain tumors and brain metastases: final report of RTOG protocol 90-05. *International Journal of Radiation Oncology Biology Physics*, 47(2):291–298, 2000.

Siebers, J. V., Keall, P. J., Nahum, A. E., and Mohan, R. Converting absorbed dose to medium to absorbed dose to water for Monte Carlo based photon beam dose calculations. *Physics in Medicine and Biology*, 45(4):983–995, 2000.

Siebers, J. V., Keall, P. J., Kim, J. O., and Mohan, R. A method for photon beam Monte Carlo multileaf collimator particle transport. *Physics in Medicine and Biology*, 47(17):3225–3249, 2002.

Spezi, E. and Lewis, G. An overview of Monte Carlo treatment planning for radiotherapy. *Radiation Protection Dosimetry*, 131(1):123–129, 2008.

Spezi, E., Lewis, D. G., and Smith, C. W. A DICOM-RT-based toolbox for the evaluation and verification of radiotherapy plans. *Physics in Medicine and Biology*, 47(23):4223–4232, 2002.

Steel, G. G. *Basic Clinical Radiobiology.* Arnold, London, second edition, 1997. ISBN 0340700203.

Stewart, B. W. and Wild, C. P. *World Cancer Report 2014.* World Health Organisation, Lyon, 2014. ISBN 978-92-832-0443-5.

Tian, Z., Shi, F., Folkerts, M., et al. A GPU OpenCL based cross-platform Monte Carlo dose calculation engine (goMC). *Physics in Medicine and Biology*, 60(19): 7419–7435, 2015.

Townson, R. W., Jia, X., Tian, Z., et al. GPU-based Monte Carlo radiotherapy dose calculation using phase-space sources. *Physics in Medicine and Biology*, 58(12):4341–56, 2013.

Tufte, E. R. *The Visual Display of Quantitative Information.* Graphics Press USA, second edition, 2001. ISBN 0961392142.

Vanderstraeten, B., Chin, P. W., Fix, M., et al. Conversion of CT numbers into tissue parameters for Monte Carlo dose calculations: a multi-centre study. *Physics in Medicine and Biology*, 52(3):539–562, 2007.

van't Riet, A., Mak, A. C. A., Moerland, M. A., et al. A conformation number to quantify the degree of conformality in brachytherapy and external beam irradiation: Application to the prostate. *International Journal of Radiation Oncology Biology Physics*, 37(3):731–736, 1997.

Verhaegen, F. and Seuntjens, J. Monte Carlo Modelling of External Radiotherapy Photon Beams. *Physics in Medicine and Biology*, 48(21):R107–R164, 2003.

Verhaegen, F., Symonds-Tayler, R., Liu, H. H., and Nahum, A. E. Backscatter towards the monitor ion chamber in high-energy photon and electron beams: charge integration versus Monte Carlo simulation. *Physics in Medicine and Biology*, 45(11):3159–3170, 2000.

Verhaegen, F. and Liu, H. H. Incorporating dynamic collimator motion in Monte Carlo simulations: an application in modelling a dynamic wedge. *Physics in medicine and biology*, 46(2):287–96, 2001.

Wang, L., Yorke, E., and Chui, C.-S. Monte Carlo evaluation of 6 MV intensity modulated radiotherapy plans for head and neck and lung treatments. *Medical Physics*, 29(11):2705–2717, 2002.

Whitley, E. and Ball, J. Statistics review 4: sample size calculations. *Critical care (London, England)*, 6(4):335–41, 2002.

Wills, L., Millin, A., Paterson, J., et al. The effect of planning algorithms in oesophageal radiotherapy in the context of the SCOPE 1 trial. *Radiotherapy and Oncology*, 93(3):462–467, 2009.

Yaikhom, G., Giddy, J., Walker, D., and Downes, P. A distributed simulation framework for conformal radiotherapy. In *Parallel and Distributed Processing, 2008. IPDPS 2008. IEEE International Symposium on*, pages 1–8, 345 E 47TH ST, NEW YORK, NY 10017 USA, 2008. IEEE. ISBN 978-1-4244-1693-6.

Yamamoto, T., Mizowaki, T., Miyabe, Y., et al. An integrated Monte Carlo dosimetric verification system for radiotherapy treatment planning. *Physics in Medicine and Biology*, 52(7):1991–2008, 2007.

Yip, S. S. F. and Aerts, H. J. W. L. Applications and limitations of radiomics. *Physics in Medicine and Biology*, 61(13):R150–R166, 2016.

Yu, M. K., Sloboda, R. S., and Mansour, F. Measurement of photon beam backscatter from collimators to the beam monitor chamber using target-current-pulse-counting and telescope techniques. *Physics in Medicine and Biology*, 41(7):1107–1117, 1996.

Zavgorodni, S., Bush, K., Locke, C., and Beckham, W. Vancouver Island Monte Carlo (VIMC) system for radiotherapy treatment planning dosimetry and research. *Radiotherapy and Oncology*, 84(Supplement 1):S49, 2007.