# INTEGRATION OF DAILY IMAGING, PLAN ADAPTATION AND RADIATION DELIVERY FOR NEAR REAL-TIME ADAPTIVE RADIATION THERAPY

by

ANTE MESTROVIC

B.Sc., Simon Fraser University, 2003

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES
(Physics)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)

January 2009

# Abstract

The primary objective of this research was to develop and implement a new approach to on-line adaptive radiation therapy (ART) in which daily imaging, plan adaptation and radiation delivery are temporally integrated and performed concurrently. The advantages of this approach are: reduction of treatment time compared to conventional on-line ART; ability to perform a complete plan re-optimization with minimal extension of treatment time; ability to detect and correct for intra-fractional patient motion.

This work was motivated by an initial study which compared four radiosurgery techniques. This study was the first quantitative analysis of the correlation between patient anatomy and the quality of treatment plans produced by different radiosurgery techniques. A number of different relationships, both qualitative and quantitative, were discovered between patient anatomy and dosimetric parameters for different techniques. The results were used to successfully predetermine the optimal radiosurgery technique based on patient anatomy.

The first step in developing a new approach to on-line ART involved accelerating plan adaptation (re-optimization) using direct aperture optimization (DAO). A series of techniques for adapting the original treatment plan to correct for the deterioration of dose distribution quality caused by the anatomical deformations were investigated. Through modification of the DAO algorithm the optimization search space was reduced and the plan adaptation was significantly accelerated. Next, a new approach to on-line ART was proposed and investigated, in which accelerated plan adaptation and radiation delivery were integrated together and performed concurrently. A fundamental advantage of this approach is that most of the plan re-optimization was performed during radiation delivery, so the time spent adapting the original plan did not significantly increase the overall treatment time. Finally, daily imaging, accelerated plan adaptation and radiation delivery were all temporally integrated using an integrated Linac/Cone Beam CT system. Intra-fractional patient images were used to successfully re-optimize the original treatment plan in near real-time to account for anatomy deformations.

This thesis concludes that integration of daily imaging, plan adaptation and radiation delivery for near real-time ART is both feasible and advantageous. With further advances in related technologies, it promises to become a part of clinical practice in the near future.

# Contents

# List of Tables

# List of Figures

# List of Abbreviations

**ARD** Adaptive Radiation Delivery

**ART** Adaptive Radiation Therapy

**AVM** Arteriovenous Malformation

**BEV** Beam's Eye View

**CAT** Conformal Arc Therapy

**CBCT** Cone Beam Computed Tomography

**CRT** Conformal Radiation Therapy

**CT** Computed Tomography

**CTV** Clinical Target Volume

**DAO** Direct Aperture Optimization

**DRR** Digitally Reconstructed Radiograph

**DSA** Digital Subtraction Angiography

**DTS** Digital Tomosynthesis

**DVH** Dose Volume Histogram

**EPID** Electronic Portal Imaging Device

**FFT** Fast Fourier Transform

**GTV** Gross Tumor Volume

**HU** Hounsfield Unit

**IMAT** Intensity Modulated Arc Therapy

**IMM/PDAF** Interacting Multiple Model Probabilistic Data Association Filter

**IMRS** Intensity Modulated Radiosurgery

**IMRT** Intensity Modulated Radiation Therapy

**KE** Kinetic Energy

**KERMA** Kinetic Energy Released in Medium

**KV** Kilovoltage

**LSF** Line Spread Function

**MDPD** ratio of Maximum Dose to Prescription Dose

**MLC** Multi Leaf Collimator

**MRI** Magnetic Resonance Imaging

**MU** Monitor Unit

**MV** Megavoltage

**NPS** Noise Power Spectrum

**NTCP** Normal Tissue Complication Probability

**OAR** Organ at Risk

**PBDD** Pencil Beam Dose Distribution

**PBK** Pencil Beam Kernel

**PET** Positron Emission Tomography

**PITV** ratio of volume within Prescription Isodose surface to Target Volume

**PSF** Point Spread Function

**PTV** Planning Target Volume

**QA** Quality Assurance

**RMSE** Root Mean Squared Error

**ROI** Region of Interest

**RTOG** Radiation Therapy Oncology Group

**SAD** Source-to-Axis Distance

**SAVR** Surface Area to Volume Ratio

**SDD** Source-to-Detector Distance

**SPECT** Single Photon Emission Computed Tomography

**SRS** Stereotactic Radiosurgery

**SSD** Source-to-Surface Distance

**TCP** Tumor Control Probability

**TPS** Treatment Planning System

**VMAT** Volumetric Modulated Arc Therapy

# Acknowledgments

I would like to acknowledge my supervisors Dr. Brenda Clark and Dr. Karl Otto for their endless support during the course of this thesis. With her continuous encouragement and guidance, Brenda has helped me stay motivated and kept me on track throughout my research. Karl's enthusiasm and passion for research have been truly inspiring. He has been both a mentor and a valued colleague.

I would also like to thank the other members of my thesis committee, Dr. Alan Nichol, Dr. Alex MacKay, Dr. Andre Marziali, and Dr. Vesna Sossi for their contribution and important feedback on my work. This research has been partly funded by NSERC and MSFHR studentships. I am appreciative and very thankful for their support.

I am deeply grateful to my fellow students, in particular my former office mates Dr. Peter Petric and Dr. Alanah Bergman. I have gotten to know you both well, thanks in part to the very little office that we shared. Thank you all for your friendship and many laughs along the way.

My fondest appreciation goes to my family, especially my mother Olga and father Dujo for their unconditional love and encouragement. To my younger brother Igor - thanks for the much needed comic relief and support. I would also like to thank my friends Maja Grubisic and Bosko Radanov for a decade of friendship and many adventures during our student years.

Most importantly, I am eternally thankful to my fiancée Andrea Maglajlija - my foremost inspiration. Thank you for your love and patience during this journey. I will love you.

# Chapter 1
# Introduction to Radiation Therapy Physics

## 1.1 Historical Background

In November of 1895, the German physicist Wilhelm Conrad Roentgen produced and detected electromagnetic radiation in a wavelength range which is known today as x-rays. Roentgen's original paper, "On a New Kind of Rays", was published on December 28, 1895, fifty days after his discovery. He published a total of 3 papers on x-rays between 1895 and 1897. In 1901, Roentgen was awarded the first Nobel Prize in physics for his work on x-rays. Roentgen died in 1923 of carcinoma of the intestine. However, it is believed that his carcinoma was not a result of his work with ionizing radiation because of the brief time spent on those investigations and because he was one of the few pioneers in the field who used protective lead shields routinely. Only one year after their discovery, it was realized that x-rays could be used for diagnostic as well as therapeutic purposes[3]. This marked the birth of the field of radiation therapy.

Another important milestone in the field of radiation therapy was the discovery of radium in 1898 by Marie and Pierre Curie. Radium (Ra) is a radioactive chemical element with the atomic number of 88. Radium has 25 different known isotopes, with radium-226 being the most stable one. The half-life of radium-226 is 1602 years. Radium is over one million times more radioactive than the same mass of Uranium. It decays in at least seven stages and the main products are polonium, lead, bismuth and radon. Radon (Rn) is a heavy gas and it is itself radioactive. During radium decay alpha particles, beta particles and gamma rays are produced. The gamma rays are used for radiation therapy with the energy spectrum peaks at around 0.55 MeV and 1.65 MeV. Soon after radium was discovered, radium tubes quickly became a popular form of radiation therapy, as they could be inserted into cavities or surgically implanted. In 1903, the Royal Swedish Academy of Sciences awarded Pierre Curie and Marie Curie (together with Henri Becquerel) the Nobel Prize in Physics. Marie Curie was the first woman to be awarded a Nobel Prize. Eight years later, she received the 1911 Nobel Prize in Chemistry,

making her the first twice-honored Nobel laureate (to this day, the only one to win the award in two different science fields). She died in 1934, at the age of 66, almost certainly due to exposure to radiation. Much of her work had been carried out with no safety measures, as the damaging effects of ionizing radiation were not yet known. She had carried test tubes containing radioactive isotopes in her pocket, remarking on the pretty blue-green light the substances gave off in the dark.

With the discovery of x-rays and radioactive radium the stage was set for the radiation therapy field. However, early 1900's are viewed as the "Dark Ages" in the evolution of the radiation therapy. Surgeons and dermatologists started administering the radiation therapy treatments, with little understanding of the physical nature and biological effects of radiation. There were no methods for calculating the radiation dosage given to patients. The equipment used to produce and deliver radiation was primitive and limited in energy. The x-rays had very weak penetrating abilities and could be used for treatment of superficial tissue only. The radiation therapy of this period involved a massive exposure of radiation to a large area of the body with the hope that the cancer would be destroyed with a single treatment. The concept of dose fractionation was not yet introduced. Not surprisingly, the success rate of radiation therapy treatments was very low, with many complications occurring after the treatments due to the unrecoverable destruction of the healthy tissue. The death incidence as a result of radiation treatments was high. Also, the rate of tumor recurrence was high.

During the 1920 -1930's some major advances and developments occurred in the field of radiation therapy. Innovations in the x-ray tube designs during the 1920's increased the peak energy of the x-rays up to 200 keV, allowing for successful treatments of superficial tumors. It became clear that machines capable of producing higher energies x-rays had to be developed to effectively treat more deeply seated tumors. In 1926, the international unit of x-ray exposure, the roentgen (R), was introduced. It was a measure of the output of an x-ray generator, defined as the charge liberated by ionizing radiation per unit mass of air. In SI units, exposure is expressed in coulombs (C) per kilogram (kg). Another major advancement of this era was the introduction of the concept of dose fractionation in the 1930's. Various experiments investigating the radiobiological basis of radiation therapy led to the realization that delivering radiation in multiple daily treatments had several radiobiological advantages over single treatments: It became evident that dose fractionation has four major effects on tumor and normal tissue cells: 1) Repair of normal tissue, 2) Reoxygenation of tumor cells to increase radiosensitivity, 3)

Redistribution of tumor cells within the cell cycle, and 4) Repopulation of tumor cells. These four processes were named the 4 R's of radiation biology and form the basis for modern radiation therapy[4].

The 1940 – 1950's marked the beginning of the megavoltage era. The machines capable of producing megavoltage x-rays were introduced, making it possible to effectively treat deeply seated tumors. One of the first devices capable of producing megavoltage x-rays was the Van de Graaff generator integrated with a particle accelerator. The generator was capable of producing high electrostatic fields, which were used to accelerate particles up to 10 MeV. Van de Graaff generators are no longer being built for radiotherapy applications. The main disadvantage of these machines is the limitation of terminal voltage due to voltage breakdowns. The cobalt-60 treatment units were introduced in 1950. The first clinical cobalt-60 treatment unit in North America was installed at the Saskatchewan University Hospital in Saskatoon, Canada in 1951. The cobalt-60 treatment unit generated an average photon energy of 1.25 MeV and had a higher specific activity than radium. It was also safer to handle and less expensive than radium. It provided a simple and reliable way of producing and delivering megavoltage radiation. While still in limited use today, cobalt-60 treatment units have been gradually phased out and replaced with the linear accelerator (linac). Linacs were introduced into the field of radiation therapy during the second half of the 1950's. Medical linacs accelerate electrons using a tuned-cavity waveguide in which the energy is transferred from an electric field of the microwaves to the moving electrons. Medical linacs produce monoenergetic electron beams between 4 and 25 MeV. The monoenergetic electrons are directed at a high-density (such as tungsten) target, producing an x-ray output with a spectrum of energies up to the electrons' energy. In addition to being able to generate much higher energy x-rays than cobalt-60 units, linear accelerators can simply be powered off when not in use.

During the 1960 -1970's linear accelerators were further developed and refined. Linacs capable of delivering adequate doses of radiation to tumors located in almost all areas of the body were now available. The introduction of computers into radiation therapy treatment planning allowed for complex mathematical calculations, necessary to determine the dose distribution within the treatment area, to be completed with accuracy and speed. Controlled clinical trials to determine the biological effect of radiation were also underway during this time period. Concurrently with the further refinements of linacs, various other therapies were being developed during the second half of the twentieth century. In 1954, protons[5] started to be used

as a treatment modality for deeply seated tumors. Negative pi-meson (pion) therapy[6, 7] started during the 1970's but was discontinued twenty years later, as no significant benefit over conventional photon therapy was observed. Neutron therapy[8, 9] trials begun during the 1970's in North America. Neutron therapy has been shown to be effective in the treatment of certain tumors, particularly slow growing tumors which are radioresistant to conventional photon therapy.

During the 1980-1990's the attention was directed towards development and improvement of techniques for treatment plan optimization and radiation delivery to the patient. The goal was to deliver a localized lethal dose of radiation to the cancerous cells while sparing the surrounding healthy tissue as much as possible. The introduction of a radiation beam shaping device, the multi-leaf collimator (MLC), and integration of 3D patient anatomical information into the radiation therapy planning process, has allowed for highly conformal 3D dose distributions to be delivered to the tumor cells. The next milestone in treatment planning and radiation delivery was the introduction of inverse planning and intensity modulated radiation therapy (IMRT)[1]. In IMRT, inverse planning is employed to optimize complex photon fluence intensity maps across each radiation beam. The MLC is used to deliver radiation based on the optimized fluence maps, thus delivering optimal, spatially-varying amount of radiation across each beam. The net result is a precisely tailored dose distribution around the tumor and maximized sparing of adjacent normal tissue.

In the 21$^{st}$ century, the main focus in radiation therapy has been directed towards the precise definition of the clinical target. With the highly advanced methods of treatment plan optimization and radiation delivery, it is the imprecision in clinical target definition that is currently the limiting factor in radiation therapy treatment[10]. Various functional imaging modalities (e.g. MRI, PET and SPECT) are currently being integrated into the field of radiation therapy to better localize the biological extent of malignant cells[11, 12]. Finally, the introduction of various integrated treatment/imaging systems (e.g. fluoroscopy, Cone Beam CT) has allowed 3D patient images to be acquired immediately before each treatment fraction[13]. The ability to accurately assess the patient anatomy prior to radiation delivery has led to the development of adaptive radiation therapy (ART). In ART, the most current patient anatomy is used to re-optimize the original treatment plan to account for anatomical deformations. With the introduction of ART, the radiation therapy treatment planning has moved from the realm of 3D

planning to 4D planning, where the optimal treatment plan changes during the course of the treatment.

## 1.2 Photon Interactions with Matter

When a photon beam passes through an absorbing medium such as human body, photons may interact with the atoms of the medium. The probability of the interaction is described by a quantity called the linear attenuation coefficient, $\mu$. For a photon beam, passing through a thickness $x$ of absorbing material with a linear attenuation coefficient $\mu$, the number of photons expected to penetrate through the medium without interacting is given by Equation 1.1.

$$\left(\frac{N}{N_o}\right) = e^{-\mu x} \tag{1.1}$$

Where:

$N_o$ = initial number of photons incident on the medium

$N$ = number of photons expected to penetrate through the medium without interacting

There are several different types of interactions that may occur between the incoming photon and the atoms of the medium[14]. For the photon energies relevant to radiation therapy (up to 25 MeV), there are six types of interactions a photon may undertake: Rayleigh scattering ($\sigma_{coherent}$), Photoelectric effect ($\tau$), Compton scattering ($\sigma_{incoherent}$), Pair production ($\kappa$), Triplet production ($\kappa_t$) and Photo-nuclear interaction ($\gamma_n$). The specific linear attenuation coefficient for each interaction type is shown in brackets. In photo-nuclear interaction an energetic photon enters and excites a nucleus, which then emits a proton or neutron. Photo-nuclear interaction only occurs for photon energies greater than about 10 MeV, but even then its occurrence is negligible[15]. Thus, photo-nuclear interaction has been commonly neglected altogether in dosimetry considerations (i.e. $\gamma_n = 0$). The total linear attenuation coefficient $\mu$ is given by the sum of the individual specific linear attenuation coefficients (Equation 1.2).

$$\mu = \sigma_{coherent} + \tau + \sigma_{incoherent} + \kappa + \kappa_t \tag{1.2}$$

Often, the linear attenuation coefficient is scaled by the physical density, $\rho$, to create the mass attenuation coefficient ($\mu/\rho$). The mass attenuation coefficient is independent of the material's density and reveals the dependence of the linear attenuation coefficient per gram of the material.

## 1.2.1 Rayleigh (coherent) Scattering

Rayleigh scattering is a process in which the energy of the scattered photon is equal to the energy of the incident photon. No energy is converted into kinetic energy and all of it is scattered. The incident photon transfers all of its energy to the entire atom and sets the electrons in the atom into momentary vibration. These oscillating electrons emit radiation of the same wavelength as the incident radiation. The emitted radiation from electrons combines to give coherent scattering. Rayleigh scattering occurs mainly in the forward direction, so the overall effect of the process is to slightly broaden the angular width of a beam. The specific linear attenuation coefficient for Rayleigh scattering, $\sigma_{coherent}$, decreases rapidly with the increase of incident photon energy and becomes negligibly small for energies greater than about 100 keV in low atomic number materials (e.g. human body).

## 1.2.2 Photoelectric Effect

In the photoelectric effect, an incident photon of energy $h\nu$ interacts with an atom and ejects one of the bound electrons from one of the electron shells (K, L, M, N, etc.). The ejected electron is called a photoelectron and emerges with the kinetic energy $KE$, given by Equation 1.3.

$$KE = h\nu - E_b \tag{1.3}$$

where $E_b$ is the binding energy of the shell from which the electron is ejected. The photoelectric effect is illustrated in Figure 1.1.

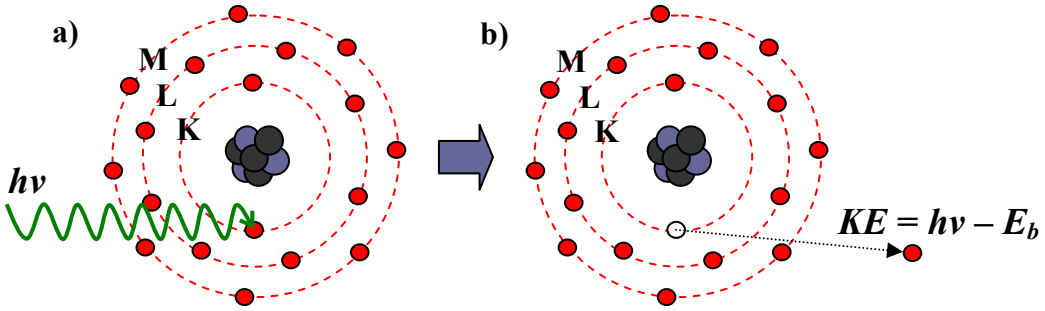


*Figure 1.1: Photoelectric effect: a) An incident photon of energy hv interacts with bound atomic electron from one of the electron shells (e.g. K shell). b) The ejected electron is called a photoelectron and emerges with the kinetic energy KE.*

The incident photon can eject an electron from any of the electron shells, but the probability of ejecting an electron from the K shell, having the strongest binding energy, is the highest (given that the incident photon has enough energy to overcome K shell binding energy). Once the photoelectron is ejected, the atom is left in an excited state and emits characteristic radiation or Auger electrons as it returns to the ground state. Characteristic radiation is emitted as an electron from a higher electron shell (e.g. L shell) drops down to the vacant hole in the K shell. The energy of the characteristic radiation is equal to the energy difference between the two electron shells. Alternatively, the characteristic radiation may be re-absorbed by an electron from an even higher electron shell (e.g. M shell) resulting in this electron, called Auger electron, to be ejected from the atom. In this case, the atom is left with two holes, one in L and the other in M shell. The mass attenuation coefficient for the photoelectric effect ($\tau/\rho$) decreases rapidly with the increase of incident photon energy ($\sim 1/E^3$ dependence) and increases with the increase of atomic number of the medium ($\sim Z^3$ dependence). It becomes negligibly small for energies greater that about 100 keV in low atomic number materials (e.g. human body).

## 1.2.3 Compton (incoherent) Scattering

Compton scattering involves the interaction of an incident photon of energy *hv* and loosely bound or free electrons (binding electron energy << *hv*). An incident photon is scattered by an electron and some of the energy of the incident photon is transferred to the recoil electron as kinetic energy *KE* (Figure 1.2).

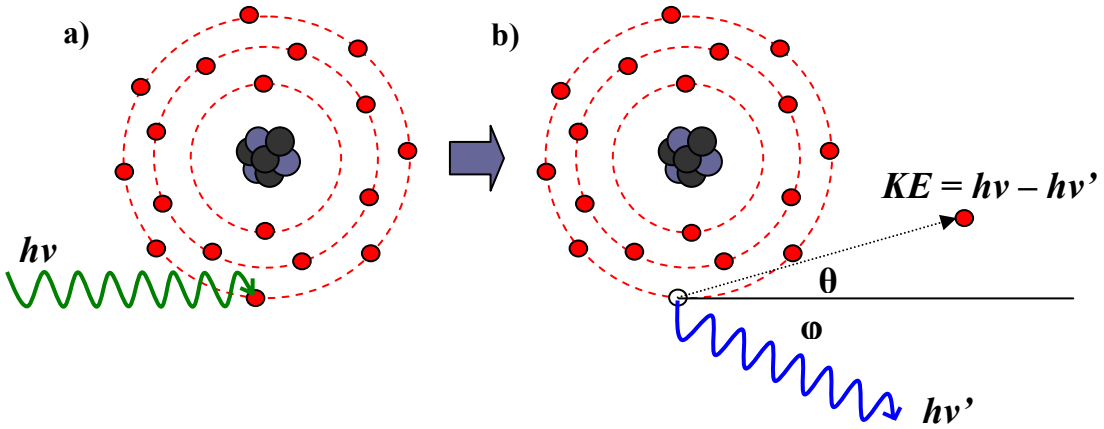


*Figure 1.2: Compton scattering: a) An incident photon of energy hv interacts with a free atomic electron (binding electron energy << hv). b) Incident photon is scattered by the electron and some of the energy of the incident photon is transferred to the recoil electron as kinetic energy KE.*

From conservation of energy and momentum equations, the kinetic energy (*KE)* and the scattering angle θ of the recoil electron can be related to the scattered photon energy *hv'* and its scattering angle φ, through Equations 1.4 to 1.6.

$$hv' = \frac{hv}{1 + (hv / m_o c^2)(1 - \cos\varphi)} \tag{1.4}$$

$$KE = hv - hv' \tag{1.5}$$

$$\cot\theta = \left(1 + \frac{hv}{m_o c^2}\right)\tan\left(\frac{\varphi}{2}\right) \tag{1.6}$$

where $m_o c^2$ is the rest energy of the electron, 0.511 MeV. In low atomic number materials (e.g. human body), the Compton scattering process is the most dominant interaction for photon energies in the range 100 keV to 10 MeV. The mass attenuation coefficient for Compton scattering ($\sigma_{incoherent}/\rho$) decreases with increase in incident photon energy ($\sim 1/E^{1/2}$ dependence) and it is almost independent of the atomic number of the medium.

### 1.2.4 Pair and Triplet Production

Pair production involves the interaction of an incident photon of energy *hv* and the strong electric field of an atomic nucleus. As the photon passes close to the nucleus and is subjected to its electric field, it may decay into an electron and positron (Figure 1.3a). For pair production to occur the incident photon must have energy greater than 1.022 MeV (twice the rest energy of the electron). If the photon has energy in excess of 1.022 MeV, the extra energy is transferred as kinetic energy to the electron ($KE^{electron}$) and the positron ($KE^{positron}$):

$$KE^{electron} + KE^{positron} = hv - 1.022\ MeV \tag{1.7}$$

During the pair production process, the atomic nucleus recoils with some momentum and acquires a very small amount of energy. This energy is negligible in comparison to the energy transferred to the electron and positron, and is therefore neglected in Equation 1.7. The total kinetic energy given to the electron and the positron ($KE^{electron} + KE^{positron}$) can be divided between the two particles in many different ways. The mass attenuation coefficient for pair production ($\kappa/\rho$) increases with the incident photon energy and the atomic number of the medium. In low atomic number materials, pair production starts to have a significant role for photon energies greater than 10 MeV.

Triplet production is similar to pair production, except that the incident photon interacts with the electric field of a bound atomic electron, instead of the nucleus. The atomic electron is also set free and joins the newly created electron-positron pair to form a triplet of particles (Figure 1.3b). From conservation of energy and momentum, it can be shown that the energy threshold for triplet production is twice the threshold for pair production, or 2.044 MeV. The probability of triplet production is relatively small compared to pair production (~1%) and the linear attenuation coefficients for the two interactions are commonly combined and quoted together ($\kappa+\kappa_t$).

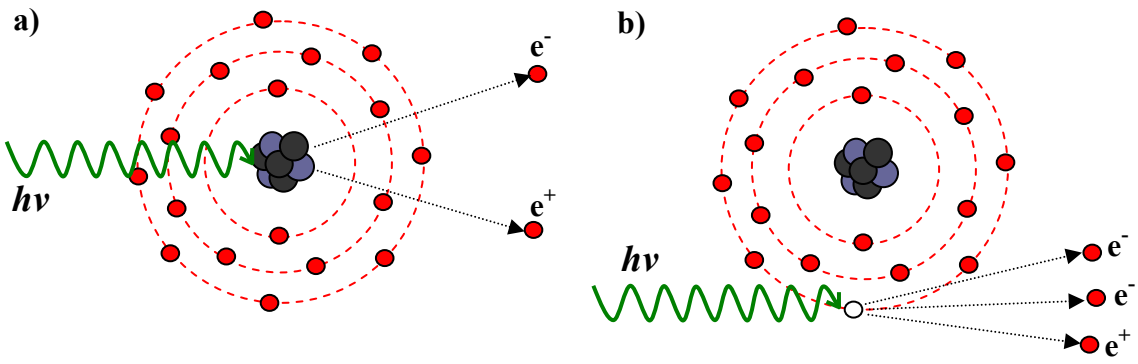


*Figure 1.3: a) Pair Production: in the presence of the strong electric field of an atomic nucleus, an incident photon of energy hv is converted into an electron $e^-$ and positron $e^+$. b) Triplet Production: in the presence of the electric field of a bound atomic electron, an incident photon of energy hv is converted into an electron $e^-$ and positron $e^+$ and the atomic electron is set free, joining the newly created electron-positron pair to form a triplet of particles.*

# 1.3 Charged Particle Interactions with Matter

In the previous section, photon interactions with matter were examined. Five types of interactions between an incident photon and the atoms of the medium were described: Rayleigh scattering, photoelectric effect, Compton scattering, pair and triplet production. In all of these processes, except Rayleigh scattering, the energy of the incident photon is transferred to the charged particles of the medium (i.e. electrons and positrons). In this section, the subsequent interactions of these energetic charged particles with the atoms of the medium are described.

## 1.3.1 Ionizational Energy Losses

As an energetic charged particle passes through the absorbing medium it collides with the atomic electrons of the medium. The collisions cause the charged particle to lose energy by ionizing the atoms along its path. The energy loss per unit thickness of the medium, $(dE/dx)_{ion}$ is called the ionizational stopping power. A closely related quantity is the mass ionizational stopping power $S_{ion}$, which is obtained by dividing ionizational stopping power by the density of the medium $\rho$.

$$S_{ion} = \frac{1}{\rho}\left(\frac{dE}{dx}\right)_{ion} \tag{1.8}$$

The $S_{ion}$ is usually expressed in $MeVcm^2/g$. In deriving the equation for $S_{ion}$ for charged particle with the relatively small mass, such as an electron, both quantum mechanics and relativity must be considered. The equation for $S_{ion}$ for electrons was first derived by Bethe (Equation 1.9).

$$S_{ion} = 2\pi r_o^2 N_e \frac{\mu_o}{\beta^2}\left[\ln\frac{E^2(E+2\mu_o)}{2\mu_o I^2} + \frac{E^2/8 - (2E+\mu_o)\mu_o \ln 2}{(E+\mu_o)^2} + 1 - \beta^2 - \delta\right] \tag{1.9}$$

Where:

$r_o$ = classical electron radius ($2.82 \times 10^{-15}$ m)

$N_e$ = number of electrons per gram

$\mu_o = m_e c^2$ = rest energy of the electron

$\beta = v/c$ = incident electron relative velocity

$E$ = kinetic energy of incident electron

$I$ = mean ionizational energy

$\delta$ = density correction

## 1.3.2 Radiative Energy Losses

In addition to the interactions with the atomic electrons of the medium, a charged particle interacts with the atomic nuclei of the medium as it passes through. As a moving charged particle passes close to the nucleus it experiences an electrostatic force, which causes it to accelerate or decelerate. For electrons, the electrostatic force causes them to decelerate. As the electrons slow down they radiate energy, a process known by the German word bremsstrahlung meaning "breaking radiation". The rate of energy loss by radiation is described by the mass radiative stopping power $S_{rad}$. For electron energies less than a 100 MeV, $S_{rad}$ is given by Equation 1.10, where $Z$ is the atomic number of the medium and all the other symbols have the same meaning as for Equation 1.9.

$$S_{rad} = \frac{1}{\rho}\left(\frac{dE}{dx}\right)_{rad} = 4r_o^2 \frac{N_e ZE}{137}\left[\ln\frac{2(E+\mu_o)}{\mu_o} - \frac{1}{3}\right] \tag{1.10}$$

The bremsstrahlung radiation continues propagating thought the medium (e.g. human body) until it is reabsorbed by the medium through one of the five photon interactions with matter described earlier. In addition to occurring in patients during radiation treatments, the production of bremsstrahlung radiation is also the main method of generating high energy x-rays in x-ray tubes and medical linacs.

Currently, most of the x-rays used for external beam radiation therapy (radiation source is outside the patient) are generated through the bremsstrahlung process. In x-ray tubes, a high voltage potential between the cathode and anode accelerates electrons towards the anode target. In medical linacs, the electrons are accelerated in a waveguide using microwave energy. In both cases the accelerated electrons are directed at a high atomic number ($Z$) target, producing bremsstrahlung radiation. From Equation 1.10 it can be seen that production of bremsstrahlung radiation is directly proportional to the atomic number of the medium, which explains the use of high Z target material. The target is usually made out of tungsten (Z=74). The energy radiated by the decelerating electrons can be any amount up to the kinetic energy of the incident electron. Bremsstrahlung x-rays therefore exhibit a spectrum of energies from 0 to a maximum equal to the kinetic energy of the incident electron beam. In x-ray tubes, electrons are accelerated up to a few hundred keV before they strike the anode, producing an x-ray beam with the peak energy of a few hundred keV. In medical linear accelerators, the electrons are accelerated in a waveguide to an energy between 4 and 25 MeV, producing a megavoltage x-ray beam with a peak energy between 4 and 25 MeV. An x-ray beam with a spectrum of energies up to, say, 10 MeV is commonly referred to as a 10 Megavoltage (MV) beam.

### 1.3.3 Positron Annihilation

During pair and triplet production an electron-positron pair is created. If the incident photon's energy is in excess of 1.022 MeV, the extra energy is divided between the electron and positron as kinetic energy (Equation 1.7). As the energetic positron travels through the medium, it interacts with the atomic electrons and atomic nuclei of the medium. Similar to the electron, the positron transfers its kinetic energy to the medium through ionizational and radiative energy

losses. Once the positron loses most of its kinetic energy and comes to near rest, it interacts with a nearby electron and both particles undergo a conversion from mass into energy, through a process known as positron-electron annihilation. Since the positron and electron rest energies are 0.511 MeV each, the total energy released is 1.022 MeV. To conserve momentum, two photons, each with 0.511 MeV, are ejected in opposite directions from the location of the annihilation. If the electron and/or positron have some small residual energy at the moment when the annihilation occurs, this energy is added to the radiation energy released. Also, to conserve momentum, the two photons may not be exactly aligned to 180 degrees. The positron annihilation process is illustrated in Figure 1.4.

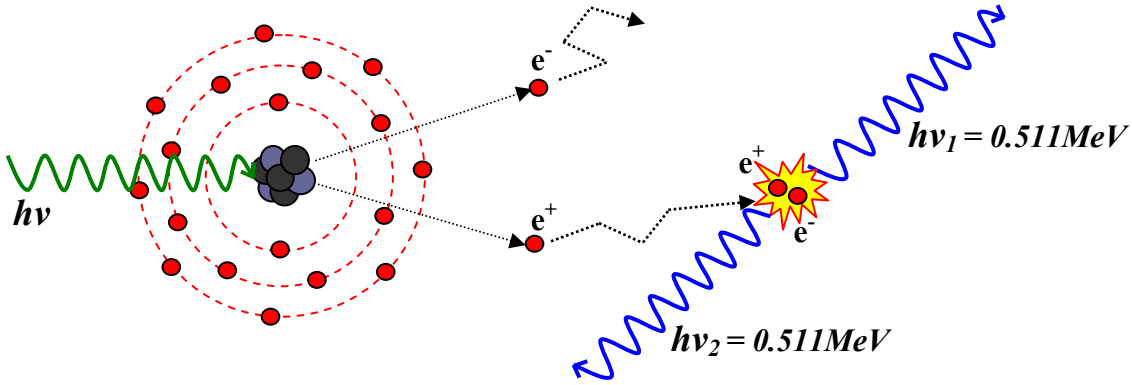


*Figure 1.4: Positron Annihilation: following a pair production, the created positron travels through the medium transferring its kinetic energy to the medium through ionizational and radiative energy losses. Once the positron loses most of its kinetic energy and comes to near rest, it interacts with a nearby electron and both particles undergo a conversion from mass into energy. To conserve energy and momentum, two photons, each with 0.511 MeV, are ejected in opposite directions from the scene of the annihilation.*

## 1.4 Energy Transfer and Absorbed Dose

In two previous sections, photon interactions with matter and charged particles interactions with matter were examined. In this section, the link between these interactions and the energy transfer and absorbed dose is established. It will be shown that: 1) photon interactions with matter are directly linked to the energy transfer to the medium, and 2) charged particles interactions with matter are linked to the absorbed dose.

### 1.4.1 KERMA - Kinetic Energy Released in the Medium

As incident photons interact with the atoms of the medium (via interactions described in Section 1.2) they transfer a relatively large fraction of their energy to the charged particles (i.e. electrons) present in the medium. This initial interaction is described by a quantity called KERMA (kinetic energy released in the medium). KERMA is defined as the kinetic energy transferred from photons to electrons ($dE_{tr}$) in a volume element whose mass is $dm$ (Equation 1.11).

$$KERMA = \frac{dE_{tr}}{dm} \tag{1.11}$$

For a beam of monoenergetic photons with energy $hv$ and fluence $\Phi$, defined as a number of photons per unit area, the KERMA is given by:

$$KERMA = \Phi \cdot \left( \frac{\mu}{\rho} \right) \cdot \left\langle E_{tr} \right\rangle \tag{1.12}$$

Where $\mu$ is the total linear attenuation coefficient, given by Equation 1.2, $\rho$ is the medium density, and $\langle E_{tr} \rangle$ is the average amount of energy transferred to electrons of the medium at each interaction. The product of $\Phi\ (\mu/\rho)$ gives the number of photon interactions that occur per unit mass of material. The units of KERMA are joules per kilogram.

For a polyenergetic photon beam with a spectrum of photon energies up to $hv_{max}$, the KERMA is given by integrating over all the photon energies:

$$KERMA = \int_0^{hv_{\max}} \frac{d\Phi(hv)}{dhv} \cdot \left( \frac{\mu(hv)}{\rho} \right) \cdot \left\langle E_{tr}(hv) \right\rangle \cdot dhv \tag{1.13}$$

## 1.4.2 Absorbed Dose

Once the energy is transferred from an incident photon to an electron of the medium, the electron is set in motion through the medium. The energetic electron deposits its energy along its track by ionizing the atoms of the medium (Equation 1.9). The energy imparted to the medium in bringing about these ionizations is the absorbed dose. Absorbed dose is defined as the energy imparted ($dE_{ab}$) by the ionizing radiation to a mass *dm* of the medium (Equation 1.14).

$$Dose = \frac{dE_{ab}}{dm} \tag{1.14}$$

Similar to KERMA, the units of absorbed dose are joules per kilogram. A special unit called *gray* (Gy) was introduced for absorbed dose, defined as 1 Gy = 1 J/kg.

There are two important differences between KERMA and absorbed dose: 1) KERMA is the total energy transferred from an incident photon to an electron, but not all of this energy is retained by the medium. Some of this energy is radiated away as bremsstrahlung radiation (Equation 1.10). Absorbed dose refers only to the fraction of the electron energy deposited to the medium through ionizational energy losses, $S_{ion}$; 2) KERMA refers to a single interaction between an incident photon and an electron of the medium and occurs at one particular location in the medium. Dose absorption, on the other hand, occurs all along electron track through multiple ionizations. A single high energy (MeV) electron set in motion can bring about hundreds of thousands of ionizations along its track. Therefore, KERMA occurs at one location but the dose is absorbed throughout the electron track (range). The electron range, *R*, can be calculated by considering the initial electron energy $E_o$ and the total energy losses per unit thickness of the medium:

$$R = \frac{1}{\rho}\int_0^{E_o} \frac{dE}{S_{ion}(E) + S_{rad}(E)} = \frac{1}{\rho}\int_0^{E_o} \frac{dE}{S_{tot}(E)} \tag{1.15}$$

Where $S_{ion}$ and $S_{rad}$ are given by Equations 1.9 and 1.10, respectively and $\rho$ is the density of the medium. As a rule of thumb, the total stopping power $S_{tot}$ is around 2 MeV/cm for electrons with

energies above 200 keV moving through water equivalent medium (e.g. human body). Therefore, a 10 MeV electron will travel through ~5 cm of water equivalent medium, ionizing thousands of atoms along the way, before it loses all of its kinetic energy and comes to rest.

Analogous to Equation 1.12, one can calculate the absorbed dose using the following equation:

$$Dose = \Phi \cdot \left( \frac{\mu}{\rho} \right) \cdot \left\langle E_{ab} \right\rangle \tag{1.16}$$

Where $\Phi$ is photon fluence, $\mu$ is the total linear attenuation coefficient, given by Equation 1.2, $\rho$ is the medium density, and $<E_{ab}>$ is the part of the average kinetic energy transferred to electrons that contributes to ionization (it excludes energy lost by bremsstrahlung). However, while the Equation 1.12 can be used to calculate KERMA at any point in the medium where the photon fluence is known, equation 1.16 is valid only for the regions where an *electronic equilibrium* is established.

## 1.4.3 Electronic Equilibrium

To illustrate an electronic equilibrium, consider a monoenergetic beam of photons incident on an absorbing medium. As photons penetrate through the medium, their energy is transferred to electrons. The electrons are set in motion and start moving through the medium, ionizing many atoms along the way, until they lose all of their kinetic energy and come to rest. The distance electrons travel before stopping is given by the electron range R (Equation 1.15). Since the electrons travel and deposit their energy "downstream" from their starting position, the number of electrons set in motion increases with depth, until depth R is reached. This is illustrated in Figure 1.5a: square A is crossed only by electrons originating in square A; square B is crossed by electrons originating in square A and B; square C is crossed by electrons originating in square A, B and C. Therefore, the number of electrons tracks increases with depth, until depth R (between square C and D) is reached. The portion of the medium from the surface to depth R is called the build-up region, and the portion beyond depth R is called the region of electronic equilibrium (square D and beyond). In the region of electronic equilibrium, the number of electrons stopping in any volume is equal to the number of electrons originating from it.

The KERMA can be calculated using Equation 1.12 at any depth, but the absorbed dose can be calculated using Equation 1.16 only in the region of electronic equilibrium (depth R and beyond). Figure 1.5b illustrates the relationship between KERMA and the absorbed dose, assuming that no bremsstrahlung losses occur (i.e. $\langle E_{ab} \rangle = \langle E_{tr} \rangle$), and no attenuation of the photon beam occurs in the medium. In this case, the absorbed dose is equal to KERMA beyond the buildup region. If the attenuation of the photon beam is taken into account the relationship between KERMA and the absorbed dose is illustrated in Figure 1.5c. Beyond the build-up region, the absorbed dose and KERMA both decrease exponentially, but the absorbed dose curve is always above the KERMA curve if bremsstrahlung losses are neglected. This is the case since the absorbed dose at any point beyond the peak is due to the electrons originating from further "upstream". A true electronic equilibrium is never established since there are always slightly more electrons stopping in any volume than originating from it. In this case, the absorbed dose differs from the KERMA by a factor slightly greater than 1.00. This factor can be included in Equation 1.16 to calculate the absorbed dose beyond the build-up region.

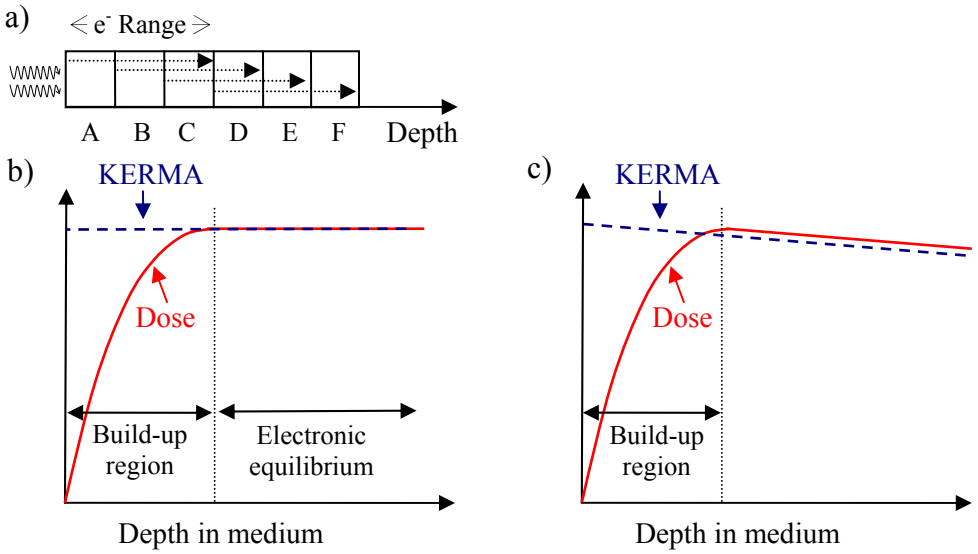


*Figure 1.5: a) Schematic diagram illustrating number of electron tracks as a function of medium depth. b) The relationship between KERMA (kinetic energy released in the medium) and the absorbed dose, assuming that no bremsstrahlung losses occur (i.e. $\langle E_{ab} \rangle = \langle E_{tr} \rangle$), and no attenuation of the photon beam occurs in the medium. c) The relationship between KERMA and the absorbed dose when attenuation of the photon beam occurs in the medium.*

# 1.5 Tumor Control and Normal Tissue Complication Probabilities

In the previous sections of this chapter, it was shown that photon interactions with matter cause charged particles (i.e. electrons) to be set in motion through the medium. These energetic electrons deposit their energy along their tracks by ionizing hundreds of thousands of atoms in the medium. The energy imparted to the medium in bringing about these ionizations is the absorbed dose. It this section, the relationship between the absorbed dose and the successful control of a cancerous tumor, without inducing negative side-effects to healthy tissue, is examined.

The main goal of radiation therapy for treatment of cancer is to deliver a highly localized lethal dose of ionizing radiation to the tumor cells while minimizing the dose delivered to the surrounding healthy, non-cancerous cells. The requirement to localize the dose to cancer cells can be understood by considering two types of probability curves commonly used in radiation therapy: *tumor control probability* (TCP) and *normal tissue complication probability* (NTCP). The TCP curve describes the probability of suppressing a cancer recurrence as a function of dose given to cancerous cells. The NTCP curve describes the probability of inducing a negative side-effect as a function of dose given to normal tissue. If these two curves were separated enough, with the TCP curve to the left and NTCP curve to the right (Figure 1.6a), the importance of localizing the dose to tumor cells, while minimizing the dose to normal tissue, would not be as crucial. For example, irradiating both the tumor and the normal tissue surrounding the tumor to a uniform dose, $D_1$, would still result in a very high TCP (close to 100%) and low NTCP (close to 0%). In this case, delivering a uniform dose, $D_1$, to a large region around the tumor site would most likely be able to successfully control the tumor, with low probability of causing negative side-effect to the surrounding normal tissue. However, in reality, the TCP and NTCP may be very close together, and often the TCP curve is placed to the right of the NTCP curve (Figure 1.6b). The exact shapes and placements of TCP and NTCP curves are specific to different tumor types, tumor sites and different healthy tissues and organs. In the second case, irradiating the tumor and the normal tissue surrounding the tumor to a uniform dose, $D_1$, would result in inadequate TCP and NTCP. In this case, a great care has to be taken to deliver a highly localized lethal dose (e.g. $D_2$) to the tumor cells while minimizing the dose delivered to the surrounding

normal tissue (e.g. < $D_3$). Since the healthy normal tissue is often in direct contact with the tumor, simultaneously satisfying both of these requirements can be a very challenging task, requiring very steep dose gradients around the tumor. Fulfilling these two somewhat opposing goals constitutes an optimization problem and leads to the field of treatment planning.

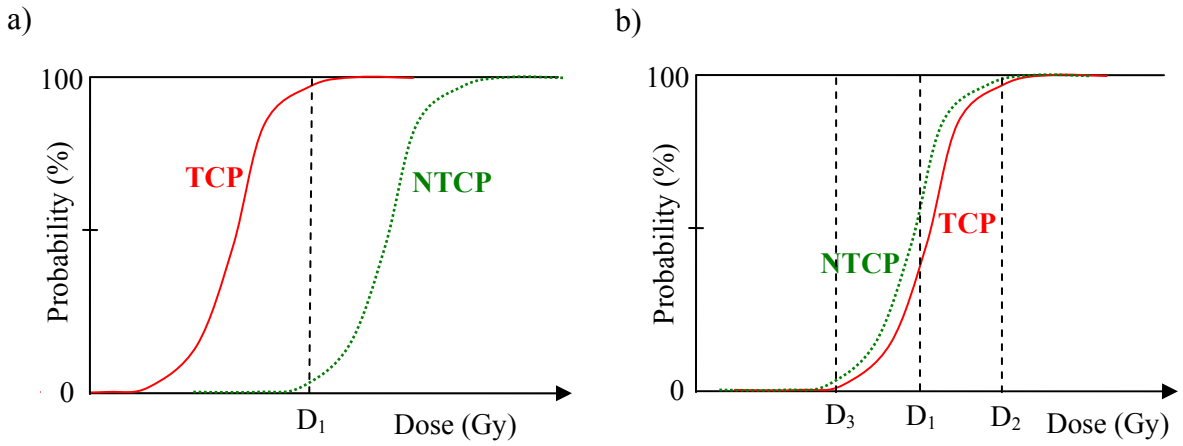


*Figure 1.6: a) Well-separated tumor control probability (TCP) and normal tissue complication probability (NTCP) curves, with TCP curve to the left of NTCP curve. b) A more realistic and challenging scenario, in which TCP and NTCP curves are very close together, often with TCP curve to the right of NTCP curve.*

# Chapter 2
# Introduction to Treatment Planning and On-line Adaptive Radiation Therapy

## 2.1 Initial Patient Imaging

The first stage of treatment planning involves imaging the patient to precisely determine patient's internal anatomy. Computed tomography (CT) is routinely used for patient imaging prior to start of treatment planning process (Figure 2.1). During a CT scan, an x-ray source emits x-rays through a patient as it rotates around a patient. On current generation CT scanners, an x-ray emitter follows a helical trajectory around a patient. On the opposite side of the x-ray emitter, a bank of detectors is used to collect radiation transmitted through a patient from many different x-ray source positions. The x-ray beam is much narrower in one dimension and is referred to as a fan beam. The fan shape of the x-ray beam results in imaging only a thin axial slice of a patient during a single revolution around a patient. The imaging slice thickness can vary from a few millimeters to a few centimeters depending on the particular CT scanner capabilities and various user adjustable parameters. By combining together the x-ray projections data, acquired from different x-ray source positions, a 3D distribution of linear attenuation coefficients within a patient is reconstructed. Such a distribution of linear attenuation coefficients constitutes a CT image. The linear attenuation coefficients are displayed in terms of CT numbers. The CT numbers range from -1000 for air to +1000 for bone, with that for water set at 0. The CT numbers normalized in this manner are referred to as Hounsfield numbers, after the English engineer G. N. Hounsfield, who is now generally recognized as the inventor of computed tomography. The relationship between a Hounsfield unit (HU) and the linear attenuation coefficient $\mu$ is given by Equation 2.1.


*Figure 2.1: Computed tomography (CT) is routinely used prior to start of treatment planning process to precisely determine patient's 3D internal anatomy (image of Philips Brilliance CT scanner from www.uchospitals.edu).*

$$HU = \frac{\mu - \mu_{water}}{\mu_{water}} \times 1000 \tag{2.1}$$

Where $\mu_{water}$ is the linear attenuation coefficient of water. From Equation 2.1 is can be seen that HU is directly proportional to the linear attenuation coefficient $\mu$. A change of 1 HU represents a change of 0.1% of the linear attenuation coefficient relative to that of water.

The information obtained from a CT scan is vital in two important aspects of the treatment planning process. First, a CT image is used for a precise delineation of the clinical target and the surrounding structures in relation to the external contour (Figure 2.2). This is possible due to the sub-millimeter in-slice spatial resolution, together with a good bone to soft tissue contrast provided by a CT image. Second, linear attenuation coefficients obtained from a CT image are used to directly infer electron density (electrons/cm$^3$) in the patient. For low Z materials (e.g. lung, muscle, fat) the photons are predominantly attenuated by Compton interactions, therefore the electron density is directly proportional to the linear attenuation coefficients. This follows from the fact that the mass attenuation coefficient for Compton scattering ($\sigma_{incoherent}/\rho$) is independent on the atomic number of the medium (Section 1.2.3). For high Z materials (e.g. bone) the relationship becomes non-linear due to the increased importance

of photo-electric effect for high Z materials. The non-linear relationship is due to the fact that the mass attenuation coefficient for photoelectric effect ($\tau/\rho$) has ~$Z^3$ dependence on atomic number of the medium (Section 1.2.2). A non-linear calibration curve is usually generated which relates the CT number (linear attenuation coefficient) to the electron density even for high Z materials[16]. The knowledge of exact electron densities in a patient is required for an accurate calculation of the dose distribution within a patient.

In addition to a CT scan, various functional imaging modalities (e.g. MRI, PET and SPECT) are currently being integrated into the field of radiation therapy to aid in better definition of the clinical target. By combining anatomical and function imaging modalities a more accurate localization of the biological extent of malignant cells is achievable. Image registration, a process of transforming different sets of data into one coordinate system, plays a key role in integrating data from various imaging modalities together and is currently an active area of research.

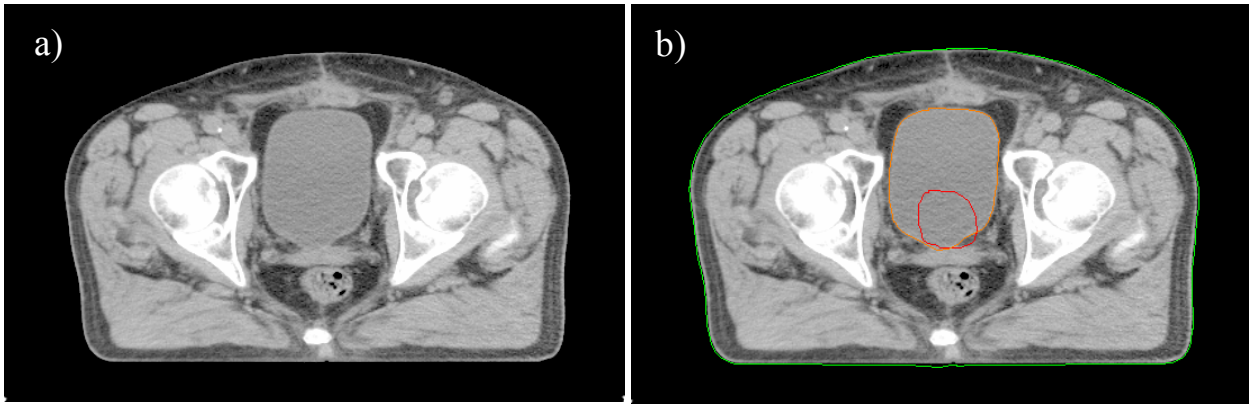


*Figure 2.2: a) An axial CT slice through a pelvic region b) A clinical target (red contour) and the bladder (orange contour) in relation to the external contour (green contour) are delineated to be used in treatment planning.*

# 2.2 Treatment Planning

## 2.2.1 Planning Volumes

After imaging the patient with a CT scanner, and possibly some other functional imaging modality, the treatment planning process commences. From the image data set, various planning volumes are contoured (delineated) to be used in treatment planning (Figure 2.3):

- Gross Tumor Volume (GTV) is the gross visible extent of the malignant growth.
- Clinical Target Volume (CTV) is the volume that contains GTV plus all the microscopic extensions of the disease, which have to be eliminated.
- Planning Target Volume (PTV) is obtained by adding a margin around the CTV. The added margin accounts for the net effect of all the possible geometrical uncertainties and inaccuracies in order to ensure that the prescribed dose is actually absorbed in the CTV. The uncertainties and inaccuracies can arise from various factors such as: patient motion during treatment; physiological movement and variations in size, shape and position of the CTV; uncertainties and variations in patient positioning; mechanical uncertainties of the equipment; dosimetric uncertainties; uncertainties associated with patient imaging, etc. Margin sizes can vary from a few millimeters up to more than a centimeter, depending on the treatment site and the treatment technique. The PTV is normally used for treatment planning and for prescription of dose.
- Organ at Risk (OAR) is a healthy organ at risk of complications, also known as a critical structure.

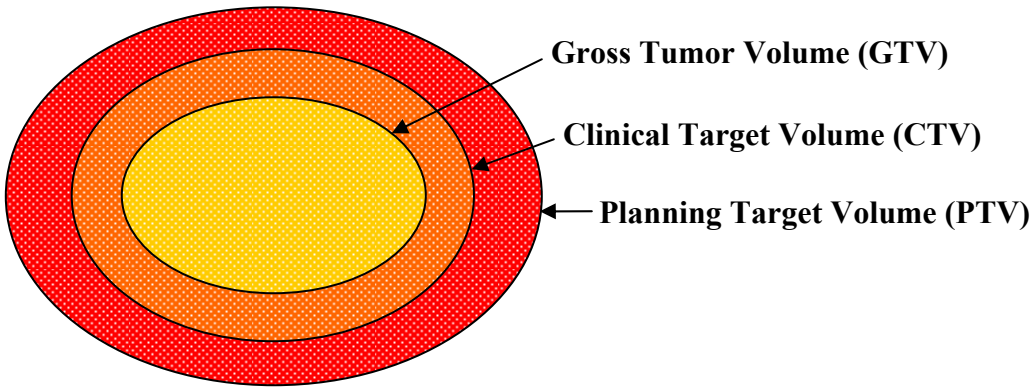


*Figure 2.3: Treatment planning volumes*

## 2.2.2 Treatment Planning Objectives

Once the planning volumes are defined, treatment planning objectives are specified in terms of the dose constraints placed on the PTV and the surrounding OARs. The prescribed dose to the PTV represents the dose which should be absorbed by the entire PTV, to ensure adequate dosage of the CTV. Dose homogeneity throughout the PTV is usually kept within +7% and -5% of the prescribed dose to avoid significant overdosing and, more importantly, underdosing of the PTV. The OARs dose constraints are specified to avoid unrecoverable damage to the OARs. The OAR dose constraints are usually specified in terms of dose-volume constraints. The dose-volume constraints are defined by specifying the maximum volume, say $V_1$, which can receive a dose greater that some specified value, say $D_1$. For each OAR, multiple dose-volume constraints are usually specified. To aid in visualization of the prescribed dose to the PTV, in relation to the multiple dose-volume constraints placed on the surrounding OARs, special types of histograms are often used. They are called cumulative *dose-volume histograms* (DVHs).

The cumulative DVH displays a volume of a specific structure V, often specified as a percentage of the total structure's volume, which receives a dose greater than a dose D. In Figure 2.4a, the prescribed dose constraints placed on the PTV and two dose-volume constraints placed on the surrounding OAR are displayed on a cumulative DVH plot. The goal is to deliver dose between $D_p$-5% and $D_p$+7% to the entire PTV, where $D_p$ is the prescribed dose, while keeping dose given to volumes $V_1$ and $V_2$ of the OAR below $D_1$ and $D_2$, respectively. In Figure 2.4b, the solid curves represent the treatment plan with dose distribution which meets the specified dose-volume constraints. On the other hand, the dashed curves represent the treatment plan with dose distribution which violates the specified dose-volume constraints for both the PTV and the OAR.

In addition to indicating the dose-volume constraints placed on the PTV and the surrounding OARs, the number of treatment fractions is also specified. As mentioned earlier, various experiments led to realization that delivering radiation in multiple daily fractions has several radiobiological advantages over the use of a single fraction. Both dose-volume constraints (e.g. prescribed dose to the PTV) and the fractionation schemes are treatment site specific and can greatly vary between treatment sites. For example, for a typical prostate cancer case the prescribed dose to the PTV is around 70 Gy and is delivered over around 35 fractions (i.e. 2 Gy per fraction). On the other hand, treatment of lesions or functional disorders in the brain requires a prescribed dose to the PTV of around 20 Gy delivered in a single fraction.

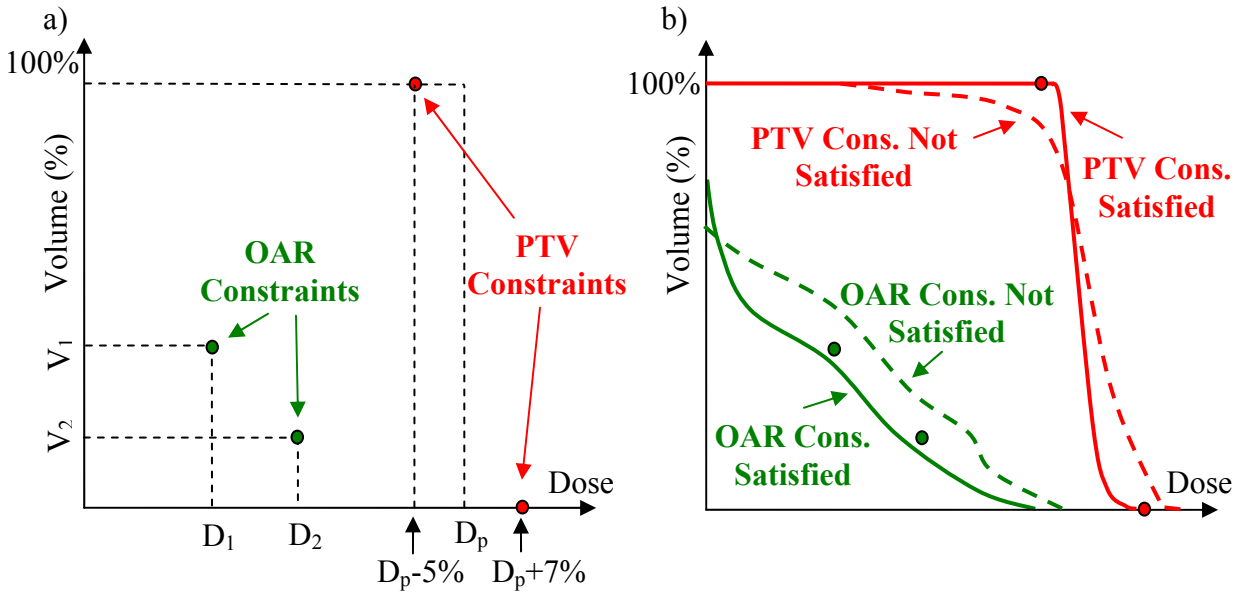


*Figure 2.4: a) The prescribed dose constraints placed on the PTV (red markers) and two dose-volume constraints placed on the surrounding OAR (green markers) are displayed on a cumulative dose-volume histogram (DVH). The goal is to deliver dose between $D_p$-5% and $D_p$+7% to the entire PTV, where $D_p$ is the prescribed dose, while keeping dose given to volumes $V_1$ and $V_2$ of the OAR below $D_1$ and $D_2$, respectively. b) The solid curves represent the treatment plan with dose distribution which meets the specified dose-volume constraints. The dashed curves represent the treatment plan with dose distribution which violates the specified dose-volume constraints for both the PTV and the OAR.*

## 2.3 Treatment Modalities in Radiation Therapy

There are various treatment modalities available in radiation therapy. They can be divided into two main categories: brachytherapy and external beam radiotherapy.

Brachytherapy is a form of radiation therapy where a radioactive source (e.g. Iridium-192, Iodine-125) is placed inside or next to the area requiring treatment. Brachytherapy is very effective in treating localized cancers and offers superior healthy tissue sparing compared to external beam radiation therapy. Brachytherapy is commonly used to treat localized prostate cancer and cervical cancer. Brachytherapy is often administered using small radioactive rods (seeds) implanted directly into the cancerous target. Some of the disadvantages of prostate brachytherapy include a risk of infection and bleeding during seed implementation, as well as

erectile dysfunction following the procedure. Erectile dysfunction is observed in 30-35 percent of men five years after seed implementation. Another disadvantage is the migration of the loose radioactive seeds out of treatment region. Seed migration can be prevented by the use of stranded seeds, which are linked together with a non-radioactive material that is eventually absorbed by the body.

External beam radiotherapy, where the radiation source is outside the patient, is the most frequently used form of radiation therapy. External beam radiotherapy is subdivided into three main categories: electron therapy, hadron (protons, neutrons and heavier ions) therapy and photon (x-rays) therapy. Each of these modalities offers different advantages for specific treatment sites and types of cancer. Electrons have a well defined range of penetration and offer a distinct advantage over photons for sites where a sharp cut-off in the dose beyond the tumor is required. Electron beams are useful for treating superficial lesions because the maximum of dose deposition occurs near the surface. Protons have desirable dose deposition characteristics, with most of the dose deposited towards the end of their range, resulting in a *Bragg peak*. This property makes them favorable for deeply seated tumors. Protons also exhibit low lateral spread in tissue. Neutrons been shown to be effective in the treatment of certain tumors, particularly slow growing tumors which are radioresistant to conventional photon therapy. Photon (x-ray) therapy is the most frequent type of external beam radiotherapy. Kilovoltage (kV) x-rays, produced by x-ray tubes, are used for treating skin cancer and superficial structures. Megavoltage (MV) x-rays are used to treat deep-seated tumors (e.g. bladder, bowel, prostate, lung, brain). Megavoltage x-rays are today most commonly produced using medical linear accelerators (linacs).

### 2.3.1 Medical Linear Accelerator (Linac)

Linacs were introduced into the field of radiation therapy during the second half of the 1950's and have gradually replaced cobalt-60 units for most types of treatments. Linacs accelerate electrons using a tuned-cavity waveguide in which the energy is transferred from the electric field of the microwaves to the moving electrons (Figure 2.5). The electrons are injected into the waveguide by an electron gun, with an initial energy of about 50 keV. The microwaves are produced by a klystron or a magnetron, which is carefully coupled to the waveguide. Medical linacs are capable of accelerating electrons to between 4 and 25 MeV. After exiting the

accelerating waveguide, the monoenergetic electrons are bent by 270 degrees by a bending magnet and directed at a high-density (such as tungsten) target. The x-rays are produced in the target by bremsstrahlung processes, as described in Section 1.3.2, resulting in a spectrum of x-ray energies. The x-rays emitted from the target are forward peaked, more intense in the centre than at the edge of the field, so an appropriately shaped flattening filter is used to flatten the profile of the x-ray beam. A monitor ion chamber is located below the flattening filter. It is used to monitor beam symmetry, dose rate and integrated dose. Linear accelerators are calibrated such that a given "beam on" time delivers a specific dose to the patient for a given set of reference conditions. "Beam on" time is measured in units called Monitor Units (MUs) and is related to the amount of charge collected by the monitor ion chamber. The amount of charge collected by the ion chamber per MU is typically calibrated such that 1 MU delivers 1 cGy of dose to a water-equivalent material at the depth of maximum dose absorption, $d_{max}$, for the field size of 10 cm x 10 cm. Under the monitor ion chamber, the primary and secondary tungsten collimators are placed to define rectangular field shapes. The independent motion of the secondary tungsten collimating jaws allows for variable rectangular field sizes to be created. Finally, an optional beam shaping device (MLC) can be added to allow for further collimation of the field shape.

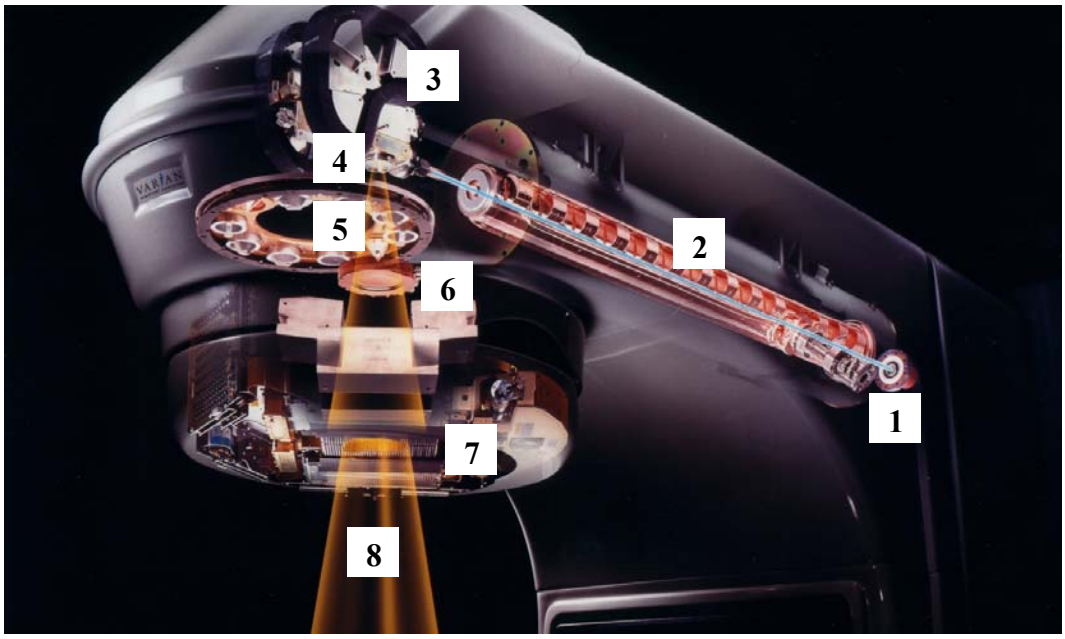


*Figure 2.5: Schematic of medical linear accelerator. The following components are labeled: 1) electron gun, 2) accelerating waveguide, 3) bending magnet, 4) tungsten target, 5) flattening filter, 6) monitor ion chamber, 7) beam shaping device: multi-leaf collimator (MLC), 8) x-ray beam (image from varian.mediaroom.com).*

### 2.3.2 Multi-Leaf Collimator (MLC)

The multi-leaf collimator is a beam shaping device used to collimate the x-ray beam produced by a linac. The MLC is composed of two banks of tungsten leaves, where each leaf can be individually moved in and out of the x-ray field (Figure 2.6). The Varian MLC (Varian Millenium 120 Model) is composed of 60 leaf pairs. The central 40 leaf pairs are 0.5 cm wide and the peripheral 20 leaf pairs (10 pairs on each side) are 1.0 cm wide. The leaves are 6.0 cm thick parallel to the beam direction.

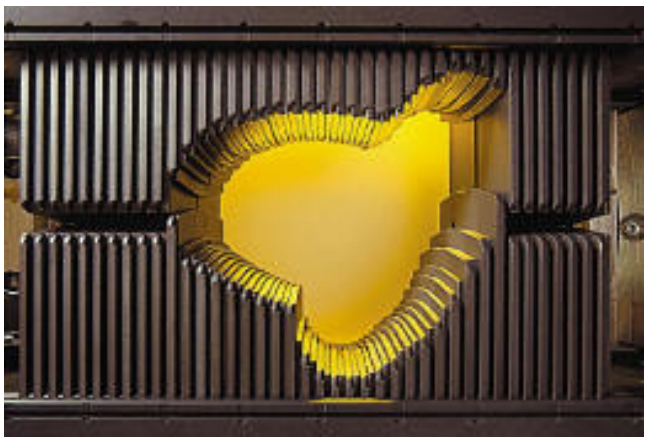

*Figure 2.6: A Multi-Leaf Collimator: a beam shaping device used to collimate the x-ray beam produced by a linac (image from varian.mediaroom.com).*

The MLC device can be used to shape the radiation beam to a desired shape by individually controlling the positions of each leaf of the MLC. Of course, the field shapes are inherently limited by the resolution (width) of the MLC leaves. In addition to being used as a beam shaping device, the MLC can also be used to modulate the beam intensity (fluence map) across the field. This is illustrated in Figure 2.7. Consider an intensity modulated fluence map consisting of three different intensity levels (1, 2 and 3), as shown in Figure 2.7. This intensity modulated fluence map can be produced by superimposing three fields (apertures) of uniform intensity. Each of these uniform fields can be delivered separately by the MLC by appropriately adjusting the leaf positions. Thus, by superimposing a number of uniform intensity fields an intensity modulated fluence map can be produced. This is the basic concept behind IMRT.

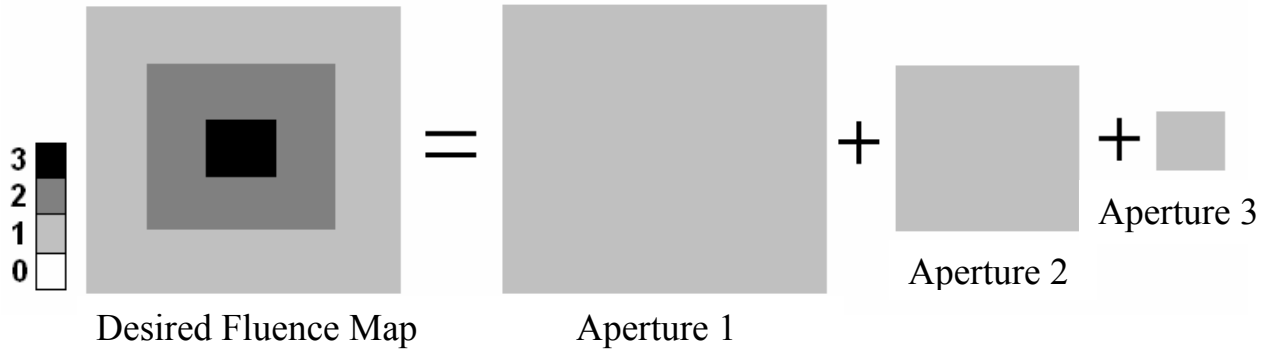


*Figure 2.7: An intensity modulated fluence map consisting of three different intensities (1, 2 and 3) can be produced by superimposing three fields (apertures) of uniform intensity. Each of these uniform fields can be delivered separately by the MLC by appropriately adjusting the leaf positions.*

## 2.3.3 X-ray Treatment Techniques

During a typical x-ray treatment, the patient lies on the treatment couch while the linac head rotates around the patient and the x-ray beam irradiates cancer cells. The linac head usually rotates in a coplanar fashion around the long axis of the treatment couch. The linac head is designed to rotate around a single point in space, called the isocentre. The beams from different linac head positions all intersect at the isocentre and therefore the patient is usually positioned such that the isocentre is placed close to the centre of the PTV. The isocentre is located 100 cm from the radiation source. Depending on the method of delivering radiation to the patient, x-ray treatment techniques can be divided into four main types:

### 3D Conformal Radiation Therapy (3D CRT)

In 3D CRT, radiation is delivered from several fixed linac head positions (beam angles). The number of beam angles is typically between 2 and 7. At each beam angle the MLC is used to shape the radiation field such that it conforms to the beam's eye view (BEV) of the PTV, as illustrated in Figure 2.8. The radiation beams have uniform intensity across the field. In some cases, the field intensity is modulated by placing simple devices, such as wedges or compensators, in the path of the radiation beam. The radiation beam intensities (weights) are obtained through a process known as forward planning.

In forward planning, the beams are manually assigned weights followed by a calculation of the 3D dose distribution. After qualitative review of the dose distribution by the treatment planner, plan improvement is performed by manually modifying the original conditions (e.g. changing the beam weights, adding wedges or compensators, adding another beam) to improve the PTV dose coverage and/or decrease the dose to OARs. This forward planning process is repeated until a satisfactory plan is generated.

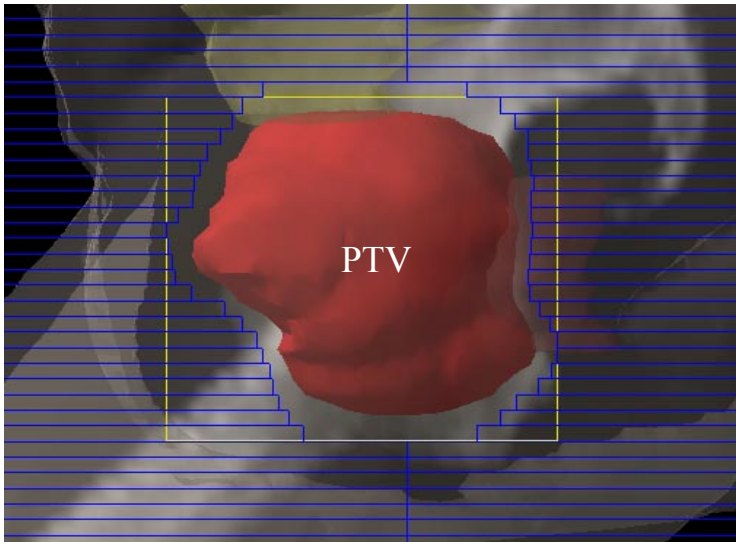


*Figure 2.8: In 3D Conformal Radiation Therapy, the MLC is used to shape the radiation field such that it conforms to the beam's eye view (BEV) of the PTV. The extra margin around the PTV is used to ensure that an electronic equilibrium (Section 1.4.3) is established throughout the whole PTV volume. In the regions near the radiation field edge there is a lateral electronic disequilibrium, caused by a lateral electron scatter. Lateral electronic disequilibrium causes the dose near the field edge to decrease due to the loss of the electron tracks.*

### 3D Conformal Arc Therapy (3D CAT)

3D CAT is similar to 3D CRT, except the radiation is delivered in a continuous arc around the patient. As the linac head rotates around the patient the radiation beam is continuously delivering radiation. This is different from the 3D CRT technique, in which the beam is turned off while the linac head rotates between treatment positions. The shape of the radiation field is continuously conformed to the BEV of the PTV, as the linac head moves through the arc. The radiation field

shapes are conformed to the BEV of the PTV at several equidistant beam angles (e.g. $10^{o}$) and interpolated between them to obtain continuously varying treatment field shape throughout the arc. The field weights throughout the arc are manually adjusted (forward planning) until a satisfactory plan is generated.

**Intensity Modulated Radiation Therapy (IMRT)**

In IMRT, radiation is delivered from several fixed linac head positions (beam angles), similar to 3D CRT. The number of beam angles is typically between 5 and 9. At each beam angle the MLC is used to deliver intensity modulated radiation field (as illustrated in Figure 2.7). The intensity profiles across each beam (fluence map) are obtained by a process known as inverse planning (explained in Section 2.4.1).

**Intensity Modulated Arc Therapy (IMAT)**

In IMAT, intensity modulated radiation fields are delivered in a continuous arc around the patient. Different methods of delivering IMAT have been proposed. In some methods radiation is delivered in a single arc around the patient and some require multiple overlapping arcs to achieve intensity modulated radiation fields. As in IMRT, the intensity profiles throughout the arc(s) are obtained using an inverse planning process.

# 2.4 Intensity Modulated Radiation Therapy (IMRT)

## 2.4.1 Inverse Planning

Inverse planning is the central concept behind IMRT and IMAT techniques. Inverse planning can be regarded as the opposite of the conventional forward planning process, used to create 3D CRT and 3D CAT treatment plans. In inverse planning the focus is on the desired outcome (e.g achieving specific target coverage) rather than how the outcome is going to be achieved. The user (planner) specifies the treatment goals and the optimization system (computer) then adjusts the intensity profiles across each beam (fluence map) iteratively in an attempt to achieve the

desired outcome. There are two main components involved in IMRT inverse planning optimization: 1) the treatment planning goals, such as PTV coverage and sparing of OARs, are used to construct a mathematical objective function, 2) an optimization algorithm, guided by the objective function, is used to generate intensity profiles across each beam.

## 2.4.2 Objective Function

The first step in IMRT inverse treatment planning is to construct a mathematical objective function, which will be used to guide the optimization process. The objective function is constructed based on the treatment planning goals. As described in Section 2.2.2, treatment planning goals are usually specified in terms of the dose-volume constraints placed on the PTV and the surrounding OARs. For example, two dose-volume constraints placed on the PTV and two dose-volume constraints placed on the surrounding OAR are displayed on a cumulative DVH plot in Figure 2.4a. The goal is to deliver dose between $D_{min}$=$D_p$-5% and $D_{max}$=$D_p$+7% to the entire PTV, where $D_p$ is the prescribed dose, while keeping dose given to volumes $V_1$ and $V_2$ of the OAR below $D_1$ and $D_2$, respectively. A quadratic objective function[1] is commonly used to implement these treatment planning goals. The quadratic objective function can be divided into different components: one for the PTV and one for each OAR.

The objective function component for the PTV can be expressed as:

$$F_{PTV} = \frac{w_{PTV}^{\min}}{N_{PTV}} \sum_{i=1}^{N_{PTV}} \left(D_i - D_{\min}\right)^2 H\left(D_{\min} - D_i\right) + \frac{w_{PTV}^{\max}}{N_{PTV}} \sum_{i=1}^{N_{PTV}} \left(D_i - D_{\max}\right)^2 H\left(D_i - D_{\max}\right) \quad (2.2)$$

Where:

$D_{min} = D_p\text{-}5\%$ and $D_{max} = D_p+7\%$, where $D_p$ is the prescribed dose to the entire PTV

$N_{PTV}$ = number of voxels in the PTV

$w_{PTV}^{\min}$ and $w_{PTV}^{\max}$ = weight factors (priorities) for satisfying $D_{min}$ and $D_{max}$ constraints

$H(x)$ = a step function, given by

$$H(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (2.3)$$

The first term of $F_{PTV}$ penalizes those PTV voxels, $D_i$, which are receiving dose less than $D_{min}$. The penalty is proportional to the square of the difference between $D_i$ and $D_{min}$, hence the name quadratic objective function. The $w_{PTV}^{\min}$ factor is used to control the importance of underdosing the PTV and is usually assigned a relatively high weight. In a similar fashion, the second term of $F_{PTV}$ penalizes those PTV voxels which are receiving dose higher that $D_{max}$.

For the OAR, the dose-volume constraints are incorporated using the method described by Bortfeld et al.[17]. Consider one of the dose-volume constraints specified for the OAR: the dose given to volumes $V_1$ should be less than $D_1$. This dose-volume constraint is displayed on a cumulative DVH plot in Figure 2.9a, together with a dose distribution which violates this dose-volume constraint. The shaded area on the cumulative DVH represents those voxels which violate the constraint. To selectively penalize only these voxels another dose D* is defined, as shown in Figure 2.9b. The following quadratic objective function can be used to penalize the OAR voxels receiving the dose between $D_1$ and D* (shaded region):

$$F_{OAR} = \frac{w_{OAR}}{N_{OAR}} \sum_{i=1}^{N_{OAR}} \left(D_i - D_1\right)^2 \cdot H\left(D_i - D_1\right) \cdot H\left(D^* - D_i\right) \tag{2.4}$$

Where:

$N_{OAR}$ = number of voxels in the OAR

$w_{OAR}$ = weight factor for the specific dose-volume constraint for the OAR

$H(x)$ = a step function, given by Equation 2.3.

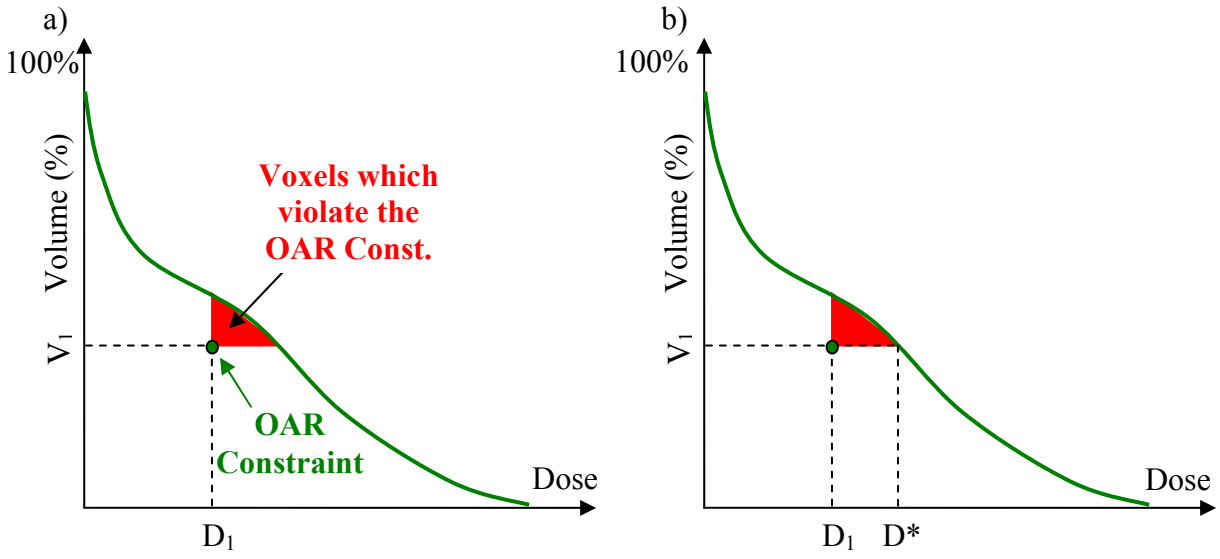


*Figure 2.9: a) A single OAR dose-volume constraint is displayed on a cumulative DVH plot, together with a dose distribution which violates this dose-volume constraint. The shaded area represents those voxels which violate the constraint. b) To selectively penalize only these voxels in the objective function, another dose D* is defined at the intersection of the volume constraint $V_1$ and the DVH curve.*

The objective function is usually scaled with the number of voxels in the OAR, $N_{OAR}$, to avoid having large organs dominate the optimization process. For multiple dose-volume constraints placed on a single OAR, the objective function component for that OAR would be a linear sum of terms similar to Equation 2.4. For each OAR, an unlimited number of dose-volume constraints can be implemented in the objective function. Finally, multiple OARs can be incorporated into the objective function by summing over all the terms for each OAR. Putting all components for PTV and OARs together, the total objective function becomes:

$$F_{TOTAL} = F_{PTV} + \sum^{n_{OAR}} \sum^{n_{constraint}} F_{OAR} \tag{2.5}$$

Where $n_{OAR}$ and $n_{constraint}$ represent the number of OARs and the number of dose-volume constraints specified for individual OARs, respectively.

Looking at Equations 2.2 and 2.4, two important properties can be observed. First, the objective function is constructed such that its value is always positive and should be minimized during the optimization process. The objective value of zero indicates that all of the treatment planning goals are completely satisfied. The optimizations methods will be discussed in Section 2.4.4. Second, the objective function involves knowledge of dose received by each voxel of the PTV and OARs from the multiple radiation beams. The dose calculation algorithm is discussed in the next section.

## 2.4.3 Pencil Beam Kernel (PBK) Dose Calculation Algorithm

To calculate the objective function value, the dose given to each voxel of the PTV and OARs from multiple radiation beams needs to be calculated, as indicated in Equations 2.2 and 2.4. In IMRT, the radiation beams are intensity modulated, with spatially varying intensity profiles across each field. To calculate the dose contribution from each field, the first step is to subdivide the radiation field into a finite number of small "beamlets", with typical dimensions of 0.25 cm x 0.25 cm (Figure 2.10). Each beamlet is regarded as a subfield with uniform intensity. By calculating the dose contribution from each individual beamlet and then summing over all the beamlets in all the fields the total dose can be calculated for each voxel of PTV and OAR.

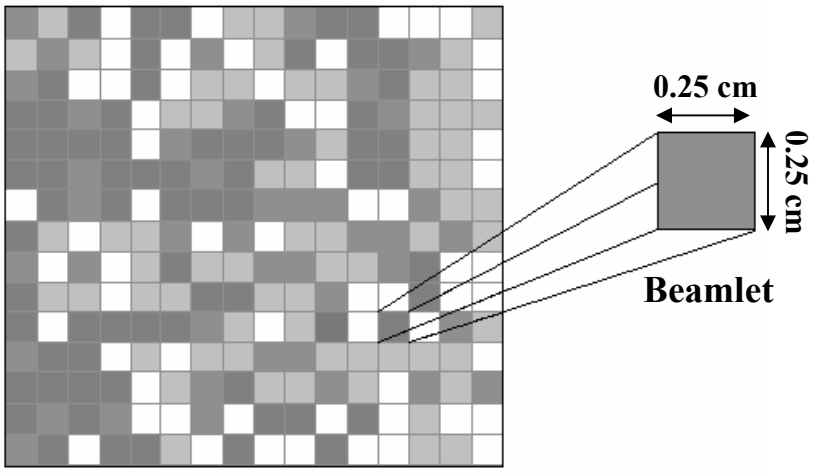


*Figure 2.10: In IMRT each intensity modulated radiation field is subdivided into a finite number of small "beamlets", with typical dimensions of 0.25 cm x 0.25 cm. Each beamlet is regarded as a subfield with uniform intensity (shade of gray).*

To calculate the dose contribution from each beamlet of radiation, a commonly used algorithm is the Pencil Beam Kernel (PBK) algorithm[18-20]. With the PBK, the dose delivered by each beamlet of radiation is modeled using a pencil beam dose kernel *K (x, y, d)*. The pencil beam dose kernel is defined at several discrete depths (e.g. *d* = 1cm, 2cm, 3cm, etc...) and is interpolated for the intermediate depths. At each depth, the pencil beam dose kernel can be approximated by a 2D Gaussian function:

$$K(x,y) = \frac{1}{2\pi\sigma^2}\exp\left(\frac{-(x^2+y^2)}{2\sigma^2}\right) \tag{2.6}$$

Where $\sigma$ is depth dependant and is derived from experimental data.

For intensity modulated radiation fields, such as the one shown in Figure 2.10, the total dose distribution at a particular depth can be calculated by superposition of all the individual pencil beam dose kernels at that depth. However, if spatial invariance of the pencil beam dose kernel (in x and y dimensions) is assumed, the superposition can be replaced with convolution. In that case, the total dose distribution *D(x, y, d)* at a particular depth *d* can be calculated by:

$$D(x,y,d) = \frac{(f+d_{ref})^2}{(f+d)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(x',y')P(x',y',d)K(x-x',y-y',d)dx'dy' \tag{2.7}$$

Where:

*f* = Source-to-Surface Distance (SSD)

$d_{ref}$ = reference depth defined by user

*F (x', y')* = idealized photon fluence field

*P (x', y', d)* = primary beam off-axis intensity profile. The actual photon fluence is obtained by multiplying the idealized photon fluence, *F (x', y'),* by a depth dependant primary beam off-axis intensity profile, *P (x', y', d)*. The *P (x', y', d)* factor accounts for the presence of the flattening filter and its effect on the photon beam's energy spectrum.

The inverse square law factor in front of the integral sign in Equation 2.7 accounts for beam divergence. The major disadvantage of the PBK dose calculation algorithm is its inability to model lateral electronic disequilibrium in heterogeneous regions. For heterogeneous regions,

the calculated dose distribution has to be corrected using one of the tissue inhomogeneity correction methods (e.g. Modified Batho correction method[21, 22]).

## 2.4.4 Optimization Methods for Minimizing Objective Function

Equation 2.7 relates the photon fluence, *F (x', y'),* across each radiation field (beam) to the 3D dose distribution in the patient, which is required to calculate the objective function value, via Equations 2.2 to 2.5. The next step is to optimize the photon fluence across each radiation beam such that the resulting 3D dose distribution in the patient will minimize (optimize) the objective function. Two common optimization techniques used in IMRT treatment planning are: Fluence-Based Optimization and Direct Aperture Optimization.

### Fluence-Based Optimization

Fluence-based optimization is the most commonly used technique in IMRT to optimize intensity (fluence) maps across radiation fields. Prior to the start of the optimization, each radiation field is segmented into a finite number of small "beamlets", with typical dimensions of 0.25 cm x 0.25 cm, as shown in Figure 2.10. Initially, each beamlet is assigned an intensity level, $x_j$, which can be written in vector form as $\boldsymbol{x} = \{x_j\}$, j = 1…J, where J is the total number of beamlets from all the radiation fields. During the optimization procedure, the elements of the intensity vector $\boldsymbol{x}$ are iteratively modified based on the gradient of the objective function. At each iteration, k, the gradient of the objective function $F_{TOTAL}(\boldsymbol{x}^{(k)})$ is calculated and a vector $\mathbf{h}^{(k)}$ is defined as:

$$\vec{h}^{(k)} = -\nabla F_{TOTAL}(\vec{x}^{(k)}) \tag{2.8}$$

Based on the direction of $\mathbf{h}^{(k)}$ vector, a new $\boldsymbol{x}^{(k+1)}$ intensity vector is defined as:

$$\vec{x}^{(k+1)} = \vec{x}^{(k)} + t \cdot \vec{h}^{(k)} \tag{2.9}$$

where *t* is the step size at each iteration. For the first iteration, k=1, the initial intensity vector **x** is used to calculate the gradient of the objective function in Equation 2.8 (i.e. $\boldsymbol{x}^{(1)} = \boldsymbol{x}$). During subsequent iterations, the intensity vector is iteratively modified according to Equation 2.9.

This optimization technique is called the Gradient Method or Steepest Descent[23]. The basic idea behind it is that by taking steps proportional to the negative of the gradient of the function at each iteration, one can find the local minimum of the function. This is illustrated in Figure 2.11. As shown in Figure 2.11, the Gradient Method will find only the closest local minimum of the objective function, which might not be the global minimum.

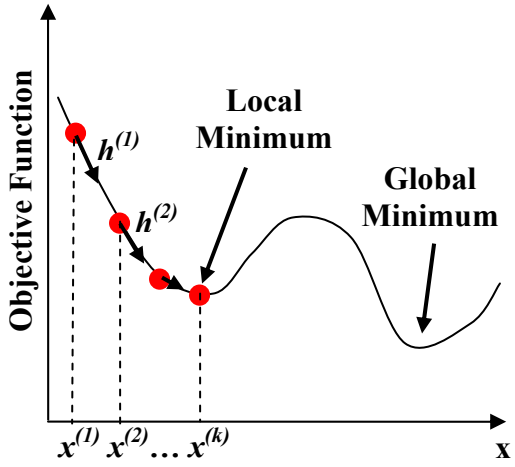


*Figure 2.11: In Gradient Method optimization one can find the local minimum of the function by taking steps proportional to the negative of the gradient of the function at each iteration.*

The Gradient Method is conceptually simple to understand but its drawback is that its convergence is relatively slow, even for well-behaved objective functions. The reason for slow convergence is that the next direction of optimization may partially spoil the optimization along the previous direction, known as the "zig-zagging" problem. The zig-zagging problem is avoided by the Conjugate Gradient Method, which produces a sequence of non-interfering directions by combining the gradient of the objective function with the previous direction of optimization[24]. Thus, fewer iterations are required compared with the Gradient Method. In the Conjugate Gradient Method, vector $\mathbf{h}^{(k+1)}$ is defined as:

$$\vec{h}^{(k+1)} = -\nabla F_{TOTAL}(\vec{x}^{(k+1)}) + \frac{[\nabla F_{TOTAL}(\vec{x}^{(k+1)}) - \nabla F_{TOTAL}(\vec{x}^{(k)})] \cdot \nabla F_{TOTAL}(\vec{x}^{(k+1)})}{\nabla F_{TOTAL}(\vec{x}^{(k)}) \cdot \nabla F_{TOTAL}(\vec{x}^{(k)})} \cdot \vec{h}^{(k)} \quad (2.10)$$

The recurrence relation of Equation 2.10 is initialized by setting:

$$\vec{h}^{(1)} = -\nabla F_{TOTAL}(\vec{x}^{(1)}) \quad (2.11)$$

The Conjugate Gradient Method is commonly used for fluence-based IMRT optimization in commercial treatment planning systems.

After the beamlets weights are optimized, the second step is to translate these optimal intensity maps into a series of deliverable MLC shapes (apertures) of uniform intensity. This step is referred to as the leaf sequencing step and is performed using a leaf motion calculator. The leaf motion calculator attempts to produce a series of MLC apertures with different weights, such that a desired intensity map is generated by superimposing these apertures. Often the optimal intensity maps are quite complex, as shown in Figure 2.12, and are not physically deliverable by the MLC. In this case, the closest match is offered by the leaf motion calculator. These deliverable fluence maps might be significantly different from the optimal fluence maps and the resulting dose distribution might be significantly different from the optimized one. This is one of the main disadvantages of Fluence-Based Optimization techniques. Another disadvantage of having the leaf sequencing step disconnected from the optimization step is the inability to include and account for MLC characteristics, such as interleaf leakage and radiation transmitted through the leaves, during the optimization. Finally, the resulting MLC leaf sequences often contain a large number of small apertures resulting in a large number of MUs required to be delivered by the linac (i.e. long "beam on" times). This translates into longer treatment times and decreased patient throughput, as well as increased leakage and scatter radiation from various components in the linac head. Increase in leakage and scatter radiation results in higher total body irradiation, which has been linked to increased risk of inducing secondary cancers[25-27].

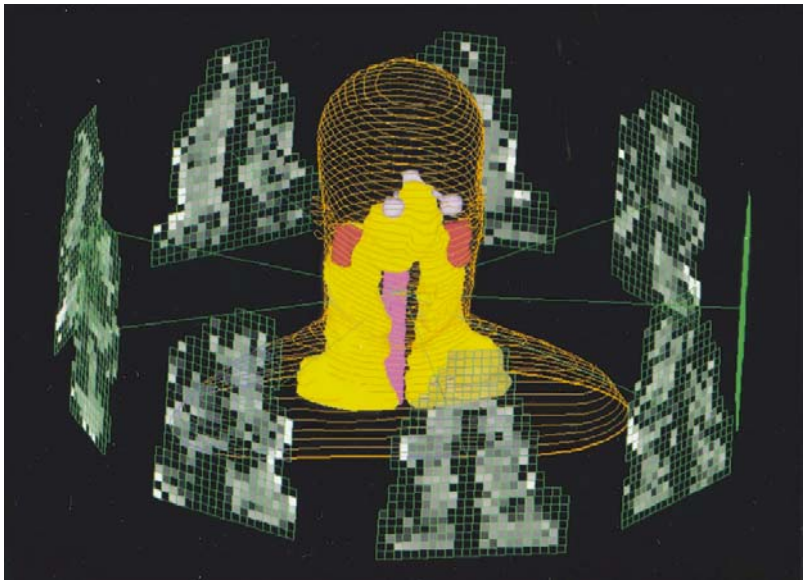

*Figure 2.12: a) Fluence-based IMRT optimization often produces complex fluence maps, which might not be physically deliverable by the MLC. The gray levels reflect the intensity value in the fluence maps (image from IMRT CWG[1]).*

## Direct Aperture Optimization (DAO)

DAO[28, 29] has been recently introduced as an alternative to fluence-based optimization in order to address some of its drawbacks. In DAO, the MLC leaf positions and aperture weights (number of MUs) are directly optimized. This is different from fluence-based optimization, in which beamlet weights are first separately optimized and then, in a second step, an attempt is made to translate these optimal intensity maps into a series of deliverable MLC shapes (apertures). In DAO, the leaf sequencing step is not required. It has been demonstrated that DAO is capable of producing high quality plans with a significant reduction in both the number of beam segments and the number of monitor units compared to traditional fluence-based optimization techniques[29-31]. This results in shorter treatment times and increased patient throughput, as well decreased leakage and scatter radiation. In DAO, delivery constraints imposed by the MLC are directly included in the optimization, ensuring that the derived optimal plan is physically deliverable by the MLC. One of the key features of the DAO technique is that the user specifies the number of apertures to be delivered as a constant in the optimization. Thus, the user is provided with greater control over the complexity of the treatment plan compared to fluence-based optimization. The following is a description of a DAO algorithm.

After specifying the number of beams and beam angles the user specifies the number of apertures per beam. Then number of apertures per beam is typically between 3 and 9, depending on the complexity of the case. During the optimization the number of apertures per beam is kept constant, but the shape and weight of each aperture is changed. Initially, all the aperture shapes usually conform to the BEV of the PTV and the aperture weights are all equal. Figure 2.13a displays one such aperture defined by the MLC. Usually, the aperture weights are initially set so that the mean dose to PTV is equal to the prescribed dose. The dose distribution and the objective function value (cost) are calculated for these initial aperture shapes and weights.

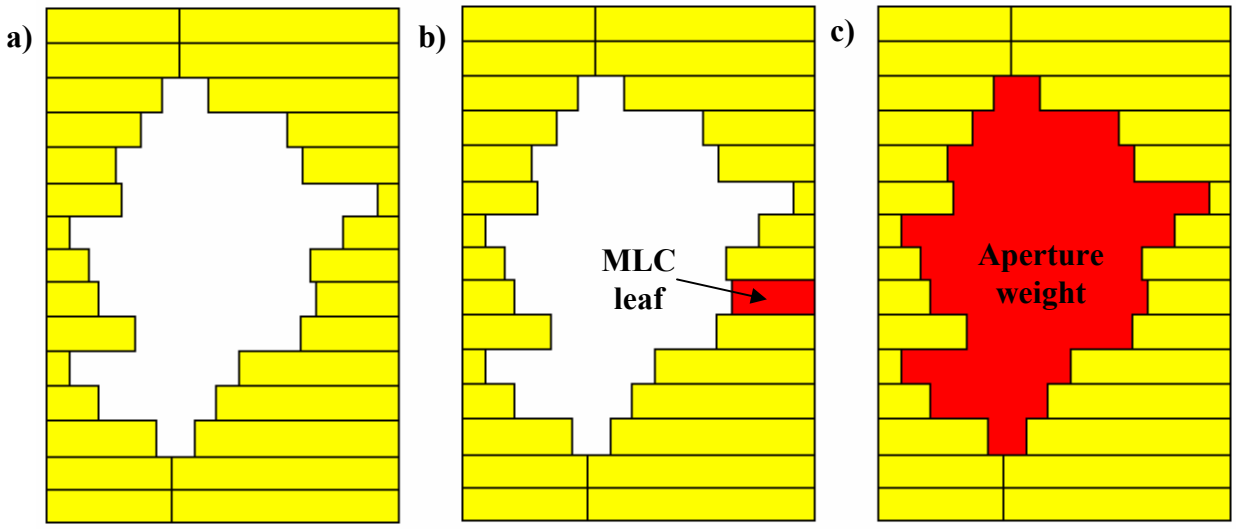


*Figure 2.13: a) One aperture defined by MLC leaves. A typical DAO treatment plan would have between 3 and 9 apertures per beam. For a treatment plan with 7 beams and 6 apertures per beam, the total number of apertures would be 42. During the optimization, the algorithm randomly selects one of the apertures. Once the aperture is selected, either b) a MLC leaf or c) aperture weight is randomly selected.*

The optimization is performed using a Simulated Annealing algorithm[32]. During the optimization the algorithm cycles through the variables to be optimized: aperture weights and MLC leaf positions for each aperture. The optimization starts with a random selection of an aperture. Once the aperture is selected, either a MLC leaf (Figure 2.13b) or the aperture weight (Figure 2.13c) is randomly selected and a random change is applied. The change is randomly

sampled from the pre-specified range for that variable. The maximum change decreases according to the following schedule:

$$M(n_{succ}) = \frac{M_o}{(1+n_{succ})^{R^{STEP}}} \tag{2.12}$$

If the selected variable is a leaf position then $M_o$ represents the initial maximum leaf step size. If the selected variable is an aperture weight then $M_o$ represents the initial maximum weight change. The parameter $R^{STEP}$ defines the rate at which the maximum change $M$ decreases as a function of the number of successful changes, $n_{succ}$. A successful change is a change which is accepted by the algorithm. The selected change must not violate the physical constraints of the MLC. Also, MLC leaves are not allowed to move outside the BEV of the PTV and negative segment weights are automatically rejected. Constraints on the minimum aperture size and on the minimum number of MU (aperture weight) for each aperture can also be added. This helps avoid inefficient segments. If the selected change violates any of these constraints, it is automatically rejected. If it meets all of the constraints the cost function is recalculated, with the selected change temporarily included. If the selected change results in a lower cost value, indicating an improvement towards achieving the planning goals, it is accepted. If the selected change results in an increase of the cost value the change is not automatically rejected, but is accepted with a probability given by the Metropolis criterion:

$$P = \exp\left(-\frac{\Delta f}{T}\right) \tag{2.13}$$

Where $\Delta f$ is the increase of the cost value and $T$ is the parameter corresponding to the "temperature" of the system. The probability of retaining a solution with a higher cost value decreases as the temperature of the system decreases. The temperature is reduced according to the following "cooling" schedule:

$$T(n_{succ}) = \frac{T_o}{(1+n_{succ})^{R^{COOLING}}} \tag{2.14}$$

Where $T_o$ is the initial system temperature, $R^{COOLING}$ defines the rate of cooling and $n_{succ}$ is the number of successful changes.

From Equations 2.13 it can be seen that when the system temperature $T$ is high, which is the case early on in the optimization, the changes that increase the cost value are allowed more frequently. As the optimization progresses, and the system temperature decreases, the probability of accepting such changes decreases. Therefore, the initial search space is relatively large, with high probability of accepting changes which result in increase of the cost value. This mechanism allows the optimization to escape from local minima (Figure 2.14). This is the inherent advantage of the simulated annealing method compared to the gradient method, which has no means of escaping from local minima. As the optimization progresses and the system temperature decreases, the search space is gradually reduced allowing the algorithm to converge to a final solution. The selection of the initial system temperature and the rate of cooling is very important in simulated annealing, since these parameters determine the algorithm's ability to escape from local minima, as well as speed of convergence to a final solution.

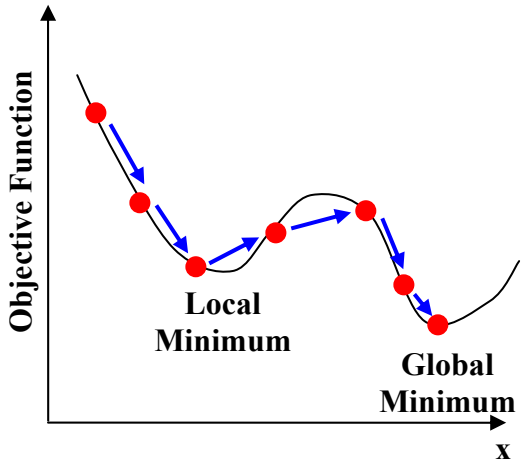


*Figure 2.14: In simulated annealing, steps which increase the objective function value are accepted with a probability given by the Metropolis criterion. This mechanism allows the optimization to escape from local minima in search of a global minimum.*

After the optimization is completed, the output is a set of deliverable apertures. The final apertures generally have different shapes and weights, as shown in Figure 2.15.

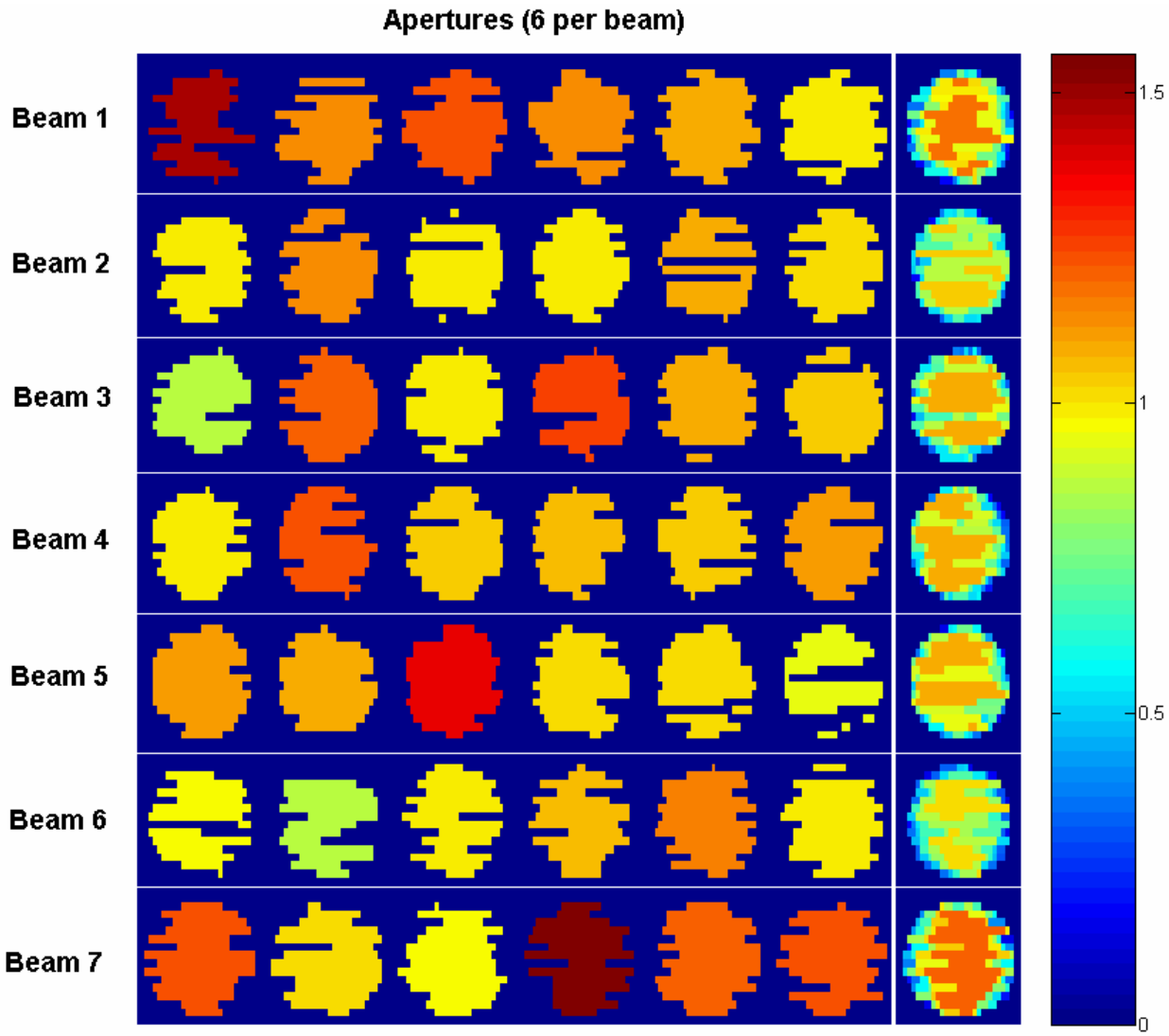


*Figure 2.15: A DAO treatment plan with 7 beams and 6 apertures per beam, created for a prostate case. The optimized aperture shapes and relative weights (color bar) are shown for each beam, together with the intensity modulated fluence map obtained by superimposing the individual apertures (last column). The intensities of the superimposed fluence maps are renormalized by the number of apertures per beam (i.e. 6).*

# 2.5 Adaptive Radiation Therapy

During the last two decades, the field of radiation therapy has witnessed some major developments and advances in treatment planning optimization and radiation delivery techniques. The introduction of the MLC and inverse planning techniques has allowed delivering optimized, intensity-modulated fluence maps across each radiation beam. The result is an unprecedented ability to precisely tailor dose distributions to the PTV volume and maximize sparing of adjacent OARs. With the highly advanced methods of treatment plan optimization and radiation delivery, it is the imprecision in clinical target definition that is currently the limiting factor in the radiation therapy treatment.

## 2.5.1 Conventional (non-adaptive) Radiation Therapy

Once the optimal treatment plan (e.g. IMRT plan) is created, precisely tailored for the individual patient anatomy, radiation is delivered to the patient in multiple fractions. Delivering radiation in multiple daily fractions has several radiobiological advantages over the single fraction. The number of fractions is treatment site specific and can greatly vary between treatment sites. For a typical prostate cancer case the prescribed dose to the PTV is delivered in 35 fractions over several weeks.

In conventional (non-adaptive) radiation therapy, radiation is delivered to the patient based on the original treatment plan, which remains the same for the whole treatment process. To account for possible anatomical variations during the course of the treatment, such as shifting/deforming of the CTV, a margin is placed around the CTV to obtain the PTV. The added margin accounts for the net effect of all the possible uncertainties and inaccuracies in order to ensure that the prescribed dose is actually absorbed in the CTV. The uncertainties and inaccuracies can arise from various factors such as:

- mechanical uncertainties of the equipment
- dosimetric uncertainties of the equipment
- uncertainties associated with patient imaging
- uncertainties in contouring (delineating) structures

However, the major sources of uncertainty arise from the inability to accurately know patient anatomy immediately prior and during the treatment, due to:

- uncertainties and variations in patient positioning
- patient motion during the treatment
- physiological movement and variations in size, shape and position of the CTV in-between (inter-fractional) and during (intra-fractional) the treatment fractions

In conventional radiation therapy, a margin added around the CTV has to account for all of these sources of uncertainties. Therefore, a relatively large margin has to be used. This imposes a limit on the ability to further tailor the dose distribution to the CTV and minimize dose received by the surrounding OARs[33].

## 2.5.2 Integrated Treatment / Imaging Systems

The introduction of various integrated treatment/imaging systems has allowed three-dimensional images to be acquired immediately before treatment fractions[34]. Systems such as Linac/CT-on-rails[35, 36] and Linac/Cone Beam CT (CBCT)[37-41] allow imaging of the patient in the treatment position (Figure 2.16). In the Linac/CT-on-rails system, the CT scanner travels on 2 parallel rails in the treatment room. In the Linac/CBCT system, the CBCT is mounted directly on the linac gantry and rotates together with the linac around a common axis of rotation. CBCT operates on the same principle as conventional CT, with the exception that an entire volumetric image is acquired through a single rotation of the source and the detector. This is made possible by the use of a 2D flat-panel detector. The ability to obtain the patient anatomy prior to the treatment fraction has led to the introduction and development of ART. There are two main forms of ART: 1) off-line ART and 2) on-line ART.

**a)**  **b)** 

*Figure 2.16: Integrated treatment/imaging systems allow imaging the patient in treatment position: a) Linac/CT-on-rails system (image from medical.siemens.com), and b) Linac/Cone Beam CT (CBCT) system (image from varian.mediaroom.com).*

## 2.5.3 Off-line Adaptive Radiation Therapy

In off-line ART[42-44], the patient is imaged in the treatment position prior to the first couple of treatment fractions (e.g. first 5 out of 35 fractions). The patient-specific systematic and random variations are estimated based on the acquired images and the original treatment plan is re-optimized prior to the delivery of the next fraction (e.g. 6$^{th}$ fraction). The plan re-optimization/adaptation is performed off-line (in-between treatment fractions). The systematic errors are used to calculate the average amount by which the patient needs to be re-positioned to account for systematic errors associated with initial patient set-up and/or change in position of the CTV. The random errors are used to calculate a patient-specific margin that needs to be added around the CTV to obtain the PTV. The margin accounts for all the random errors associated with patient set-up and variations in size, shape and position of the CTV. Once the original plan is re-optimized (adapted), the adapted plan is used for all the subsequent treatment fractions. The main advantages of this approach are: 1) the ability to detect and correct for the systematic variations; and 2) the use of the customized patient-specific margin instead of a margin based on variations averaged from patient population. As a result, the margins added around the CTV to obtain the PTV are generally tighter compared to the conventional radiation therapy.

Another proposed approach to off-line ART is to image the patient in the treatment position prior to each treatment fraction[45, 46]. The treatment plan re-optimization/adaptation is performed off-line after each treatment fraction and the adapted plan is used for the subsequent treatment fraction. Any dosimetric errors caused by the anatomical deformation in fraction n are attempted to be corrected in the subsequent fraction n+1. This approach requires deformable registration[47-50] to track the dose delivered to deformed region of interest (ROI) between the fractions. The disadvantage of this approach is the fact that the plan adaptation always lags the radiation delivery by "one fraction" and attempts to correct for the previously introduced dosimetric errors. In some cases, the dosimetric errors introduced in fraction n can not be corrected for in fraction n +1. For example, incorrectly delivering high dose of radiation to a sensitive critical structure in fraction n can not be "undone" in the subsequent fraction n+1.

## 2.5.4 On-line Adaptive Radiation Therapy

In on-line ART[51-54], patient images are acquired immediately prior to each treatment fraction. The patient is imaged in a treatment position using an integrated treatment/imaging system, such as Linac/CT-on-rails or Linac/CBCT. The plan re-optimization is performed on-line, while the patient is in the treatment position. Based on the current (daily) patient images the following actions might be required: 1) the patient might need to be re-positioned, by translating the treatment couch, to account for improper initial positioning and/or change in position of the CTV; 2) the original treatment plan might need to be re-optimized, to account for non-rigid anatomical deformations such as variations in size and shape of the CTV and change of relative position of the CTV with respect to other OARs. Once the original treatment plan is re-optimized, radiation is delivered based on the adapted treatment plan.

As a result of knowing the exact patient anatomy immediately prior to each treatment fraction, many sources of uncertainty are eliminated and a much tighter margin can be used around the CTV to obtain the PTV. Therefore, the major advantage of on-line ART is being able to confidently use smaller margins around the CTV, compared to conventional radiation therapy and even off-line ART (Figure 2.17). This allows the dose distribution to be more closely tailored to the CTV, thus minimizing the dose to the surrounding OARs. Reducing dose to OARs results in a decrease of NTCP and allows for potential prescribed dose escalation. Escalating the

prescribed dose has been shown to increase TCP and is directly related to the success of the treatment[55-58].

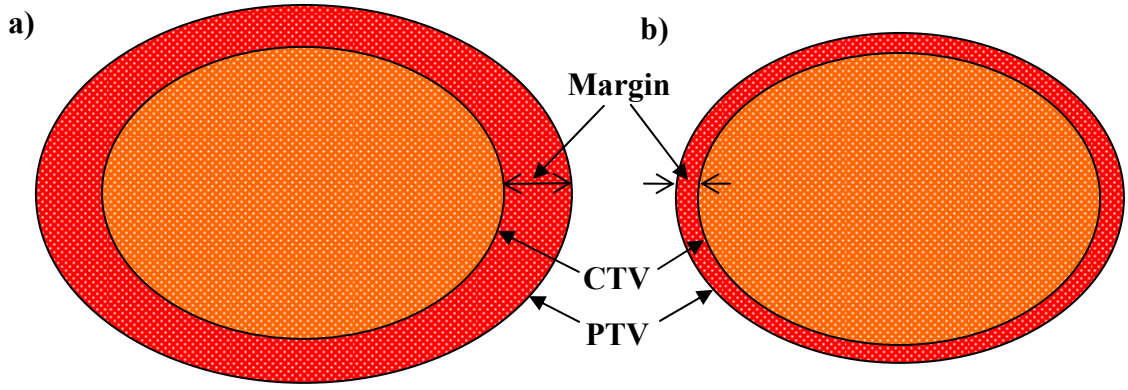


*Figure 2.17: Margin added around the CTV to obtain the PTV in: a) conventional radiation therapy and b) on-line ART.*

The on-line ART procedure can be divided into three stages:

1) **Imaging** the patient after they have been positioned on the treatment couch to obtain the current (daily) internal anatomy. This stage also includes analyzing the images and identifying/contouring the current CTV and OARs.

2) **Adapting** (re-optimizing) the original treatment plan based on the current anatomy.

3) **Delivering** the radiation to the patient based on the adapted treatment plan.

Currently, these three steps are performed sequentially, after the patient has been set-up in the treatment position (Figure 2.18).

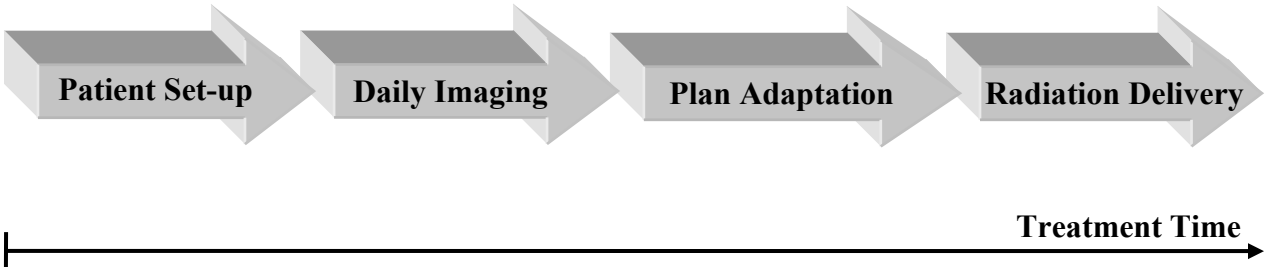


*Figure 2.18: Current on-line ART. Daily imaging, plan adaptation and radiation delivery are performed sequentially, after the patient has been set-up in the treatment position.*

There are several drawbacks with this current approach to on-line ART:

1) The overall treatment time is substantially extended, as compared to conventional (non-adaptive) radiation therapy. The addition of daily imaging and plan adaptation steps roughly doubles the overall treatment time. As a consequence, the patient throughput is reduced by half.

2) In the current clinical implementation of on-line ART, the plan adaptation is limited only to patient re-positioning due to the lack of methods for efficient plan re-optimization. The time required for plan re-optimization has to be on the order of a few minutes or less to be clinically acceptable.

3) Prolonged treatment times result in increased susceptibility to intra-fractional patient motion (Figure 2.19)[2]. This is a major drawback of current on-line ART, since intra-fractional patient motion increases the uncertainty in patient anatomy during the treatment. This uncertainty requires the margin around the CTV to be increased, undermining the main advantage of on-line ART – reduced margin size.

4) Inability to account for intra-fractional patient motion. Since daily imaging is performed only at the beginning of the treatment fraction, any intra-fractional patient motion can not be detected and accounted for with the current approach to on-line ART.

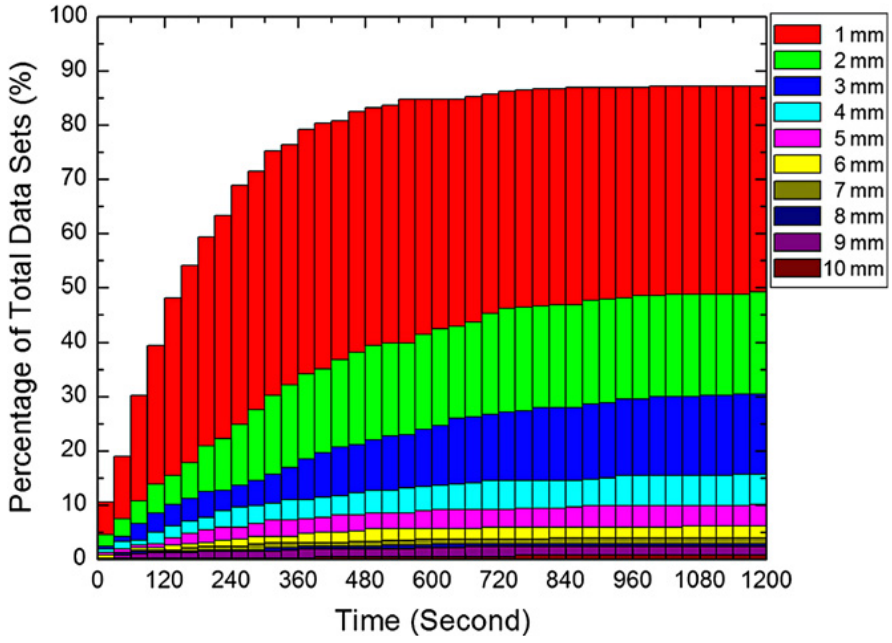


*Figure 2.19: Intra-fractional prostate motion as a function of treatment time. Figure shows a plot of the percentage for the prostate targets to move more than 1 mm, 2 mm, 3 mm, 4 mm, 5 mm, etc ., after 30 sec, 60 sec, 90 sec, 120 sec, etc. of treatment. The total number of data sets used in this study was 427[2]. (Courtesy of Xing et al.).*

## 2.6 Thesis Objectives

The primary objective of this thesis is to develop and implement a new approach to on-line ART, which has the potential to overcome the shortcomings and drawbacks of the current on-line ART procedure (described in the previous section). In this new approach to on-line ART, daily imaging, plan adaptation and radiation delivery are temporally integrated and performed concurrently (Figure 2.20d). The advantages of this approach are:

1) Reduction of treatment time compared to conventional on-line ART. Shorter treatment time increases patient throughput and decreases susceptibility to intra-fractional patient motion.

2) Ability to perform a complete plan re-optimization with minimal extension of treatment time.

3) Ability to detect and correct for intra-fractional patient motion.

This work was motivated by an initial study which compared four radiosurgery techniques: Circular Collimator Arcs, Dynamic Conformal Arcs, Static Conformal Fields and Intensity Modulated Radiosurgery (IMRS). The study investigated the correlation between the patient anatomy and the quality of the treatment plan produced by different radiosurgery techniques. Data from 18 patients was used in this study. This work is presented in Chapter 3 of this thesis. In this study, an inherent assumption was that the patient anatomy does not change during the entire treatment process. For radiosurgery treatments, this assumption is somewhat justified since all the radiation is delivered in only one fraction. However, for a majority of other radiation therapy treatments the radiation delivery is fractionated (30–40 fractions), and the assumption that the patient anatomy remains the same throughout the whole treatment process (several weeks) is no longer valid. This leads to the second part of this thesis, on-line ART (Chapters 4 to 6).

In Chapter 4, methods for accelerating plan adaptation (re-optimization) using DAO technique are investigated (Figure 2.20b). This work is the first investigation of using DAO for on-line ART. In Chapter 5, the concept of integrating accelerated plan adaptation and radiation delivery is investigated (Figure 2.20c). Finally, in Chapter 6, daily patient imaging, accelerated plan adaptation and radiation delivery are all temporally integrated and performed concurrently (Figure 2.20d). The ideas and concepts investigated in this thesis lead toward the ultimate goal of near real-time ART.

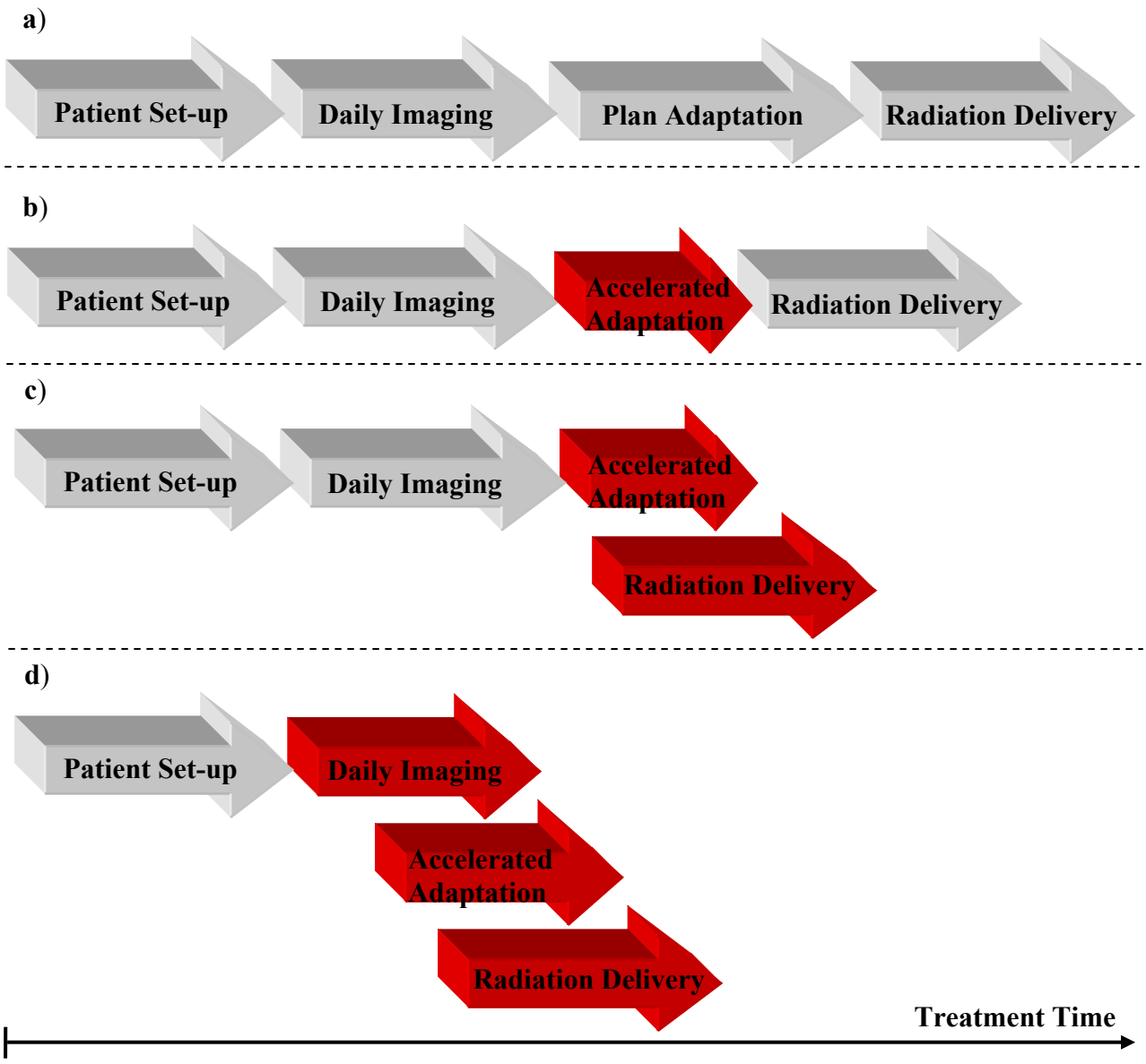


*Figure 2.20: a) Current approach to on-line ART. b) In Chapter 4, methods for accelerating plan adaptation (re-optimization) using DAO technique are investigated. c) In Chapter 5, the concept of integrating accelerated plan adaptation and radiation delivery is investigated. d) Finally, in Chapter 6, daily patient imaging, accelerated plan adaptation and radiation delivery are all temporally integrated and performed concurrently.*

# Chapter 3
# Correlation between Patient Anatomy and Treatment Plan Quality for different Stereotactic Radiosurgery Techniques

## 3.1 Introduction

### 3.1.1 Introduction to Stereotactic Radiosurgery (SRS)

SRS[59-61] is a treatment modality that uses accurately directed beams of radiation for the treatment of benign and malignant tumors, vascular lesions and functional disorders in the brain. The goal of SRS is to precisely deliver high radiation dose to the target volume while minimizing dose to the surrounding healthy tissue. Radiation is usually delivered in a single fraction. Compared to conventional radiation therapy, the target volumes involved in SRS are relatively small (few ccm). Delivering high radiation dose in a single fraction to a small target volume demands great accuracy in treatment delivery (within 1mm). This is achieved through very accurate target localization, patient immobilization and radiation delivery techniques. At the beginning of the SRS procedure, a carbon fibre or metallic head-ring is rigidly attached to the patient's skull using four sharp pins. The head ring serves to immobilize the head during both imaging and planning. Attached to the head ring is the localizer box, which defines a 3D coordinate system used for treatment planning and image registration between different imaging modalities (e.g. CT, MRI, angiography, etc.). Accurate radiation delivery is commonly achieved through the use of a micro multi-leaf collimator (μMLC). The μMLC has tungsten leaves which are 2-3 mm wide, compared to the 5 mm leaf width of a conventional MLC.

Arteriovenous malformation (AVM)[62-64] is a vascular disorder that can be treated with SRS. Normally, the arteries carry oxygen-rich blood at a relatively high pressure. Arteries divide

and sub-divide repeatedly, eventually forming a capillary bed. Blood moves through the capillaries, giving up oxygen and taking up waste products from the neighboring cells. Capillaries successively join together to form the veins that carry blood away at a relatively low pressure (Figure 3.1a). In an AVM, blood vessels form a direct connection between high pressure arteries and low pressure veins. This results in a collection of blood vessels with abnormal connections and no capillaries (Figure 3.1b). This collection can be very fragile and prone to rupturing and causing hemorrhage. AVMs can form anywhere in the body, but are of the greatest concern when they occur in the brain. Bleeding of an AVM in the brain may result in permanent disability and death. The treatment of choice for AVMs is surgery[65], but in the cases where they are inaccessible to surgery they may be treated using SRS.

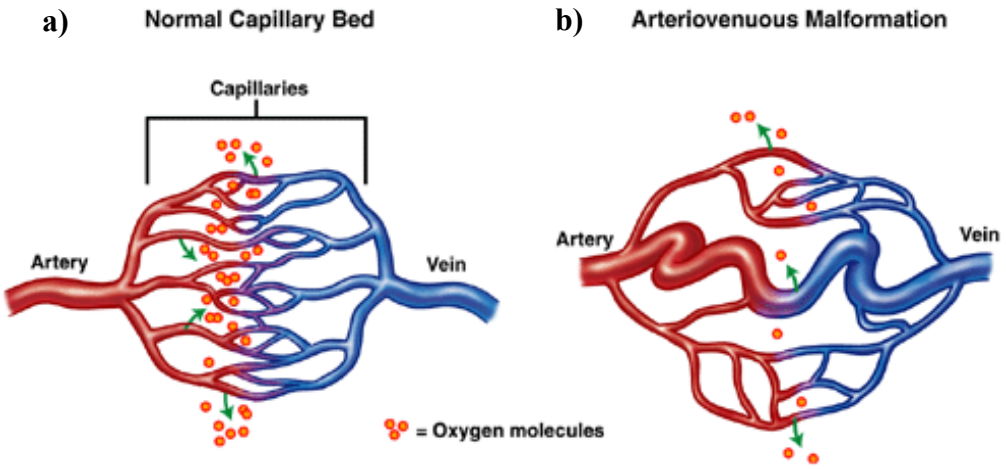


*Figure 3.1: a) Normally, the blood from arteries goes through capillaries before entering veins. b) In AVM, blood vessels form an abnormal direct connection between arteries and veins, bypassing capillaries (image from www.avmsupport.org).*

Digital Subtraction Angiography (DSA) is commonly used to localize AVMs[66]. DSA is a type of fluoroscopy technique used to visualize blood vessels. Images are produced by subtracting a 'pre-contrast' image from the images obtained with contrast medium (e.g. iodine-based) introduced into the blood vessels. 3D localization of AVMs can be achieved by acquiring sets of orthogonal DSA images (Figure 3.2). During SRS treatment planning, DSA images are registered to CT and MRI images, using fiducial markers on the localizer box. As a result, AVM contours are directly fused with CT and MRI images, which are used to identify/contour other surrounding OARs.

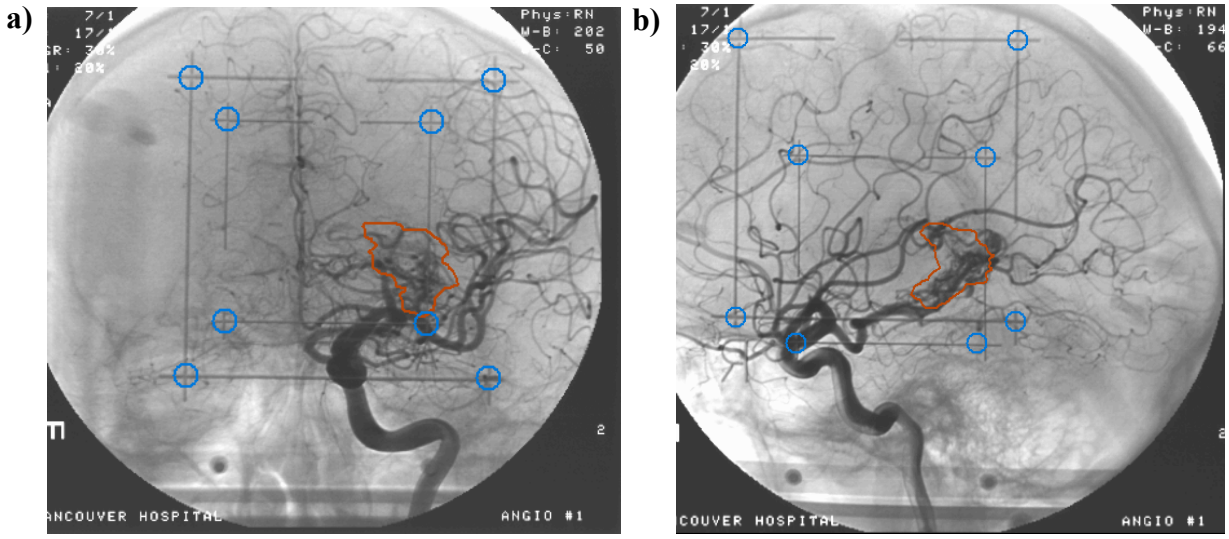


*Figure 3.2: Digital Subtraction Angiography (DSA) is commonly used to localize AVMs. 3D localization of AVMs can be achieved by acquiring set of orthogonal DSA images: a) Frontal image and b) Lateral Image. The red contour represents the AVM volume to be treated (PTV).*

There are several SRS techniques presently available for treatment of AVMs. One common procedure uses radiation produced by a large number of Cobalt-60 sources (Leksell Gamma Knife[67]). Other techniques use radiation produced by a linac. Linac-based techniques include:

1) Dynamic Conformal Arcs[68]
2) Fixed Conformal Fields[69]
3) Intensity Modulated Radiosurgery (IMRS)[70]
4) Circular Collimator Arcs[71]

The first three techniques are equivalent to radiation therapy treatment techniques (3D CRT, 3D CAT and IMRT, respectively) discussed in Section 2.3.3. The main difference is that SRS techniques use the μMLC, instead of the MLC, to more precisely define radiation fields. Also, SRS arc techniques generally employ multiple non-coplanar arcs, with single or multiple isocentres, to deliver radiation, whereas radiation therapy arc techniques usually deliver radiation in a single coplanar arc. The fourth technique, Circular Collimator Arcs, is similar to Dynamic Conformal Arcs except circular collimators are used to define the treatment field instead of the μMLC. Circular collimators with a range of diameters are available, from a few millimetres to a

few centimetres, and the appropriate one is selected based on the size of the AVM. Dynamic conformal arcs and IMRS are generally regarded as superior to circular arcs and fixed conformal field techniques in terms of their ability to conform the dose distribution to the target volume. The extra degree of freedom provided by motion through the arc in dynamic conformal arcs and intensity modulating across the fields in IMRS offers the potential for sparing of sensitive organs and increasing the target dose conformity, especially for cases where organs at risk are in close proximity to the irregularly shaped targets. However, the increased complexity of these plans results in various shortcomings, such as longer treatment times and complicated treatment accuracy verifications. Furthermore, IMRS beam profiles are generally complex and delivering such beam profiles results in certain practical difficulties: They generally require larger number of segments and monitor units, which makes the overall treatment delivery less efficient; high gradients in the intensity profiles may result in plans that are sensitive to treatment uncertainties such as target movement and patient set-up. Also, the prolonged treatment time results in increased contribution of indirect radiation to the total delivered dose. As a result, more of the surrounding normal tissue may be exposed to the significant levels of radiation. Therefore, even though dynamic conformal arcs and IMRS are generally superior compared to circular arcs and fixed conformal field techniques in terms of their ability to conform the dose distribution to the target, there might be cases where the simpler method of dose delivery proves to be more beneficial. For instance, fixed conformal field techniques might be suitable for cases where the target is relatively distant from the critical structures, and circular arcs may still be optimal for very small (<~2 cm in maximum dimension) and nearly spherical lesions.

## 3.1.2 Study Objectives

There are various studies in the literature that compare different radiosurgery techniques[72-76]. The treatment plans produced by the competing techniques are generally evaluated by comparing the isodose distributions and DVHs. The plan quality is usually determined based on the conformity of the dose to the target volume, dose homogeneity throughout the target volume and sparing of the surrounding normal tissue. An ideal treatment plan would deliver perfectly homogenous dose, equal to the prescribed dose, to the target volume and deliver no dose to the surrounding normal tissue. The plan that most closely approaches these ideal conditions is considered to be superior to the other plans. However, these findings are often limited to the particular case,

patient or phantom, considered in the study. It would be useful to generalize this result and conclude that a particular treatment technique is always superior to the other techniques for a particular type of treatment. However, the situation is frequently complex, since there are many variables affecting the plan quality from patient to patient. These variables include size, shape complexity of the target, proximity of the critical structures to the target, overlap of the critical structures with the target, etc. Thus, even though one treatment technique might prove to be better in a particular patient, another technique might be more effective for the next patient. The situation gets more complicated and unpredictable as the number of competing plans and the number of variables between patients increase. Therefore, to successfully predict which treatment technique is superior on a patient-to-patient basis, a systematic approach is required.

This study is the first quantitative analysis of the correlation between the patient anatomy and the quality of the treatment plan produced by different radiosurgery techniques. Eighteen previously treated AVM patients were used as a sample group in this study. The patients exhibited a wide range of target size, shape and location. A Multiple Linear Regression model was used to establish the correlation between the patient anatomy and the plan quality for different treatment techniques. The model was employed to predict the plan quality of different radiosurgery techniques for treatment of AVMs based on the individual patient anatomy. The results were used to predetermine the optimal radiosurgery technique for treatment of AVMs on a patient-to-patient basis. This work was published in the International Journal of Radiation Oncology, Biology and Physics [77].

# 3.2 Methods and Materials

## 3.2.1 Treatment Planning

Eighteen AVM patients, previously treated with SRS in Vancouver BC Cancer Agency between February 1998 and March 2001, were used as a sample group in this study. For each patient, four different treatment plans were created using the following techniques: 1) Circular Collimator Arcs, 2) Dynamic Conformal Arcs, 3) Fixed Conformal Fields and 4) IMRS. All the plans were created with BrainSCAN (BrainLAB AG, Heimstetten, Germany) by a single experienced physicist to assure the consistency between the treatment plans and common dosimetric criteria (i.e. target coverage). The original plans used for treatment were not used in this study. Figure

3.3 displays axial, coronal and sagital dose distributions produced by the four techniques for one patient. The 80%, 50%, 30% and 10% isodose lines are shown, together with the PTV volume (red contour). Prescribed doses ranged from 16 to 25 Gy between the eighteen patients. For the circular collimator arc technique, the dose was prescribed to the 50% isodose line and for the other three techniques the dose was prescribed to the 80% isodose line. The required target coverage was that at least 99% of the target volume be inside the prescribed isodose surface, for all four techniques.

In each patient the following six critical structures were identified: right and left motor strip, optic pathway, speech (Broca's and Wernicke's areas), memory (hippocampal gyrus) and brainstem (posterior clinoid to foramen magnum). No specific dose-volume constraints were specified for the critical structures and normal tissue for the three forward planning techniques. For the IMRS technique, the objective of inverse planning was to reduce the maximum dose to the critical structures and normal tissue as much as possible, while maintaining adequate target coverage.

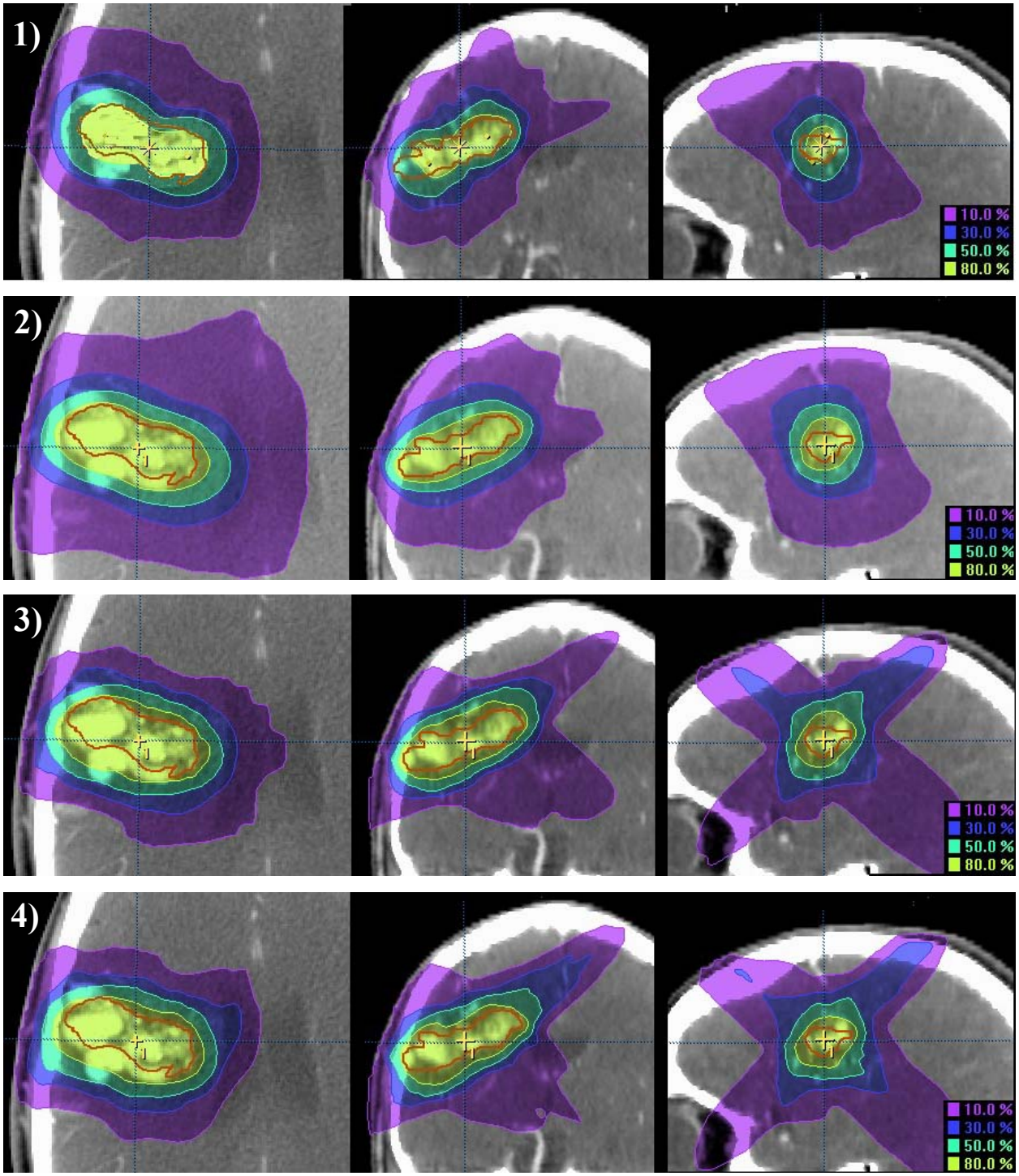


*Figure 3.3: Axial, coronal and sagital dose distributions produced by the four techniques for one patient: 1) Circular Collimator Arcs, 2) Dynamic Conformal Arcs, 3) Fixed Conformal Fields and 4) IMRS. The 80%, 50%, 30% and 10% isodose lines are shown, together with the PTV volume (red contour).*

### 3.2.2 Treatment Plan Quality (Dosimetric Parameters)

All of the 72 plans, 4 plans for each of the 18 patients, were analyzed and evaluated based on the DVHs for the target, normal tissue and critical structures. To evaluate the dose conformity to the target volume and the dose homogeneity in the target volume two parameters were used:

1) The ratio of the volume within the prescription isodose surface to the target volume (PITV)
2) The ratio of the maximum dose to the prescription dose (MDPD)

These two parameters are defined by the Radiation Therapy Oncology Group (RTOG) 93-05 protocol. To assess the doses to the normal tissue (all the healthy tissue outside the target volume) and the critical structures, volumes treated to 3, 6, 12, 18 and 24 Gy ($V_3$, $V_6$, $V_{12}$, $V_{18}$ and $V_{24}$) were used. Also, the maximum dose ($D_{max}$) to normal tissue and critical structures was obtained from DVHs.

### 3.2.3 Patient Anatomy (Anatomical Parameters)

To calculate the anatomical parameters for each patient, an algorithm was developed in MATLAB (Mathworks Inc., Natic, MA). The algorithm starts by generating 3D target and critical structure volumes by interpolating (shape-based interpolation[78]) contours defined on 2D CT and MR slices (Figure 3.4). The complete 3D anatomy for one of the patients is shown is Figure 3.5.

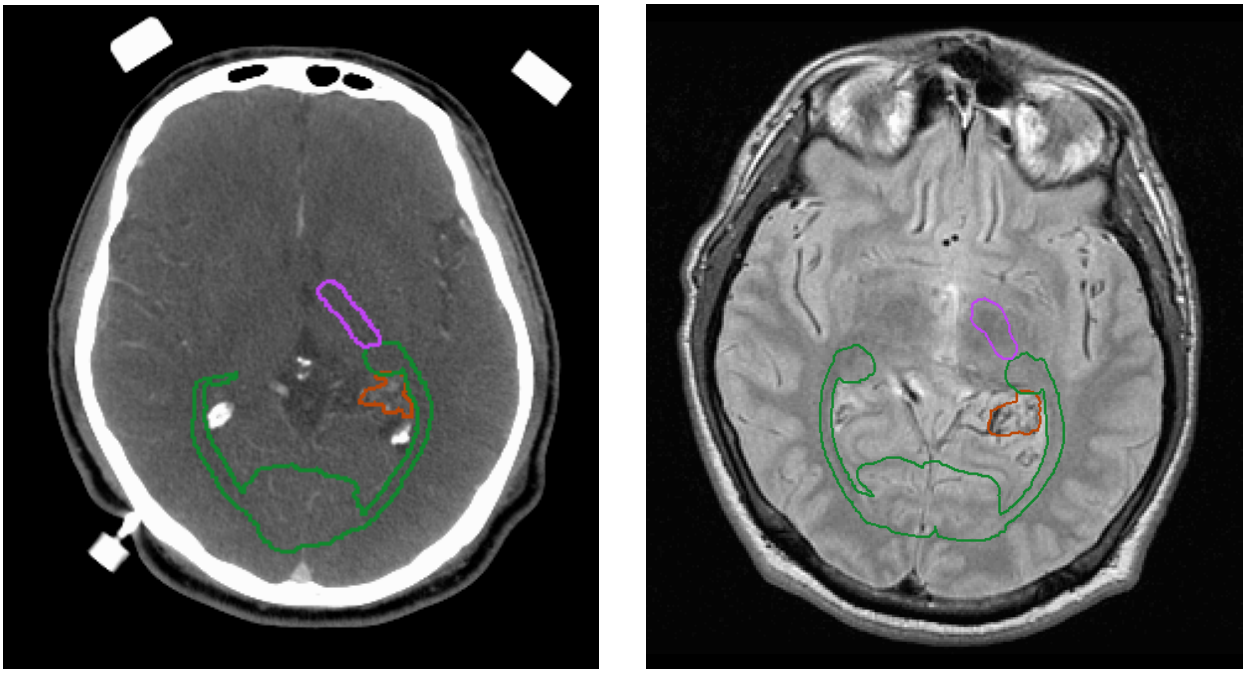

*Figure 3.4: The target (PTV) and two critical structures, the optic pathway and the left motor strip, are shown as red, green and purple contours, respectively on: 1) CT slice and 2) MR slice. The contours are slightly different on two slices since the slices do not exactly coincide, due to different slice thicknesses in CT and MR images.*

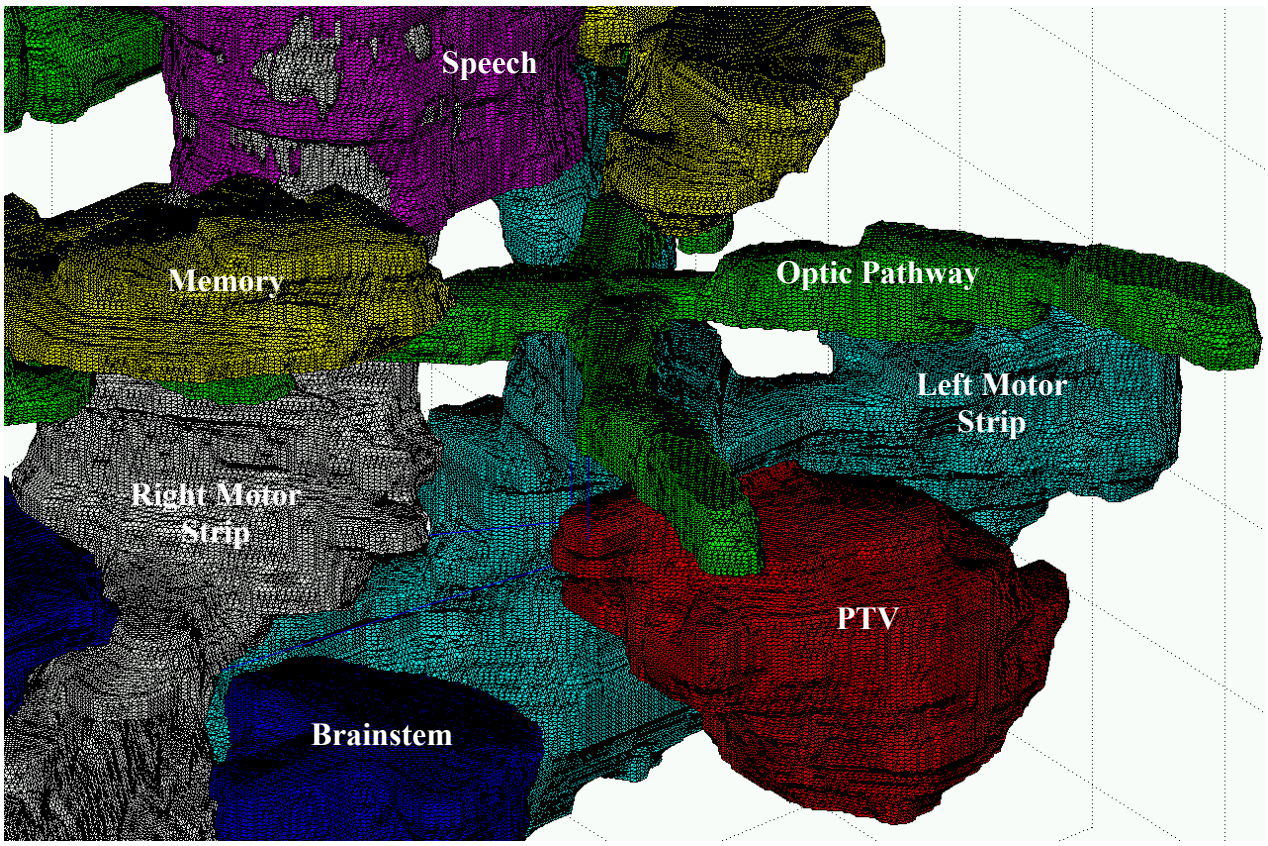


*Figure 3.5: Reconstructed 3D anatomy for one of the AVM patients.*

After the complete 3D patient anatomy was reconstructed in MATLAB, the algorithm calculated the following five parameters to characterize the size and shape of the target and each critical structure:

1) *Volume and surface area*

2) *Normalized surface area to volume ratio (SAVR)*:

$$SAVR = \frac{(SurfaceArea)^{3/2}}{6\pi^{1/2} Volume} \tag{3.1}$$

*SAVR* assesses the irregularity of the object. As the object's shape becomes more irregular, the value of *SAVR* increases. A sphere has a minimum value of *SAVR*, which is by definition equal to 1.

3) *Normalized convex hull surface area to volume ratio (convex hull SAVR)*: the convex hull of a set of points is defined as the smallest convex set containing all the points. The convex hull is insensitive to the local variations of the object's surface roughness. Therefore, the convex hull *SAVR* measures the overall shape of the object: the objects that have roughly spherical overall shape have convex hull *SAVR* values close to 1. As the object's overall structure deviates from that of a sphere, the value of the convex hull *SAVR* increases.

4) *Concave volume*: Concave volume is defined to be the difference between the convex hull volume and the actual volume. It quantifies the concave regions of a particular object.

5) *Percentage of Concave volume*: the ratio of the concave volume to the convex hull volume of the object. This parameter quantifies the relative amount of the concave regions.

Figure 3.6 shows the distribution of *Target SAVR*, *Target convex hull SAVR* and *Target Percentage of Concave volume* for the eighteen patients used in this study. For each parameter a target shape corresponding to the minimum and the maximum value is also shown. The algorithm also calculated the positions of all the critical structures with respect to the target and the amount of overlap between the target and each critical structure. The following parameters were used to characterize the location of the target with respect to the critical structures:

6) *Minimum distance*: the minimum distance between the target and each of the critical structures.

7) *Mean distance*: the mean distance is defined as the average of the minimum distances between the target and every pixel on the critical structure's surface.

8) *Overlap volume*: the volume of the overlap between the target and each critical structure.

9) *Percentage of Overlap volume*: the ratio of the overlap volume to the actual volume of the object.

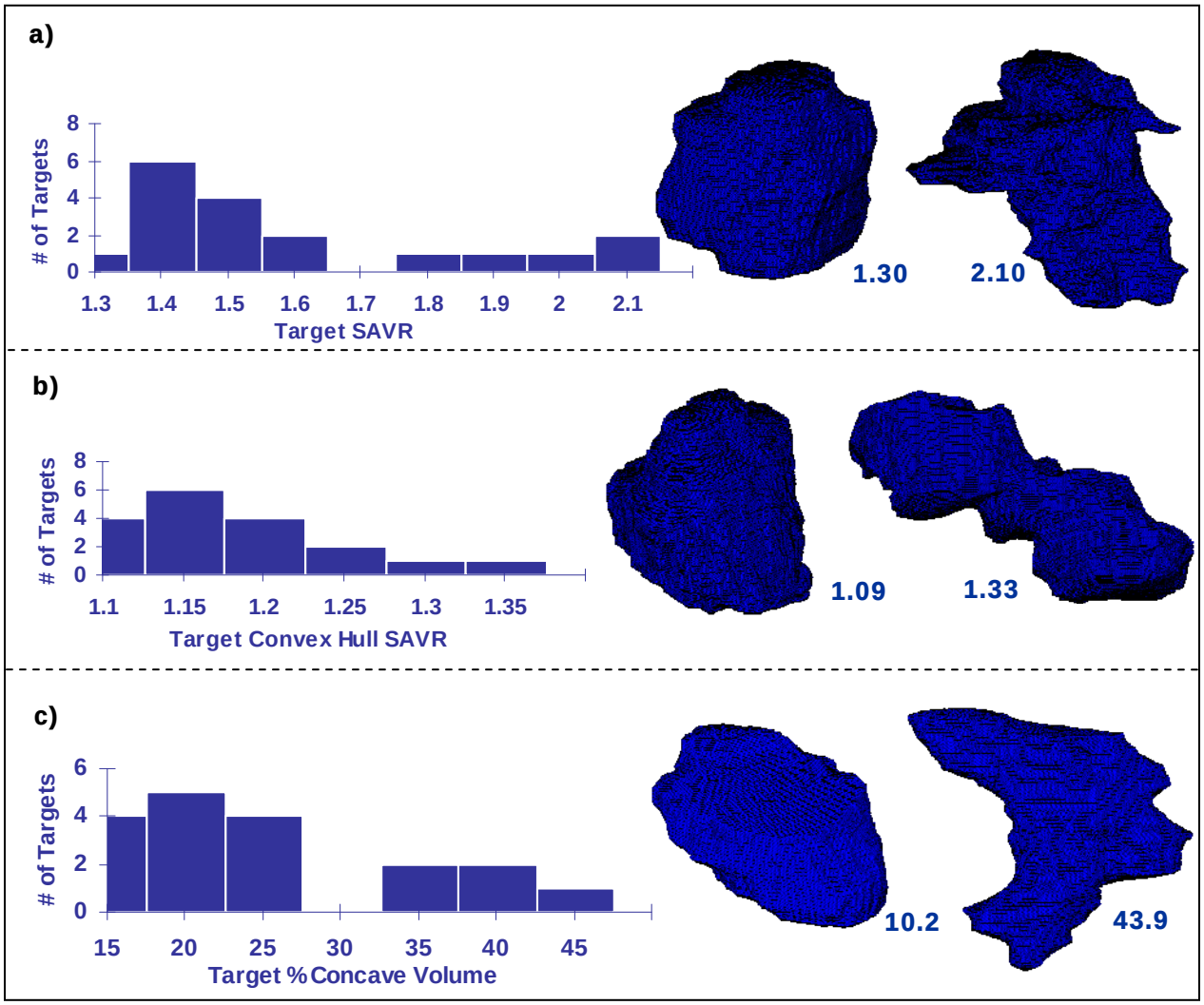


*Figure 3.6: Distribution of a) Target SAVR (normalized surface area to volume ratio), b) Target convex hull SAVR and c) Target percentage of concave volume for the eighteen patients. A target shape corresponding to the minimum and the maximum value for each parameter is shown on the right.*

### 3.2.4 Correlation between Patient Anatomy and Treatment Plan Quality

The correlation between the patient anatomy (anatomical parameters) and the quality of the treatment plan (dosimetric parameters) was examined. The anatomical parameters were used to define the independent variables (predictors). Two sets of predictors were defined: one used to predict the doses delivered to the target and the normal tissue and the other used to predict doses delivered to the critical structures. The dosimetric parameters were used to define the dependent variables. Two sets of dependent variables were defined: one to assess the dose delivered to the target and the normal tissue and the other to assess the dose delivered to the critical structures. All of the anatomical parameters (predictors) and the dosimetric parameters are listed in Table 3.1. Eighteen AVM patients were used as a sample group, resulting in eighteen data points for target and normal tissue (n = 18). In addition, six critical structures in each patient were identified, resulting in 108 data points for critical structures (n = 108). All the critical structures were analysed together to obtain a larger sample pool.

| **Target and Normal Tissue** | | **Critical Structures** | |
|---|---|---|---|
| Anatomical Parameters (Predictors) | Dosimetric Parameters | Anatomical Parameters (Predictors) | Dosimetric Parameters |
| 1) Target volume | 1) MDPD | 1) Structure volume | 1) Structure $D_{max}$ |
| 2) Target SAVR | 2) PITV | 2) Structure SAVR | 2) Structure $V_{24}$ |
| 3) Target convex hull SAVR | 3) Normal Tissue $D_{max}$ | 3) Structure convex hull SAVR | 3) Structure $V_{18}$ |
| 4) Target % concave volume | 4) Normal Tissue $V_{24}$ | 4) Structure % concave volume | 4) Structure $V_{12}$ |
| 5) Target concave volume | 5) Normal Tissue $V_{18}$ | 5) Structure concave volume | 5) Structure $V_6$ |
| 6) Target overlap volume with all critical structures | 6) Normal Tissue $V_{12}$ | 6) Structure overlap volume with the target | 6) Structure $V_3$ |
| 7) Target % of overlap volume with all critical structures | 7) Normal Tissue $V_6$ | 7) Structure % of overlap volume with the target | |
| 8) Average minimum distance between target and structures | 8) Normal Tissue $V_3$ | 8) Minimum distance between the structure and the target | |
| 9) Average mean distance between target and structures | | 9) Mean distance between the structure and the target | |

*Table 3.1: The list of the anatomical parameters (predictors) and the dosimetric parameters. Abbreviations: SAVR = normalized surface area to volume ratio, MDPD = the ratio of the*

*maximum dose to the prescription dose, PITV = the ratio of the volume within the prescription isodose surface to the target volume, $D_{max}$ = the maximum dose, $V_n$ = the volume treated to n Gy.*

For each target and normal tissue dosimetric parameter (second column in Table 3.1) all the target and normal tissue predictors (first column in Table 3.1) showing some statistically significant correlation with that dosimetric parameter were identified. Similarly, for each critical structure dosimetric parameter (fourth column in Table 3.1) all the critical structure predictors (third column in Table 3.1) showing some statistically significant correlation with that dosimetric parameter were identified. One option was to examine the correlation between each dosimetric parameter and each predictor independently, but this method was rejected due to the following reasons. First, the correlation of dosimetric parameter with some predictors can be "masked" by the variations of the other predictors. This would be the case when the two predictors are correlated between themselves, as well as correlated with the dosimetric parameter. Correlation between the predictors poses a potential problem which must be addressed. Secondly, in terms of volume, with 9 possible predictors for both target and normal tissue and critical structures and with 8 dosimetric parameters for target and normal tissue and 6 dosimetric parameters for critical structures, 126 possible relationships for each treatment plan or over 500 possible relationships for all four treatment plans would be generated. Therefore, a more systematic method to analyze the correlation between each dosimetric parameter and all the predictors simultaneously, as well as address the possible correlation between the predictors themselves, was used: Multiple Linear Regression[79].

Each one of the dosimetric parameters ($Y$) from Table 3.1 was fitted with the following regression equation:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \ldots + b_kX_k \tag{3.2}$$

Where $X_1$ , $X_2$ , $X_3$ ,...$X_k$ are the associated predictors for that dosimetric parameter, $a$ is the regression constant and $b_1$, $b_2$, $b_3$, ...$b_k$ are the regression coefficients. In this study, the multiple regression model in which the treatment plan characteristic $Y$ is linearly related to the predictors $X_1$ , $X_2$ , $X_3$ ,...$X_k$ was used. However, the same general techniques can be applied to models involving quadratic or higher order terms in some or all of the predictors. The interpretation of data becomes more subjective as the complexity of the regression model increases, thus we

decided to restrict the analysis to the linear model in this study. To determine which predictors are effective in predicting the dosimetric parameter and should be included in the model and which predictors should be eliminated, forward stepwise regression[80] was used. At the end of the process, all the predictors that showed statistically significant correlation with the dosimetric parameter ($p<0.05$) were identified and the ones that did not show statistically significant correlation ($p>0.05$) were eliminated. As already mentioned, multicollinearity, the correlation between the predictors, is a potential problem with having multiple predictors in the model. The forward stepwise regression technique addresses this problem by excluding predictors from the model if they are highly correlated with some of the predictors that are already included in the model. Therefore, if we have a couple of predictors that are highly correlated to the dosimetric parameter, as well as to each other, only the predictor with the highest degree of correlation with the dosimetric parameter will be included in the model and the others will be eliminated.

In addition to identifying the predictors that show statistically significant correlation with the dosimetric parameter ($p < 0.05$), the regression constant and the regression coefficient for each predictor were calculated. Several diagnostic statistics parameters were also calculated to measure the goodness of the fit[79]:

- $\sigma$ : standard deviation of the regression coefficients
- RMSE : the root mean squared error of the model
- $R^2$ : the amount of dosimetric parameter variability explained by the model

## 3.2.5 Prediction of Plan Quality for Different Treatment Techniques

Once the correlation between the patient anatomy and the dosimetric parameters was established, the model was applied to predict the plan quality of different treatment techniques on a patient-to-patient basis. To examine the effectiveness of the model, treatment plans of all eighteen AVM patients were re-analyzed. For each patient, the dosimetric parameters were calculated (predicted) using the model and compared to the actual dosimetric parameters for all four treatment techniques.

## 3.3 Results

Multiple linear regression model correlation results are given in Appendix A. Table A.1 shows the results for the target and the normal tissue dosimetric parameters and Table A.2 shows the results for the critical structure dosimetric parameters. The tables list all non-zero regression coefficients ($b_1$, $b_2$, $b_3$,...$b_k$), labelled according to Table 3.1, and the regression constant (*a)* for different dosimetric parameters. All the other predictors not shown in the tables did not display statistically significant correlation ($p < 0.05$) with the dosimetric parameters (i.e. their regression coefficients were not significantly different from zero) and were eliminated from the model. Also shown in the tables are the standard deviations of the regression coefficients (σ values) and the root mean squared errors of the model (RMSE). The dosimetric parameters that have no regression coefficients listed either showed no statistically significant correlation ($p > 0.05$) with any of the predictors, or the amount of dosimetric parameter variability explained by the model was less that 50 % ($R^2 < 0.5$).

The data from Tables A.1 and A.2 may be used to quantitatively compare the variation of a particular dosimetric parameter as a function of a single predictor for different treatment techniques (investigation of dosimetric parameter variation as a function of multiple predictors could also be performed, but that would require graphs of dimensions > 2). In cases where a dosimetric parameter has multiple significant predictors, the effect of a single predictor ($X_j$) can be isolated by first subtracting the effects of other predictors from dosimetric parameter. By defining a corrected dosimetric parameter ($Y^{corrected}$) as

$$Y^{corrected} = Y - \left( \sum_{\substack{i=1 \\ i \neq j}}^{k} b_i X_i + a \right) \tag{3.3}$$

a simplified equation that correlates a dosimetric parameter only to a single predictor $X_j$ is obtained:

$$Y^{corrected} = b_j X_j \tag{3.4}$$

To demonstrate how to quantitatively compare the variation of a particular dosimetric parameter as a function of a single predictor for different treatment techniques, two corrected dosimetric parameters ($Y^{corrected}$) are examined. Figure 3.7 displays corrected *Normal tissue* $V_{12}$ vs. *Target volume* and Figure 3.8 displays corrected *Critical structure* $V_{18}$ vs. *Critical structure overlap volume with the target*. In both figures, the slope of the simplified linear regression line, Equation 3.4, is given by the regression coefficient for the associated predictor and the uncertainty of the slope is given by the standard deviation of the regression coefficient. The error bars for the simplified linear regression equation are given by the RMSE values from the model. The $R^2$ values shown in the figures assess the amount of dosimetric parameter variability explained by that predictor. Figures 3.7 and 3.8 are individually examined in the next section.

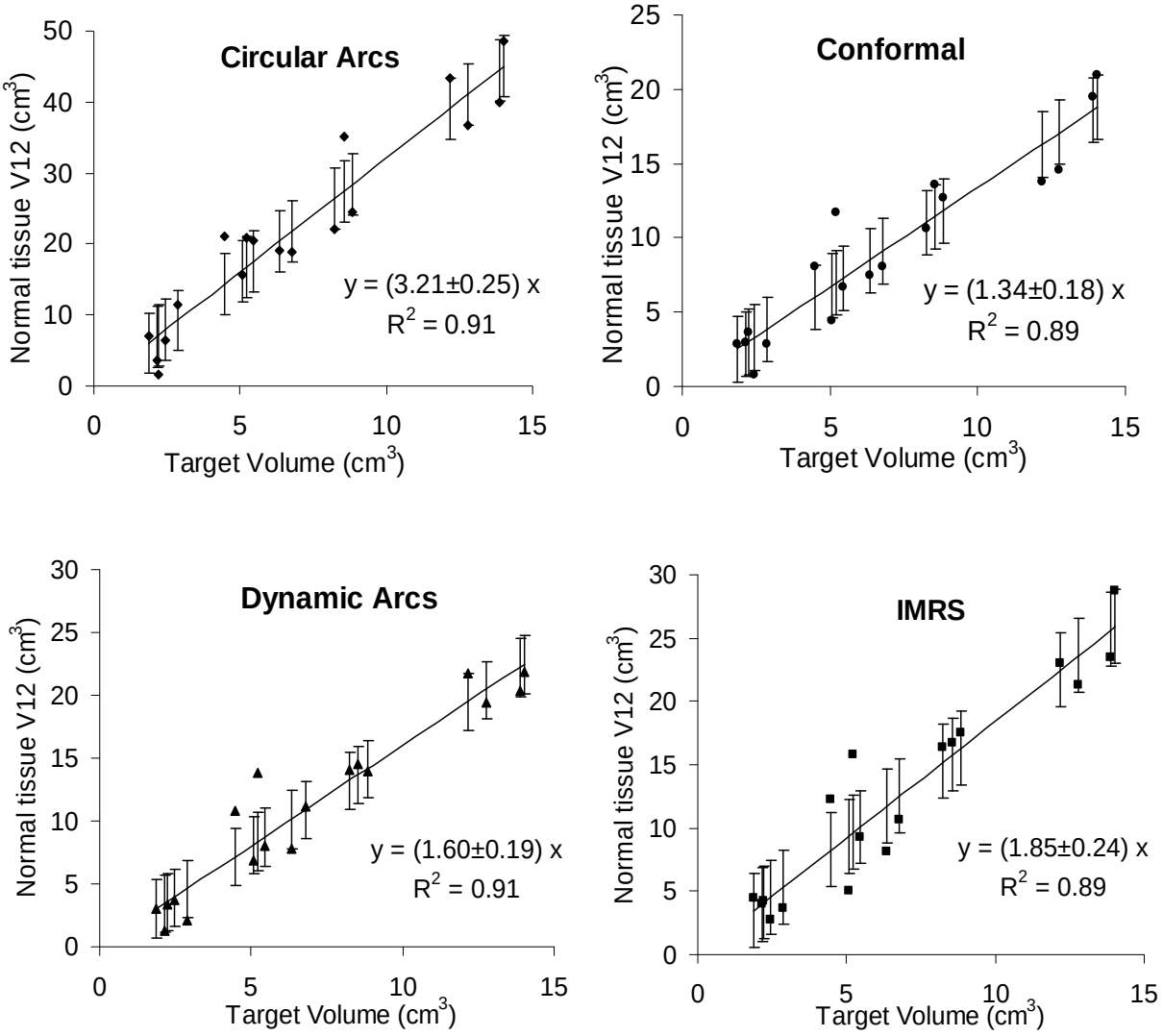


*Figure 3.7: Normal tissue $V_{12}$ (the volume treated to 12 Gy) vs. Target volume for different treatment techniques. The data points represent the corrected Normal tissue $V_{12}$ dosimetric parameter, calculated using Equation 3.3. The straight line for each treatment technique is the simplified linear regression fit, Equation 3.4. The slope of the simplified linear regression line is given by the regression coefficient for the associated predictor (target volume) and the uncertainty of the slope is given by this regression coefficient's standard deviation. The error bars for the simplified linear regression equation are given by the RMSE values from the model. The $R^2$ values shown in the figures assess the amount of treatment plan characteristic variability explained by target volume predictor.*

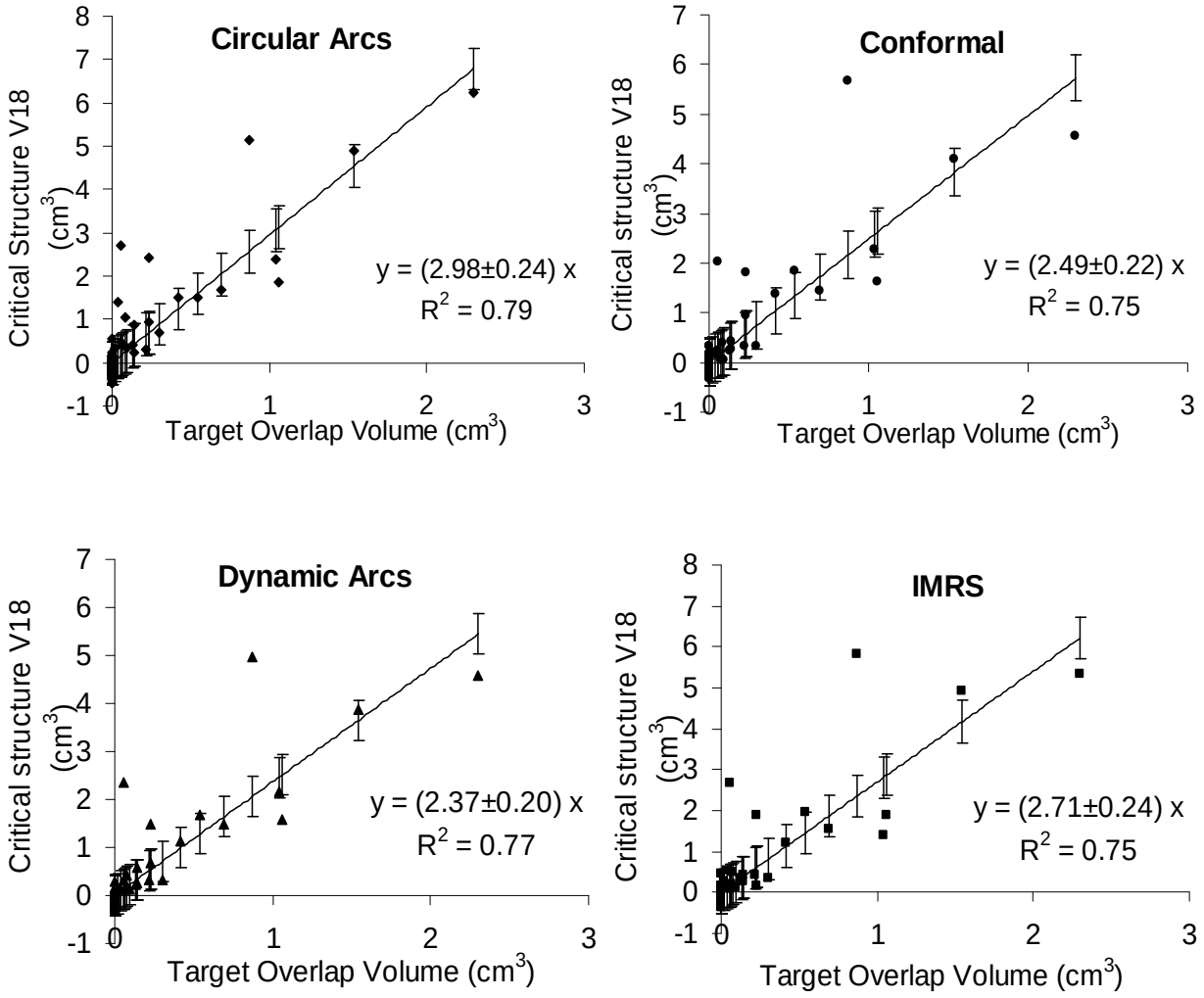


*Figure 3.8: Critical structure $V_{18}$ (the volume treated to 18 Gy) vs. Target Overlap Volume (structure's overlap volume with the target) for different treatment techniques. The data points represent the corrected Critical structure $V_{18}$ dosimetric parameter, calculated using Equation 3.3. The straight line for each treatment technique is the simplified linear regression fit, Equation 3.4. The slope of the simplified linear regression line is given by the regression coefficient for the associated predictor (target overlap volume) and the uncertainty of the slope is given by this regression coefficient's standard deviation. The error bars for the simplified linear regression equation are given by the RMSE values from the model. The $R^2$ values shown in the figures assess the amount of treatment plan characteristic variability explained by target overlap volume predictor.*

# 3.4 Discussion

The data from Tables A.1 and A.2 may be used to obtain various types of information about different treatment techniques. This ranges from: 1) qualitative relationships to 2) quantitative relationships and finally to 3) predicting the plan quality of different radiosurgery techniques based on the individual patient anatomy. Each one is examined separately in the following sections.

## 3.4.1 Qualitative Relationships

Tables A.1 and A.2 may be used to determine which predictors are significantly correlated to various dosimetric parameters for different treatment techniques. For example, the maximum dose delivered to the critical structures (critical structure $D_{max}$ in Table A.2) is significantly correlated only to the overlap of the structure with the target and the minimum distance between the critical structure and the target for all four treatment techniques. Furthermore, the fact that the calculated regression coefficients for structure overlap with the target are positive for all four treatment techniques indicates that the correlation is positive: as the structure overlap with the target increases, the maximum dose to that critical structure also increases. Similarly, the negative regression coefficients for the minimum distance between the critical structure and the target for all four treatment techniques indicate that the correlation is negative: as the minimum distance between the critical structure and the target increases, the maximum dose to the critical structure decreases. These relationships are intuitive and not very surprising.

A less intuitive correlation may be observed by looking at the maximum dose delivered to the normal tissue (normal tissue $D_{max}$ in Table A.1). The normal tissue $D_{max}$ is significantly correlated only to the average mean distance between the target and the critical structures for all techniques except the circular collimator arcs, which displays no significant correlation with any of the predictors. The fact that the regression coefficients are positive indicates that the correlation is positive: as the average mean distance between the target and the critical structures increases, the maximum dose to the normal tissue increases. One possible explanation for this could be that in the patients where the critical structures are in close proximity to the target, great care is taken by the planner to reduce the dose around the target. This is generally achievable by

spreading the beams around the target, which decreases the maximum dose delivered to the normal tissue. Similar observations could be made for other dosimetric parameters by analyzing the regression coefficients of their associated predictors. This information may be used to obtain qualitative relationships between the patient anatomy and different dosimetric parameters.

It should be noted here that a dosimetric parameter being correlated "only" to a single predictor implies that if that predictor is included in the model then no other predictor is significantly correlated to that dosimetric parameter any longer. All the other predictors either show no correlation with this dosimetric parameter or are so highly correlated with the original predictor in the model that their inclusion does not add any new information to the model. If the latter is true, then taking out the original predictor from the model would allow some other predictor to become significantly correlated to this dosimetric parameter and enter the model.

## 3.4.2 Quantitative Relationships

Plots, such as the ones shown in Figures 3.7 and 3.8, quantitatively compare the variation of a particular dosimetric parameter with specific predictors for different treatment techniques. For example, from Figure 3.7 it is observed that for all four techniques the normal tissue $V_{12}$ varies linearly with the target volume. The $R^2$ value is around 0.90 for all four techniques indicating that the linear fit accounts for about 90% of the variability in the observations. Also, the slopes for all four techniques are different from each other, up to the standard deviation, indicating that the normal tissue $V_{12}$ correlates differently with the target volume for different techniques. In Figure 3.8, the linear fit accounts for 75 – 79 % of the variability in the observations and the slopes are different for different techniques, indicating that the critical structure $V_{18}$ correlates differently with the target overlap volume for different techniques (except for conformal fields and dynamic arc techniques, whose slopes are equal to within the standard deviation).

## 3.4.3 Prediction of Plan Quality for Different Treatment Techniques

The multiple linear regression model was applied to predict the plan quality of different treatment techniques on a patient-to-patient basis. The model predicted dosimetric parameters based purely on individual patient anatomy, without the need to create multiple treatment plans

using different treatment techniques. Creating multiple treatment plans to choose the optimal technique is time consuming and somewhat impractical, given the time constraints imposed by the procedure. With this method, however, the prediction of dosimetric parameters for different treatment techniques was readily available. The predicted dosimetric parameters were generally in close agreement, within associated standard deviations, with the actual data. Based on the predicted results, most of the patients could be classified into one of the following three groups: 1) one technique was clearly superior over the other three techniques; 2) there were two competing techniques with similar dosimetric parameters; and 3) different techniques performed better for different dosimetric parameters. A case from each one of these three categories is presented and discussed.

Table 3.2 shows the actual dosimetric parameters, the predicted dosimetric parameters and the uncertainty in the predicted dosimetric parameters for the target, the normal tissue and one critical structure for a patient belonging to the first category. The other critical structures displayed similar results and are omitted from the table for clarity. To further reduce the amount of data displayed in the table, only the dosimetric parameters for which the four techniques had different results are displayed. The dosimetric parameters for which the four techniques had comparable results, within associated standard deviations, are not useful in distinguishing the effectiveness of different techniques and are therefore not displayed in the table. From Table 3.2, we can see that most of the predicted data agrees well, to within associated standard deviations, with the actual data. For each dosimetric parameter, the technique that performed the best can be identified by the lowest dosimetric parameter value. The lowest values are highlighted for both actual and predicted dosimetric parameters in Table 3.2. It is clear that for this patient, Dynamic Arcs is the optimal treatment technique since it displays the lowest values for almost all dosimetric parameters. The same conclusion is reached from both the actual and predicted dosimetric parameter values, demonstrating the effectiveness of this method.

| | **Circular Arcs** | | | **Conformal** | | | **Dynamic Arcs** | | | **IMRS** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Actual | Predicted | Diff | Actual | Predicted | Diff | Actual | Predicted | Diff | Actual | Predicted | Diff |
| **Target** | | | | | | | | | | | | |
| MDPD | 2.00 | none | - | 1.30 | 1.29±0.01 | 1.1 | ***1.28*** | ***1.28±0.01*** | 0.2 | 1.29 | none | - |
| PITV | 2.87 | none | - | 2.03 | 2.11±0.15 | 0.5 | ***1.89*** | ***1.99±0.07*** | 1.3 | 2.21 | 2.23±0.14 | 0.1 |
| **Normal Tissue** | | | | | | | | | | | | |
| $D_{max}$ (Gy) | 39.2 | none | - | 25.5 | 24.2±2.1 | 0.6 | 25.1 | 24.0±2.1 | 0.4 | ***25.0*** | ***22.7±1.9*** | 1.2 |
| $V_{18}$(cm $^3$) | 19.6 | 16.0±2.9 | 1.2 | 12.4 | 12.1±1.6 | 0.2 | ***10.9*** | ***10.4±1.6*** | 0.3 | 13.1 | 12.1±2.9 | 0.4 |
| $V_{12}$(cm $^3$) | 38.8 | 31.1±4.3 | 1.8 | 31.3 | 29.1±2.2 | 1.0 | ***26.7*** | ***25.8±2.3*** | 0.4 | 33.4 | 32.4±2.9 | 0.3 |
| $V_6$(cm $^3$) | 113 | 93±11 | 1.8 | 121 | 107±29 | 0.5 | ***85*** | ***80±7*** | 0.8 | 149 | 130±21 | 0.9 |
| $V_3$(cm $^3$) | 309 | 282±31 | 0.9 | 338 | 359±62 | 0.3 | ***235*** | ***217±24*** | 0.8 | 351 | 394±63 | 0.7 |
| **Critical Structure (Optic nerve)** | | | | | | | | | | | | |
| $V_6$(cm $^3$) | 4.0 | 3.7±1.2 | 0.3 | 2.8 | 2.8±1.3 | 0.0 | ***2.1*** | ***2.1±1.2*** | 0.0 | 3.7 | 3.2±1.4 | 0.4 |
| $V_3$(cm $^3$) | 11.6 | 7.3±2.2 | 1.9 | 11.9 | 10.2±2.5 | 0.7 | ***8.2*** | ***6.1±2.1*** | 1.0 | 12.5 | 10.7±2.6 | 0.7 |

*Table 3.2: Actual and predicted dosimetric parameters for the target, the normal tissue and the optic nerve are shown for all four techniques. Highlighted are the lowest dosimetric parameters among all four techniques. Abbreviations: MDPD = the ratio of the maximum dose to the prescription dose, PITV = the ratio of the volume within the prescription isodose surface to the target volume, $D_{max}$ = the maximum dose, $V_n$ = the volume treated to n Gy, Diff = |Predicted value – Actual value| / (Standard deviation of the predicted value).*

Table 3.3 displays the results for a patient belonging to the second category (two techniques with superior results). Again, we can see that most of the predicted data agrees well, within associated standard deviations, with the actual data. The two techniques with the lowest dosimetric parameters (highlighted) are the Circular Arcs and the Dynamic Arcs. The Dynamic Arcs perform slightly better than Circular Arcs for target dosimetric parameters, but the two techniques have comparable dosimetric parameters for normal tissue and critical structure. In such cases, the user might want to consider other factors, such as plan complexity, treatment time, treatment accuracy verification, treatment delivery efficiency, treatment uncertainties, etc., besides the dose distribution parameters when choosing between the two techniques. Knowing that the dose distribution will not be significantly compromised compared to a more complex dose delivery method, the user might decide it is more practical to choose a simpler method of dose delivery.

| | **Circular Arcs** | | | **Conformal** | | | **Dynamic Arcs** | | | **IMRS** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Actual | Predicted | Diff | Actual | Predicted | Diff | Actual | Predicted | Diff | Actual | Predicted | Diff |
| **Target** | | | | | | | | | | | | |
| MDPD | 1.39 | none | - | 1.28 | 1.30±0.01 | 1.6 | 1.28 | ***1.28±0.01*** | 0.3 | ***1.19*** | none | - |
| PITV | 2.59 | none | - | 2.48 | 2.39±0.15 | 0.6 | ***2.20*** | ***2.25±0.07*** | 0.6 | 2.65 | 2.48±0.14 | 1.2 |
| **Normal Tissue** | | | | | | | | | | | | |
| $D_{max}$ (Gy) | 27.0 | none | - | 25.0 | 26.1±2.1 | 0.5 | 25.0 | 26.2±2.1 | 0.6 | ***23.3*** | ***24.8±1.9*** | 0.8 |
| $V_{18}$(cm $^3$) | 11.6 | 11.0±2.9 | 0.2 | 11.0 | 10.7±1.6 | 0.2 | ***9.1*** | ***9.1±1.6*** | 0.0 | 12.9 | 10.6±2.9 | 0.8 |
| $V_{12}$(cm $^3$) | *22.1* | *21.2±4.3* | 0.2 | 24.4 | 25.0±2.2 | 0.3 | ***20.2*** | ***20.9±2.3*** | 0.3 | 26.7 | 27.6±2.9 | 0.3 |
| $V_6$(cm $^3$) | ***67*** | ***60±11*** | 0.6 | 107 | 74±29 | 1.2 | ***63*** | ***63±7*** | 0.0 | 101 | 100±21 | 0.1 |
| $V_3$(cm $^3$) | ***199*** | ***185±31*** | 0.4 | 431 | 357±62 | 1.2 | ***189*** | ***212±24*** | 1.0 | 467 | 391±63 | 1.2 |
| **Critical Structure (Left Motor Strip)** | | | | | | | | | | | | |
| $D_{max}$ (Gy) | ***3.6*** | ***3.1±2.7*** | 0.2 | 9.3 | 5.2±2.8 | 1.4 | ***2.8*** | ***2.6±2.6*** | 0.1 | 8.0 | 5.7±2.2 | 1.0 |
| $V_3$(cm $^3$) | ***0.1*** | ***2.2±2.2*** | 0.9 | 11.3 | 8.9±2.5 | 1.0 | ***0.0*** | ***1.2±2.1*** | 0.6 | 11.9 | 9.4±2.6 | 1.0 |

*Table 3.3: Actual and predicted dosimetric parameters for the target, the normal tissue and the left motor strip are shown for all four techniques. Highlighted are the lowest dosimetric parameters among all four techniques. Abbreviations: MDPD = the ratio of the maximum dose to the prescription dose, PITV = the ratio of the volume within the prescription isodose surface to the target volume, $D_{max}$ = the maximum dose, $V_n$ = the volume treated to n Gy, Diff = |Predicted value – Actual value| / (Standard deviation of the predicted value).*

Table 3.4 shows the results for a patient from the third category (different techniques perform better for different dosimetric parameters). Again, we can see that most of the predicted data agrees well, within associated standard deviations, with the actual data. For this patient, the lowest dosimetric parameter values (highlighted) are scattered across the table indicating that different techniques perform better for different dosimetric parameters. For example, the fixed conformal field technique is superior for target dose distribution and normal tissue $V_{18}$ and $V_{12}$, dynamic arcs technique is superior for lower doses to normal tissue ($V_6$ and $V_3$), while IMRS is superior for minimizing the maximum dose ($D_{max}$) to the normal tissue and optic nerve. In such a case, the user must decide which goals are of the highest importance and choose the technique that is optimal for realizing them (ie. technique that has the lowest values for those dosimetric parameters). For example, dose conformity to the target volume is an important dosimetric parameter, while the dose homogeneity throughout the target volume may be of slightly less

concern and importance. Also, the tolerance dose of a particular structure can indicate which Vn values might be the most relevant. For example, the tolerance dose to the optic nerve is about 10 Gy for a single fraction radiosurgery. Therefore, choosing the technique that has a lower value for $V_{12}$ (Dynamic Arcs technique) might be more important than minimizing the $V_3$ value (Circular Arc technique).

| | **Circular Arcs** | | | **Conformal** | | | **Dynamic Arcs** | | | **IMRS** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Actual | Predicted | Diff | Actual | Predicted | Diff | Actual | Predicted | Diff | Actual | Predicted | Diff |
| **Target** | | | | | | | | | | | | |
| MDPD | 1.96 | none | - | 1.26 | 1.28±0.01 | 1.8 | 1.28 | 1.29±0.01 | 0.7 | ***1.23*** | none | - |
| PITV | 2.15 | none | - | ***1.53*** | ***1.65±0.15*** | 0.7 | 1.64 | 1.66±0.07 | 0.2 | 1.91 | 1.85±0.14 | 0.4 |
| **Normal Tissue** | | | | | | | | | | | | |
| $D_{max}$ (Gy) | 27.2 | none | - | 24.3 | 24.7±2.1 | 0.2 | 24.5 | 24.5±2.1 | 0.0 | ***23.3*** | ***22.7±1.9*** | 0.3 |
| $V_{18}$(cm $^3$) | 20.4 | 16.7±2.9 | 1.3 | ***10.1*** | ***9.4±1.6*** | 0.4 | 11.7 | 10.9±1.6 | 0.5 | 15.1 | 14.5±2.9 | 0.2 |
| $V_{12}$(cm $^3$) | 47.1 | 42.7±4.3 | 1.0 | ***25.3*** | ***27.8±2.2*** | 1.1 | 29.3 | 28.0±2.3 | 1.0 | 34.2 | 33.7±2.9 | 0.2 |
| $V_6$(cm $^3$) | 136 | 123±11 | 1.2 | 116 | 137±29 | 0.7 | ***101*** | ***91±7*** | 1.4 | 128 | 143±21 | 0.8 |
| $V_3$(cm $^3$) | 395 | 361±31 | 1.1 | 438 | 450±62 | 0.2 | ***309*** | ***294±24*** | 0.7 | 491 | 503±63 | 0.2 |
| **Critical Structure (Optic nerve)** | | | | | | | | | | | | |
| $D_{max}$ (Gy) | 25.6 | 25.6±2.7 | 0.0 | 25.0 | 22.4±2.8 | 0.9 | 25.0 | 22.1±2.6 | 1.1 | ***23.3*** | ***20.0±2.2*** | 1.5 |
| $V_{12}$(cm $^3$) | 3.3 | 3.1±0.7 | 0.3 | 2.7 | 2.4±0.7 | 0.4 | ***2.4*** | ***2.0±0.7*** | 0.6 | 3.0 | 2.7±0.9 | 0.3 |
| $V_3$(cm $^3$) | ***12.2*** | ***10.1±2.2*** | 1.0 | 15.6 | 13.1±2.5 | 1.0 | 14.0 | 12.4±2.1 | 0.8 | 16.1 | 13.7±2.6 | 0.9 |

*Table 3.4: Actual and predicted dosimetric parameters for the target, the normal tissue and the left motor strip are shown for all four techniques. Highlighted are the lowest dosimetric parameters among all four techniques. Abbreviations: MDPD = the ratio of the maximum dose to the prescription dose, PITV = the ratio of the volume within the prescription isodose surface to the target volume, $D_{max}$ = the maximum dose, $V_n$ = the volume treated to n Gy, Diff = |Predicted value – Actual value| / (Standard deviation of the predicted value).*

The predicted data could be used to construct DVHs for different techniques and allow graphical comparison of the results. Figure 3.9 shows the predicted normal tissue DVHs for different techniques for one of the AVM patients. The predicted values are shown as points, with error bars given by the RMSE values, and the smooth lines represent the actual data obtained from the normal tissue DVHs for different techniques. Both the actual data and the predicted

values indicate that the Dynamic Arcs technique is the optimal technique for reducing the dose to the normal tissue for this patient.

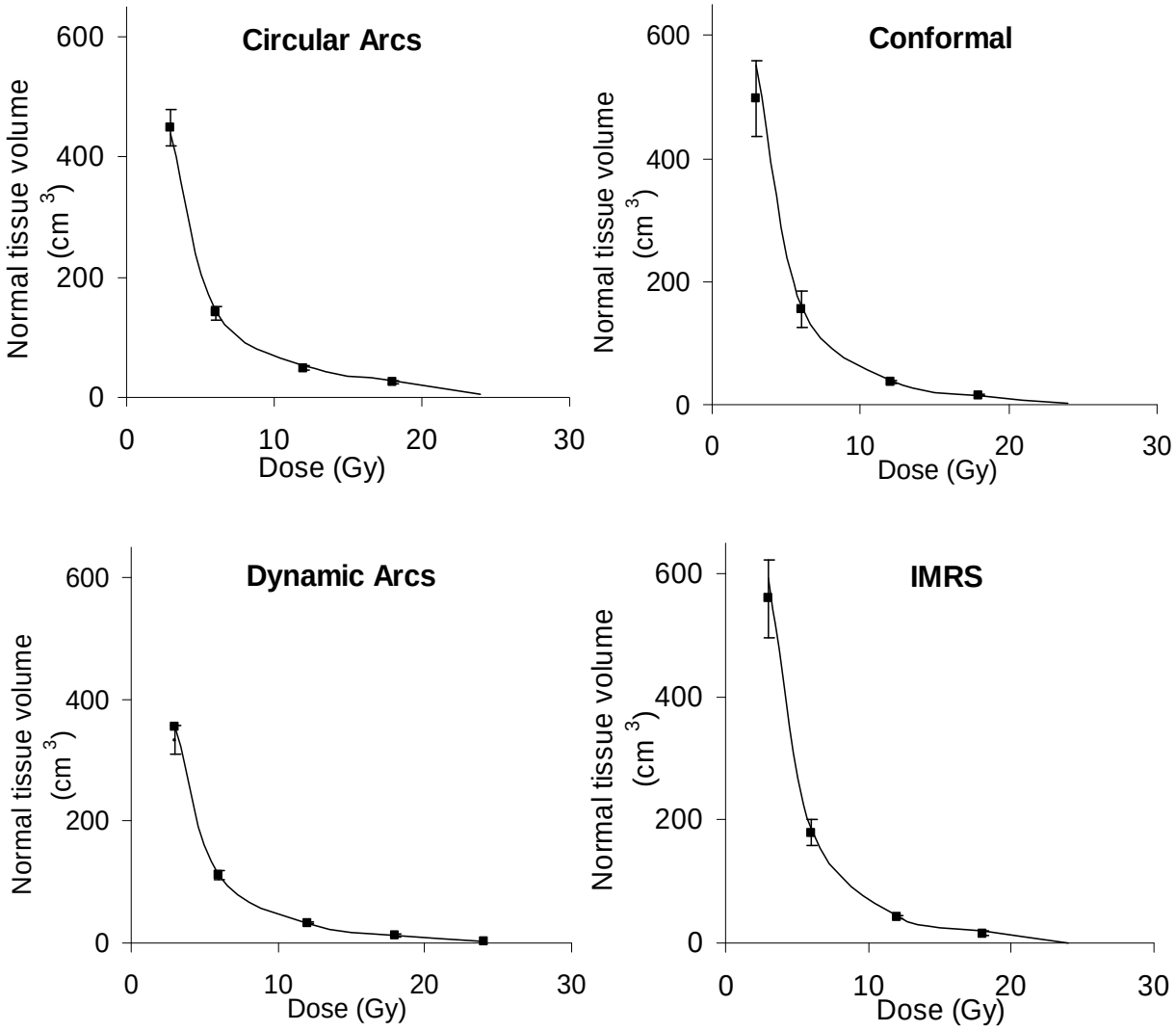


*Figure 3.9: The predicted (points with error bars) and the actual (lines) normal tissue DVHs for different treatment techniques for one of the patients. The error bars are given by the root mean squared error (RMSE) values from the model.*

## 3.5 Conclusion

This study is the first quantitative analysis of the correlation between the patient anatomy and the quality of the treatment plan produced by different radiosurgery techniques. Eighteen previously treated AVM patients were used as a sample group in this study. The patients exhibited a wide range of target size, shape and location. Multiple linear regression model was used to establish the correlation between the patient anatomy and the plan quality for different treatment techniques. The results from the model were used to compare the effectiveness of different treatment techniques based on the patient anatomy. A number of different relationships, both qualitative and quantitative, were discovered between the patient anatomy and dosimetric parameters for different techniques.

The model was employed to predict the overall plan quality of different radiosurgery techniques based on the individual patient anatomy. The predicted dosimetric parameters were in close agreement with the actual data, indicating the effectiveness of the model. The results were used to predetermine the optimal radiosurgery technique/techniques for a number of patients, without having to create multiple treatment plans with different techniques. Therefore, this method shows great potential to be a valuable tool for selecting the most effective treatment technique. Even though this method was developed for radiosurgery of AVMs, the technique is completely general and could be used for other types of radiation therapy treatments.

# Chapter 4
# On-line Adaptive Radiation Therapy Using Direct Aperture Optimization

## 4.1 Introduction

This work was motivated by an initial study, presented in Chapter 3, which investigated the correlation between the patient anatomy and the quality of the treatment plan produced by different radiosurgery techniques. In that study, an inherent assumption was that the patient anatomy does not change during the entire treatment process. For radiosurgery treatments, this assumption is to some extent justified since the whole treatment is delivered in only one fraction. Typically, the entire radiosurgical process, including patient imaging, treatment planning and radiation delivery are completed in a single day. This greatly reduces the susceptibility to anatomical deformations. Furthermore, at the beginning of radiosurgery procedure, a carbon fibre head-ring is rigidly attached to the patient's skull using four sharp pins. The head ring serves to immobilize the head during radiation delivery, which essentially eliminates patient motion during the treatment. Attached to the head ring is the localizer box, which defines a rigid 3D coordinate system and reduces the uncertainties in patient positioning. Therefore, for radiosurgery treatments, the uncertainties in patient positioning, patient motion and anatomical deformations during the treatment are all reduced by design of the treatment.

However, for a majority of other radiation therapy treatments the radiation delivery is fractionated (30–40 fractions) and the assumption that the patient anatomy remains the same throughout the whole treatment process (several weeks) becomes unrealistic. Also, the patient is not immobilized prior to radiation delivery which increases susceptibility to patient set-up inaccuracies and patient motion during the treatment. One way to deal with these sources of increased uncertainty is to use large margins around the CTV to obtain the PTV. This imposes a limit on the ability to closely tailor the dose distribution to CTV and minimize dose received by

the surrounding OARs[33]. To overcome this limit of conventional radiation therapy, a new form of radiation therapy, on-line ART, has recently emerged (refer to Section 2.5).

As discussed in Section 2.5.3, one of the current drawbacks for clinical implementation of on-line ART is the substantially extended treatment time, as compared to conventional (non-adaptive) radiation therapy. The addition of daily imaging and plan adaptation (re-optimization) steps roughly doubles the overall treatment time, reducing the patient throughput by half. Also, prolonged treatment times result in increased susceptibility to patient motion during the treatment. This increases the uncertainty in patient anatomy and position during the treatment, which requires the margin around the CTV to be expanded, undermining the key advantage of on-line ART – reduced margin size. Therefore, before on-line ART can be successfully put into clinical practice, the time allotted for daily imaging and plan adaptation (IMRT plan re-optimization) has to be drastically reduced. For on-line ART to be clinically acceptable the time allocated for plan adaptation has to be on the order of a few minutes or less. In this chapter, methods for accelerating plan adaptation are investigated (Figure 2.20b).

Recently, the importance of accelerating IMRT plan adaptation has been recognized by several research groups and different methods have been proposed to rapidly adapt the original treatment plan[51-54]. In the approach taken by Mohan et al.[52], the original intensity distributions are modified based on the overlap of the PTV with the OAR in the BEV. Court et al.[53, 54] modify the MLC leaf positions based on the geometrical relationship between the anatomy deformation and the MLC leaves. Feng et al.[51] use the deformation matrix, which associates the planning image set and the images acquired on the day of treatment, to morph the treatment apertures. However, these methods do not perform a true re-optimization of the original plan. Therefore, even though they are generally time efficient, there is no guarantee that the adapted plan is clinically acceptable. Also, these methods do not indicate how close the adapted plan is to the optimal plan.

Our approach to accelerating plan adaptation is to develop a method which rapidly and efficiently re-optimizes the original IMRT plan. Instead of generating and adapting fluence-based IMRT plans, we have focused our attention on IMRT plans obtained by DAO (refer to Section 2.4.4). All the other studies mentioned earlier focused on fluence-based IMRT plans. In addition to the advantages offered by DAO over fluence-based optimization, discussed in Section 2.4.4, we have recognized that DAO offers additional advantages that make it highly suitable for on-line ART. These advantages include: 1) bypassing of the leaf sequencing step

reduces the overall time needed for generating a new plan, 2) decreased number of MUs results in shorter delivery times and 3) control over the degree of plan adaptation and optimization scheme allows for accelerated plan adaptation. The investigation of the potential for accelerated plan adaptation is the main focus of this chapter. All of these advantages ultimately contribute to reducing the overall treatment time required for on-line ART. Finally, by performing the plan adaptation using DAO the adapted plans are actually optimized, in contrast to the methods mentioned earlier. These features make DAO an attractive and promising approach for on-line ART. This study is the first investigation into the potential of using DAO for on-line ART. This work was published in the journal of Medical Physics[81].

# 4.2 Methods and Materials

## 4.2.1 Treatment Planning System

### Direct Aperture Optimization

Our DAO algorithm is implemented in the MATLAB programming environment (Mathworks Inc., Natick, MA). The implementation follows the description given in Section 2.4.4. Initially, all the aperture shapes are conformed to the BEV of the PTV and the aperture weights are all equal. The aperture weights are initially set so that the mean dose to PTV is equal to the prescribed dose. The PTV, OARs and body contours are exported from the Eclipse (Varian Inc. Palo Alto, CA - version 6.5) treatment planning system (TPS) into MATLAB using DICOM RT. Random point sampling[82] is used to distribute points in each structure that is included in the optimization. The voxel resolution is set to 3 mm x 2.5 mm x 2.5 mm.

### Objective (Cost) Function

A quadratic cost function is used to guide the optimization towards the optimal solution. The cost function is composed of two sets of terms: one for the PTV and one for each OAR, as described in Section 2.4.2. For the PTV, the minimum and the maximum allowed dose ($D_{min}$ and $D_{max}$) are specified together with their respective priorities. For each OAR, multiple dose-volume

constraints are specified together with their respective priorities. The cost function is constructed such that its value is always positive. The goal of the optimization is to minimize the cost function value. The cost value of zero indicates that all of the treatment planning goals are completely satisfied.

### Dose Calculation Algorithm

Dose calculation is performed using the PBK algorithm, described in Section 2.4.3. One of the unique features of our PBK algorithm is the dose calculation method. Generally, PBK algorithms pre-calculate the pencil beam dose distribution (PBDD) before the start of the optimization. PBDD is the dose distribution associated with each beamlet of radiation (Figure 2.10), obtained by scaling the pencil beam dose kernel *K (x, y, d)* with various factors, such as inverse square law and primary beam off-axis intensity profiles (Equation 2.7). The pre-calculated PBDDs are usually truncated or a coarser voxel resolution is used, due to the computer storage memory requirements. In our optimization system, we use a fast pencil beam dose calculation algorithm[83] with full scatter conditions. The pencil beam dose kernel *K (x, y, d)* is directly used during the optimization, thus eliminating the pre-calculation step. Variable beamlet sizes are used to speed up the dose calculation process. An average MLC transmission of 1.6% was experimentally measured and included in our dose calculation model, accounting for radiation transmitted through MLC leaves (Varian Millenium 120).

## 4.2.2 Anatomical Deformation Model

A model simulating a typical prostate case was created in the TPS. A prostate case was used as an example because significant changes in anatomy from day to day are well documented[84-88]. However, the approach and ideas developed here are completely generalizable and could be applied to other treatment sites. In consultation with a radiation oncologist, a model was created where a prostate, rectum and a bladder are represented by an ellipsoid, cylinder and a sphere, respectively, as shown in Figure 4.1. We have chosen to use a virtual patient model, instead of real patient data, to permit systematic and controlled deformation of anatomical structures.

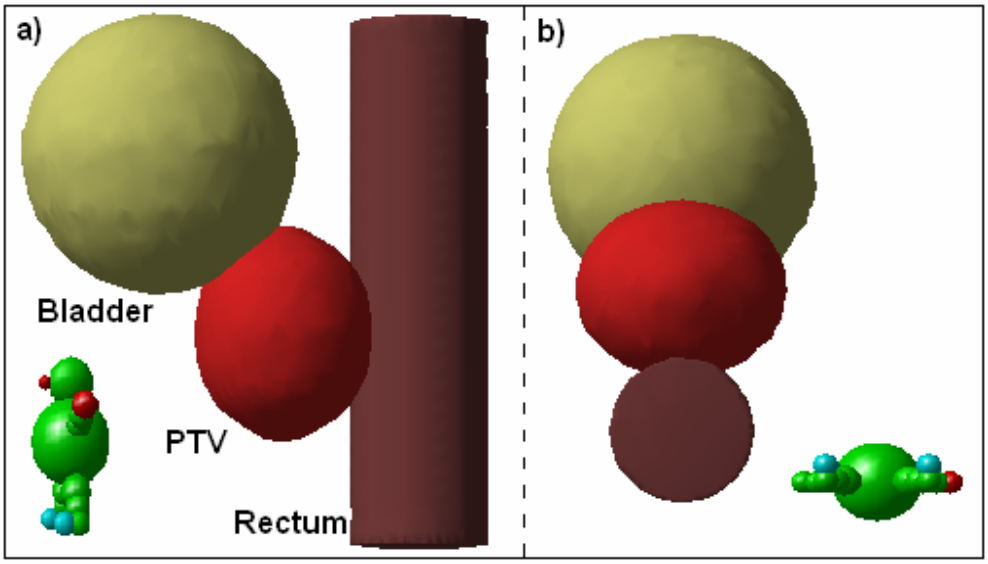


*Figure 4.1: A model simulating a prostate case: a) Left-Right view; b) Inferior-Superior view*

The dimensions and positions of these structures were chosen to be as realistic as possible and represent a typical prostate patient with a full bladder. The prostate was represented by an ellipsoid of 2.5 cm radius in the left-right direction, 2 cm in the anterior-posterior direction and 2.5 cm in the superior-inferior direction. The bladder was represented by a sphere of radius 4 cm and the rectum was represented by a cylinder of radius 2 cm. The prostate, which represented the CTV, was positioned so that it just touches the bladder and the rectum. The PTV was created by adding a 5 mm margin to the CTV, resulting in a region of overlap between the PTV and the rectum and the PTV and the bladder. The 5 mm margin is the smallest acceptable CTV to PTV margin in the RTOG 0415 prostate IMRT protocol[89]. Since on-line ART effectively eliminates the inter-fractional anatomy variations and the uncertainties associated with them, it is reasonable to use the smallest acceptable margin. The model anatomy resided in a square, 40 cm by 40 cm, phantom. The PTV, rectum, bladder and phantom (body) contours were exported from the TPS into MATLAB using DICOM RT. This configuration represented the "original anatomy" and was used to create the original DAO treatment plan (discussed in the next section).

Next, the original anatomy was deformed by enlarging the rectum[90, 91]. As a result of enlarging the rectum, the prostate was expanded in the left-right and superior-inferior directions and compressed in the anterior-posterior direction. Also, the shape of the prostate was changed from convex to concave, due to the localized pressure from the rectum. Four different "deformed anatomies" were created by systematically deforming the original anatomy by various amounts

(0.25, 0.50, 0.75 and 1.00 cm). For example, 0.25 cm deformation involved the increase of the rectum's radius by 0.25 cm, increase of the prostate's L-R radius by 0.25 cm, increase of the prostate's S-I radius by 0.25 cm and decrease of the prostate's A-P radius by 0.125 cm, half of 0.25 cm. The decrease of the prostate's A-P radius was required in order to keep the volume of the prostate roughly constant for all four deformed anatomies, within 5% of the original volume. As a result of the above mentioned deformations, the shape of the PTV was highly distorted, but the displacement of the centroid of the PTV was less than 2 mm for all four deformed anatomies. The chosen range of deformations was intended to cover a range of clinically relevant, non-rigid prostate deformations. The largest deformation represented an extreme case, rarely observed in patients[88] and was included mainly for algorithm testing purposes. As with the original anatomy, the deformed anatomies were first created in the TPS and then the contours of the deformed structures were exported from the TPS into MATLAB using DICOM RT. The deformed anatomies are shown in Figure 4.2.

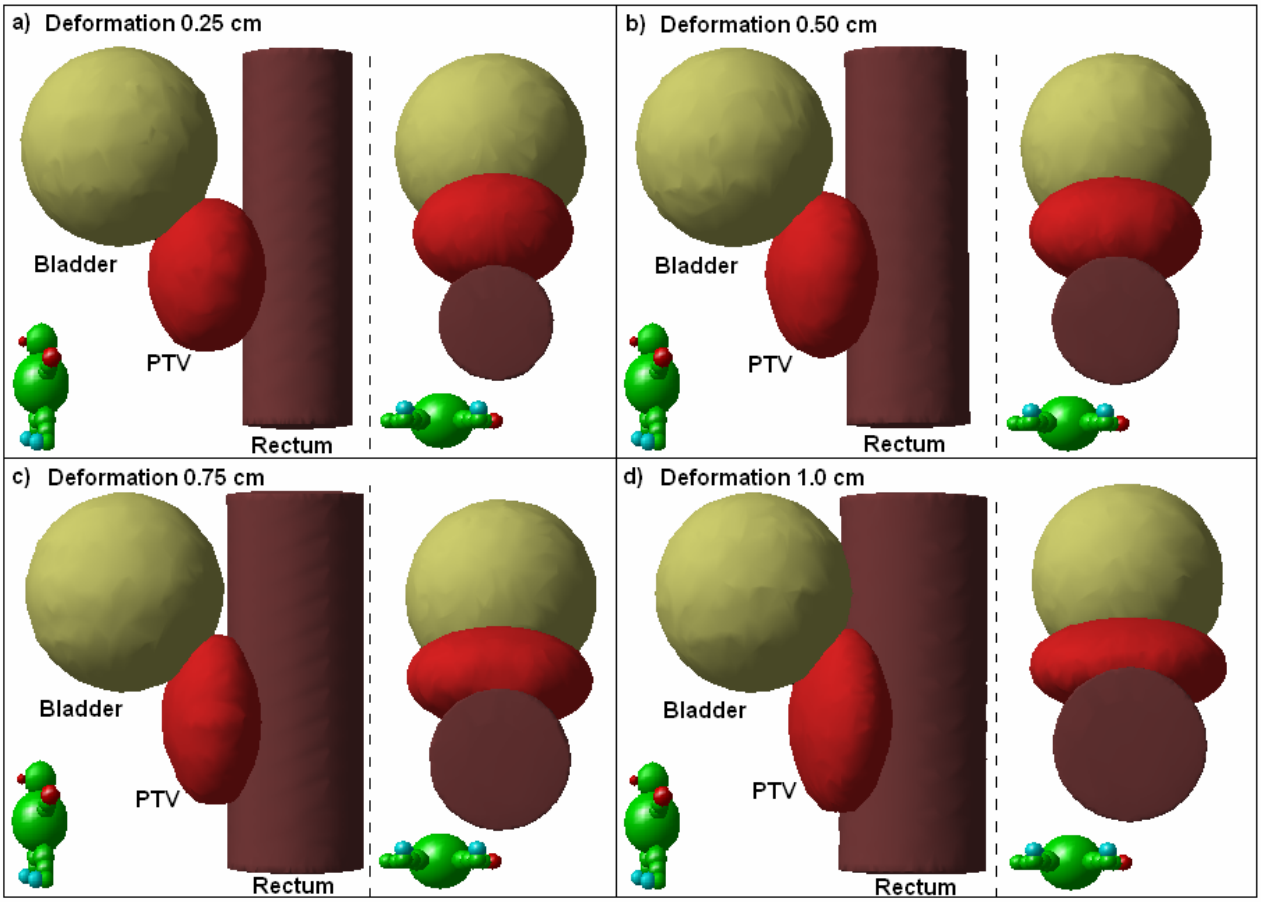


*Figure 4.2: Four deformed anatomies created by systematically deforming the original anatomy by various amounts: a) 0.25 cm; b) 0.50 cm; c) 0.75 cm; d) 1.00 cm.*

### 4.2.3 Original Treatment Plan

Our DAO system was used to create the "original treatment plan" for the original anatomy. Seven beams were used with gantry angles of 40, 80, 110, 250, 280, 310 and 355 degrees, with six apertures per beam. The number of beams and gantry angles were obtained from our institutional prostate IMRT protocol in clinical use. Our choice of six apertures per beam was based on the study of the required number of apertures for DAO by Jiang et al.[92] and our own experience. The dose-volume constraints for the bladder and the rectum were based on the RTOG 0415 prostate IMRT protocol.

For the purposes of this study, the prescribed dose to the PTV was escalated from 73.8 Gy, which is specified by the RTOG 0415 protocol, to 83.8 Gy. This was done for two reasons. First, we wanted to investigate the possibility of using a higher prescribed dose while still meeting the dose-volume constraints set for the critical structures. This is reasonable since on-line ART effectively eliminates the inter-fractional variations in the patient anatomy, allowing more precise targeting of the PTV and avoiding the surrounding critical structures. This was reflected in the relatively small margin of 5 mm used to obtain the PTV from the CTV. Another reason for increasing the prescribed dose was to have treatment planning goals that were more challenging for our DAO algorithm.

To further challenge the DAO algorithm, one additional dose-volume constraint was placed on the bladder and the rectum: no more than 5% of the volume could receive a dose that exceeded 85 Gy. Finally, the 50% volume constraint for the bladder and the rectum was removed from the optimization since it was automatically attained without a constraint. The prescribed dose to the PTV and the dose-volume constraints used in this study are summarized in Table 4.1. The RTOG 0415 Prostate IMRT Protocol constraints are shown in the parentheses for comparison purposes. According to the RTOG 0415 Prostate IMRT Protocol, the maximum dose to the PTV could be up to 10% higher than the prescribed dose. This was classified as a minor variation and the plan was still considered to be acceptable. Also, if the prescription isodose surface covered between 95% and 98% of the PTV the coverage was again classified as a minor variation and the plan was acceptable.

| Normal Organ Constraints | No more than 5% volume receives dose that exceeds | No more than 15% volume receives dose that exceeds | No more than 25% volume receives dose that exceeds | No more than 35% volume receives dose that exceeds |
|---|---|---|---|---|
| Bladder | 85 Gy (n.a.) | 80 Gy (80 Gy) | 75 Gy (75 Gy) | 70 Gy (70 Gy) |
| Rectum | 85 Gy (n.a.) | 75 Gy (75 Gy) | 70 Gy (70 Gy) | 65 Gy (65 Gy) |
| Prescription | Minimum PTV dose (encompassing > 98% of PTV) | Maximum dose to PTV (No variations = 83.8 Gy +7%) | Maximum dose to PTV (Minor variations = 83.8 Gy +10%) | Maximum dose to PTV (Major variations > 83.8 Gy +10%) |
| PTV | 83.8 Gy (73.8 Gy) | 89.7 Gy (79.0 Gy) | 92.2 Gy (81.2 Gy) | > 92.2Gy (>81.2 Gy) |

*Table 4.1: The prescribed dose to the PTV and the dose-volume constraints for the bladder and the rectum used in this study. The RTOG 0415 Prostate IMRT Protocol constraints are shown in parentheses, for comparison purposes.*

To control the dose delivered to the unspecified normal tissue around the PTV, a "shell" was created around the PTV and included in the optimization. The shell was placed 5 mm from the PTV and surrounded the entire PTV. A 5 mm distance from the PTV was chosen based on our previous experience with IMRT planning. The purpose of the shell was to eliminate hot spots outside the target. The maximum dose allowed to the shell was set to 80 Gy. Also, for the purposes of optimization, the rectum and bladder were truncated so that they extended only up to 1.5 cm superior to the PTV. The lower portion of the rectum was also truncated so that it extended down to the bottom of the CTV, since the anal canal extended caudal to the coplanar fields. The anal canal, the rest of the bladder and rectum received negligible dose, since they were sufficiently distant from the field of view of the 7 coplanar beams, and thus were not included in the optimization. By truncating the rectum and bladder, the number of voxels used in the optimization was decreased, which resulted in faster optimization. Also, by excluding the portions of the bladder and rectum that received negligible dose, the dose-volume constraints

specified in Table 4.1 became harder to meet, thus further challenging the DAO algorithm. The shell and the truncated bladder and rectum are shown in Figure 4.3.

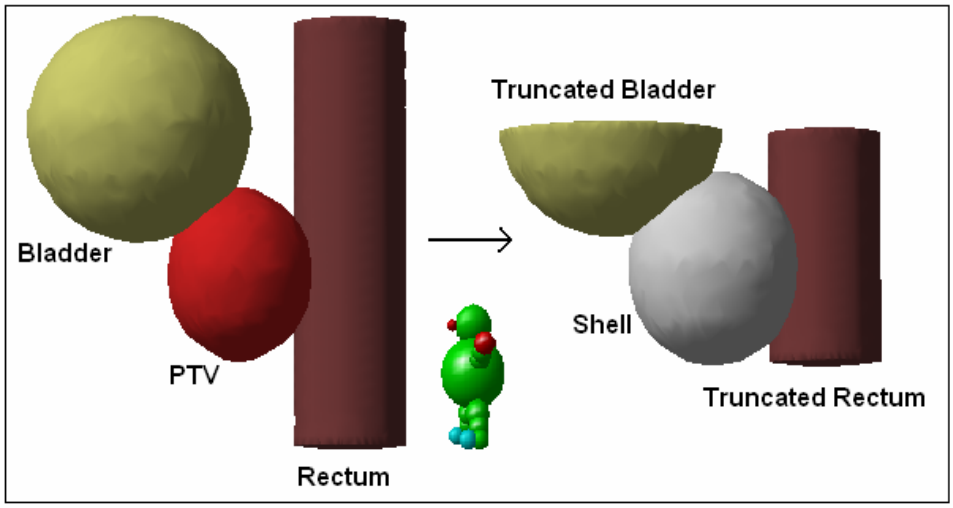


*Figure 4.3: For the purposes of plan optimization the rectum and bladder were truncated. Also, the shell was placed 5 mm from the PTV and surrounded the entire PTV. Placing a limit on the maximum dose to the shell in the optimization eliminated hot spots outside the PTV.*

### 4.2.4 Non-Adapted Original Plan

First, the consequence of using the original treatment plan for the four deformed anatomies was investigated. This was done to assess and quantify the degree of plan quality deterioration when no adaptation of the original treatment plan was performed.

### 4.2.5 Original Plan Adaptation

To correct for the original plan quality deterioration caused by the anatomical deformation, the original treatment plan was adapted. For each deformed anatomy, the original treatment plan was used as the "starting point" for the optimization. The aperture shapes and weights optimized for the original anatomy were used as the starting point (Figure 4.4a) instead of starting the optimization with aperture shapes conformed to the BEV of the deformed PTV and equal aperture weights (Figure 4.4b).

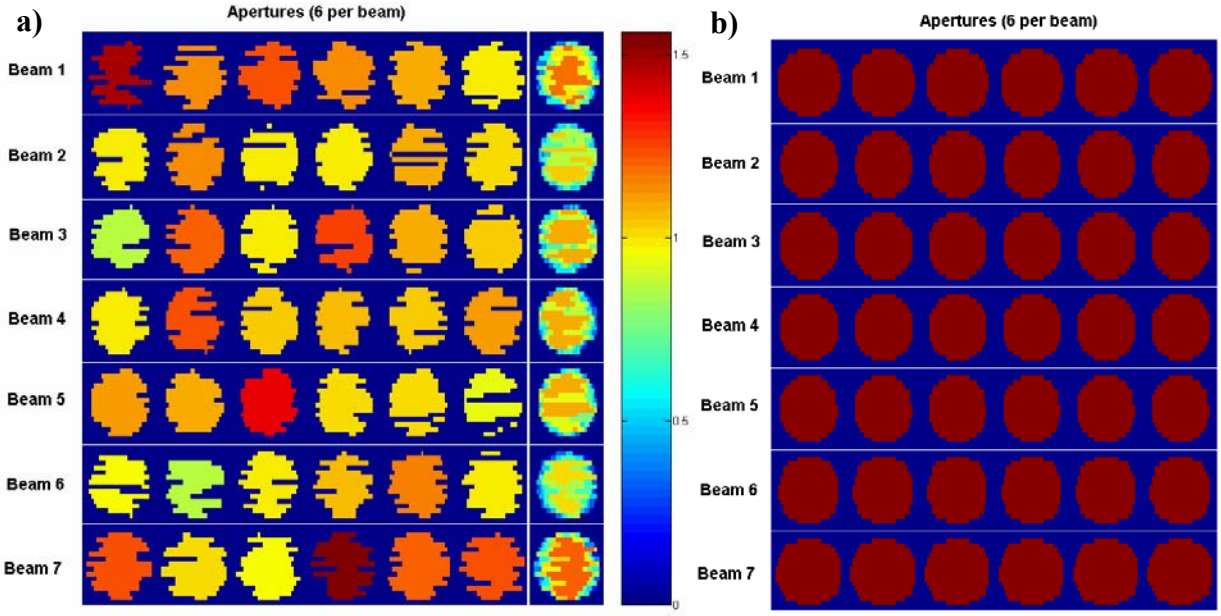


*Figure 4.4: a) The aperture shapes and weights (color bar) optimized for the original anatomy were used as the starting point for optimization instead of b) starting the optimization with aperture shapes conformed to the BEV of the deformed PTV and equal aperture weights.*

The hypothesis behind using the original treatment plan as the starting point for the optimization is illustrated in Figure 4.5. The original treatment plan is assumed to be the optimal, or very close to the optimal, treatment plan for the original anatomy. This corresponds to the bottom of the objective function well (Figure 4.5a). The anatomical deformation alters the objective function well shape, displacing the original treatment plan from the optimal configuration to some higher objective function value (Figure 4.5b). However, if the anatomical deformation is not too large the original plan may still be in the proximity of the new objective function minimum. In this case, the original treatment plan can be quickly adapted and returned to the bottom of the new objective function well (Figure 4.5c).

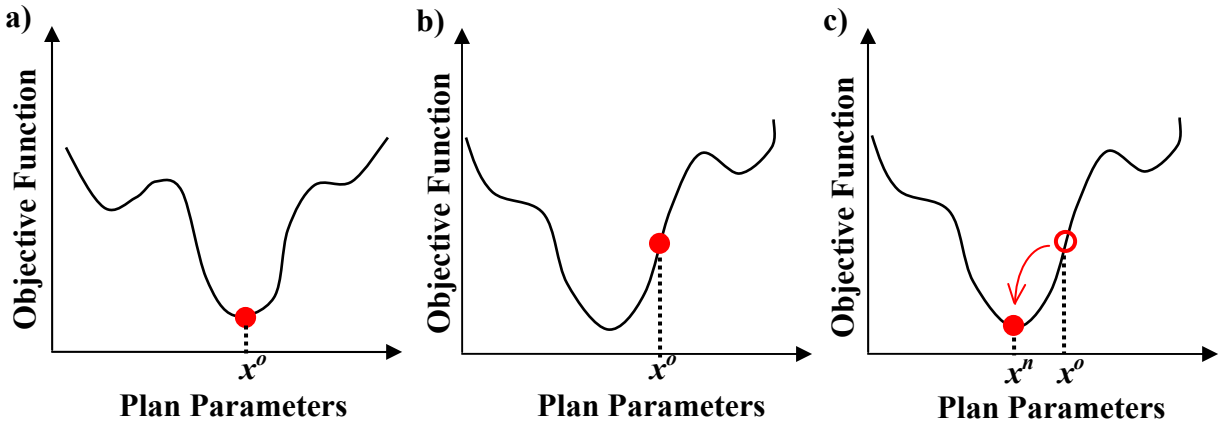


*Figure 4.5: a) The original treatment plan is assumed to be the optimal treatment plan, $x^o$, for the original anatomy. b) The anatomical deformation displaces the original treatment plan from the optimal position. If the anatomical deformation is not too large the original plan may still be in the proximity of the new objective function minimum. c) In this case, the original treatment plan can be quickly adapted and returned to the bottom of the new objective function well, $x^n$.*

To test the above hypothesis the original plan adaptation was compared to the complete plan regeneration (Figure 4.4b) for each deformed anatomy. The comparison was done in terms of the time required to arrive at the clinically acceptable plan. Our objective function was constructed so that it approached zero as the plan approached the constraints specified in Table 4.1. The objective function value of exactly zero indicated that all of the constraints were completely satisfied. However, the plans were generally clinically acceptable before the objective function value of zero was reached. As mentioned earlier, according to the RTOG 0415 Prostate IMRT Protocol, the maximum dose to the PTV could be up to 10% higher than the prescribed dose and the prescription isodose surface could cover between 95% and 98% of the PTV. Such a plan would be classified as 'a minor variation' and still be clinically acceptable. To generate the original treatment plan, the optimization was performed until the objective function reached a plateau. The minimum objective function value was around 0.5, indicating that the constraints in Table 1 were almost entirely met. However, for the plan adaptation, in the interest of reducing adaptation time, the optimization was stopped as soon as the plan became clinically acceptable. In other words, the plan adaptation was performed until the constraints for critical structures specified in Table 4.1 were completely met and the PTV coverage was classified as a minor deviation. The reasoning behind this approach is that the clinical benefits

of continuing plan adaptation until the objective function asymptotically reaches a plateau would be negated by the considerable extension of the treatment time. In other words, the marginal improvement of plan quality gained by continuing plan adaptation past a certain threshold is not worth the risk associated with considerably extended treatment time (e.g. increased susceptibility to intra-fractional patient motion).

Examination of the relationship between the plan's objective function value and the plan quality revealed that the treatment plan became clinically acceptable once the objective function value of 10 was reached. This was the case for all deformed anatomies. Investigation of using different objective function thresholds (1 to 10) to define the clinically acceptable plan revealed that the overall findings of this study are insensitive to the exact value of the threshold used. For the purposes of this study, treatment plan adaptation was performed until a threshold objective function value of 5 was reached, at which point a plan was considered clinically acceptable. Due to the random nature of the DAO algorithm, each optimization was slightly different from the previous one. To obtain the average time required to arrive at the clinically acceptable plan for each deformed anatomy, the original plan adaptation and the complete plan regeneration were both repeated ten times.

### 4.2.6 Accelerated Plan Adaptation

In the previous section, the adaptation of the original plan was performed using the same DAO algorithm that was used to generate the original treatment plan. In this section, the possibility of accelerating the plan adaptation by modifying the DAO algorithm is investigated. The motivation behind accelerating plan adaptation is an extension of the hypothesis made earlier (Figure 4.5). If the anatomy deformation is not too large the original plan may still be in the proximity of the new objective function minimum. In that case, it might be beneficial to reduce or constrain the allowed search space in order to further accelerate the optimization towards the new objective function minimum. Based on this hypothesis, three different methods of reducing the search space were investigated.

## 1. Reduced Leaf Step Size

As described in Section 2.4.4, the optimization starts with a random selection of an aperture. Once the aperture is selected, a random selection of either MLC leaf or the aperture weight is performed. Once the variable to be changed is selected, a random change is applied to that variable. The change is randomly sampled from the pre-specified range of allowed changes. The maximum change decreases according to Equation 2.12. If the selected variable is an MLC leaf then $M_o$ in Equation 2.12 is the initial maximum leaf step size and $R^{STEP}$ defines the rate at which the maximum leaf step size $M(n_{succ})$ decreases as a function of the number of successful changes, $n_{succ}$.

During the original plan optimization the initial maximum leaf step size, $M_o$, was set to 7.5 cm. This value was chosen to allow each leaf to move to any position across the field in each iteration. It corresponds to the average initial aperture size. However, during treatment plan adaptation it might not be desirable to have such drastic leaf steps. To reduce the search space, the initial maximum leaf step size was reduced (Figure 4.6). For each deformed anatomy, the plan adaptation was repeated with ten reduced values of $M_o$: 0.75 cm, 1.5 cm, 2.25 cm, etc. to 7.5 cm.

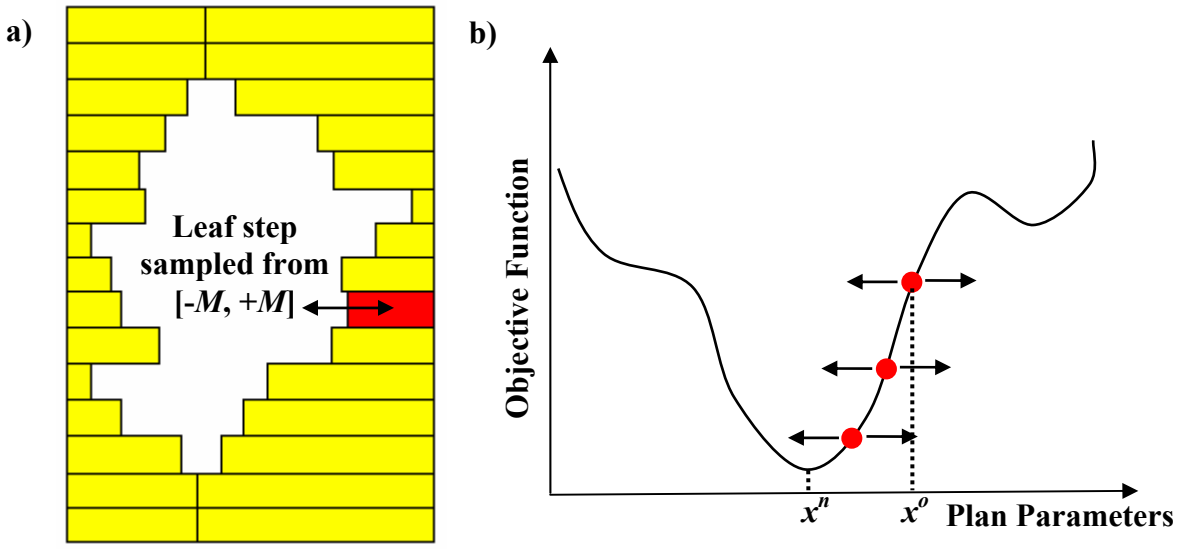


*Figure 4.6: a) At each iteration the leaf step size is randomly sampled from [-M, +M] where M is given by Equation 2.12. b) The leaf step size was reduced during plan adaptation to shrink the search space in an attempt to accelerate plan adaptation from the original plan, $x^o$, towards the new optimal solution, $x^n$.*

In addition to reducing the initial maximum leaf step size, $M_o$, the effect of using different $R^{STEP}$ rates was also investigated. Since $R^{STEP}$ defines the rate at which the maximum leaf step size decreases as a function of the number of successful changes (Equation 2.12), larger values of $R^{STEP}$ result in faster decrease of the maximum leaf step size. During the original plan optimization $R^{STEP}$ was set to 0.1. This value was selected based on our prior experience with complete plan regeneration, where the optimization starts with aperture shapes conformed to the BEV of the PTV and equal aperture weights (Figure 4.4b). However, for plan adaptation it might be advantageous to increase the value of $R^{STEP}$, which results in further reduction of the search space. For each reduced value of $M_o$, the plan adaptation was repeated with five different values of $R^{STEP}$: 0.1, 0.2, 0.3, 0.4 and 0.5 (Figure 4.7).

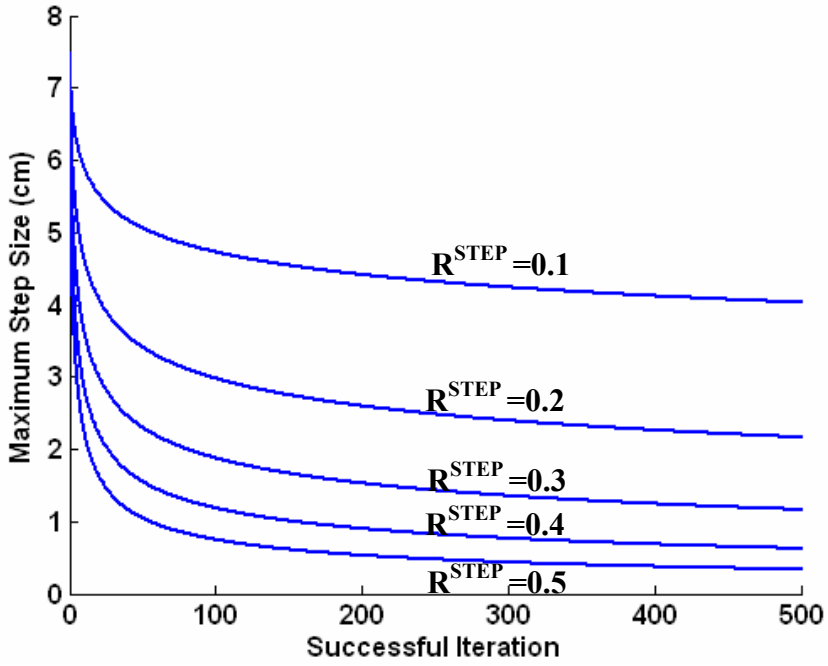


*Figure 4.7: During the original plan optimization $R^{STEP}$ was set to 0.1. However, for plan adaptation it might be advantageous to increase the value of $R^{STEP}$, which results in faster decrease of the maximum leaf step size and further reduction of search space. Plan adaptation was repeated with five different values of $R^{STEP}$: 0.1, 0.2, 0.3, 0.4 and 0.5, for each reduced value of the initial maximum leaf step size, $M_o$ (only the curves for $M_o$ = 7.5 cm are displayed).*

## 2. Constrained Leaf Range

Another approach to reduce the search space during plan adaptation was to constrain the allowed leaf range. During the original plan optimization the only constraint on the allowed leaf range was that the leaves were not allowed to move outside the BEV of the PTV. However, for plan adaptation it might be beneficial to put an additional constraint on the allowed leaf range. If the original plan is in the proximity of the new objective function minimum, only small changes in the leaf positions might be necessary to re-optimize the plan. Restricting the DAO algorithm to explore only a portion of the search space around the original treatment plan reduces the search space and might accelerate plan adaptation (Figure 4.8). For each deformed anatomy, the plan adaptation was repeated with ten different constraints on the allowed leaf range: 0.5 cm, 1.0 cm, 1.5 cm, etc. to 5.0 cm. The maximum leaf step size was varied such that the leaf could move to any position across the range in each iteration.

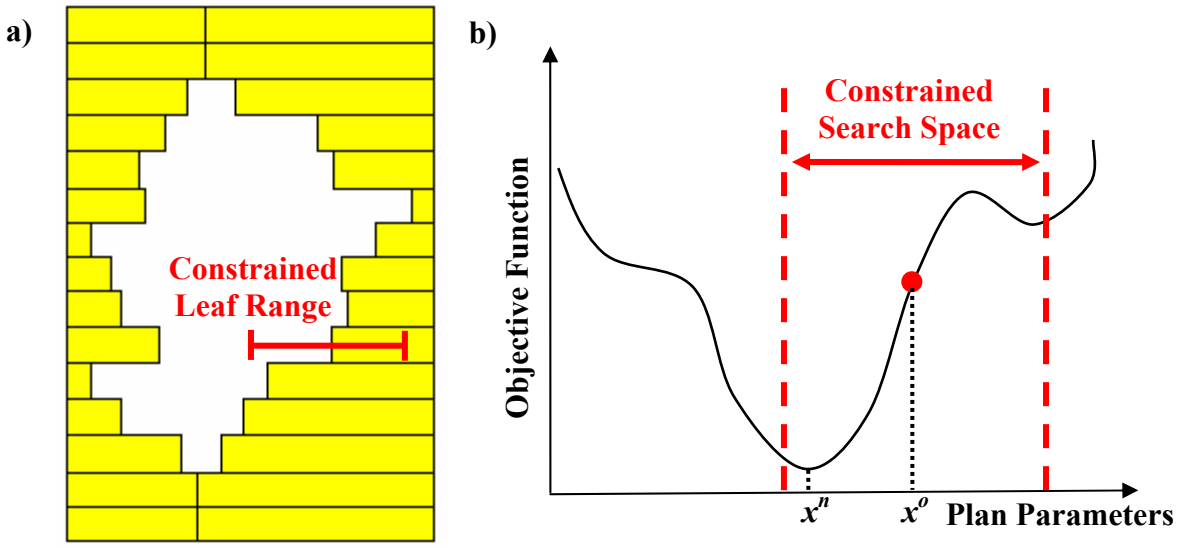


*Figure 4.8: a) The allowed range for each MLC leaf was constrained during plan adaptation. b) Restricting the DAO algorithm to explore only a portion of the search space around the original treatment plan, $x^o$, reduces the search space and might accelerate plan adaptation towards the new optimal solution, $x^n$.*

In both of these approaches to accelerate plan adaptation (Reduced Leaf Step Size and Constrained Leaf Range) the search space was reduced by modifying the change of aperture leaf

positions during the optimization. The effect of modifying the change of aperture weights in addition to aperture leaf positions was also investigated. It was found that the plan adaptation was not sensitive to constraining and modifying the change of aperture weights. Therefore, the aperture weight changes in the modified DAO algorithm remained the same as in the original DAO algorithm.

### 3. Reduced System Temperature

The third approach to accelerate plan adaptation was to reduce the system temperature by decreasing the initial system temperature, $T_o$, and increasing the rate of cooling, $R^{COOLING}$. System temperature determines the probability of accepting variable changes that result in increase of the cost value, via Equation 2.13. During the original plan optimization, the system temperature was initially relatively high so that the changes that increased the cost value were allowed more frequently. As the optimization progressed, and the system temperature decreased, the probability of accepting such changes decreased. Therefore, the initial search space was relatively large, with high probability of accepting changes which result in increase of the cost value. This mechanism allowed the optimization to escape from local minima and find the global minimum. As the optimization progressed and the system temperature decreased, the search space was gradually reduced allowing the algorithm to converge to a final solution. The system temperature was reduced according to the cooling schedule, Equation 2.14. During the original plan optimization, the initial system temperature was set to $T_o$ = 0.005 and the rate of cooling was set to $R^{COOLING}$ =0.25. Both of these values were selected based on our prior experience with complete plan regeneration, where the optimization starts with aperture shapes conformed to the BEV of the PTV and equal aperture weights.

However, if the anatomical deformation is not too large and the original plan stays in the vicinity of the global minimum then the optimization might be able to identify this global minimum without the need to explore a very large search space. The risk of getting trapped in a local minimum is lower than for complete plan regeneration. Therefore, to accelerate plan adaptation search space was reduced by performing the optimization with lower initial system temperature, $T_o$, and increased rate of cooling, $R^{COOLING}$. The initial temperature was reduced to $T_o$: 0.004, 0.003, 0.002, 0.001 and 0.000. Having the system temperature equal to zero effectively rejected all the variable changes that resulted in increase of the cost value. Only the changes that

lowered the value of the objective function were accepted. For each reduced value of $T_o$, except $T_o$ = 0, plan adaptation was repeated with five different values of $R^{COOLING}$: 0.25, 0.35, 0.45, 0.55 and 0.65. Figure 4.9a displays the system temperature as a function of the number of successful changes, $n_{succ}$, for original plan optimization ($T_o$=0.005 and $R^{COOLING}$=0.25) and accelerated plan adaptations (only four combinations of $T_o$ and $R^{COOLING}$ are shown for clarity). Figure 4.9b displays the corresponding probability of accepting a variable change which increases the cost value by 0.001, calculated via Equation 2.13 with $\Delta f$=0.001.

In this approach, the initial maximum leaf step size, $M_o$, and the rate at which the maximum leaf step size decreased as a function of the number of successful changes, $R^{STEP}$, were selected based on the results from the "Reduced Leaf Step Size" Section (discussed in the "Results and Discussion" Section). The leaf range was not constrained.

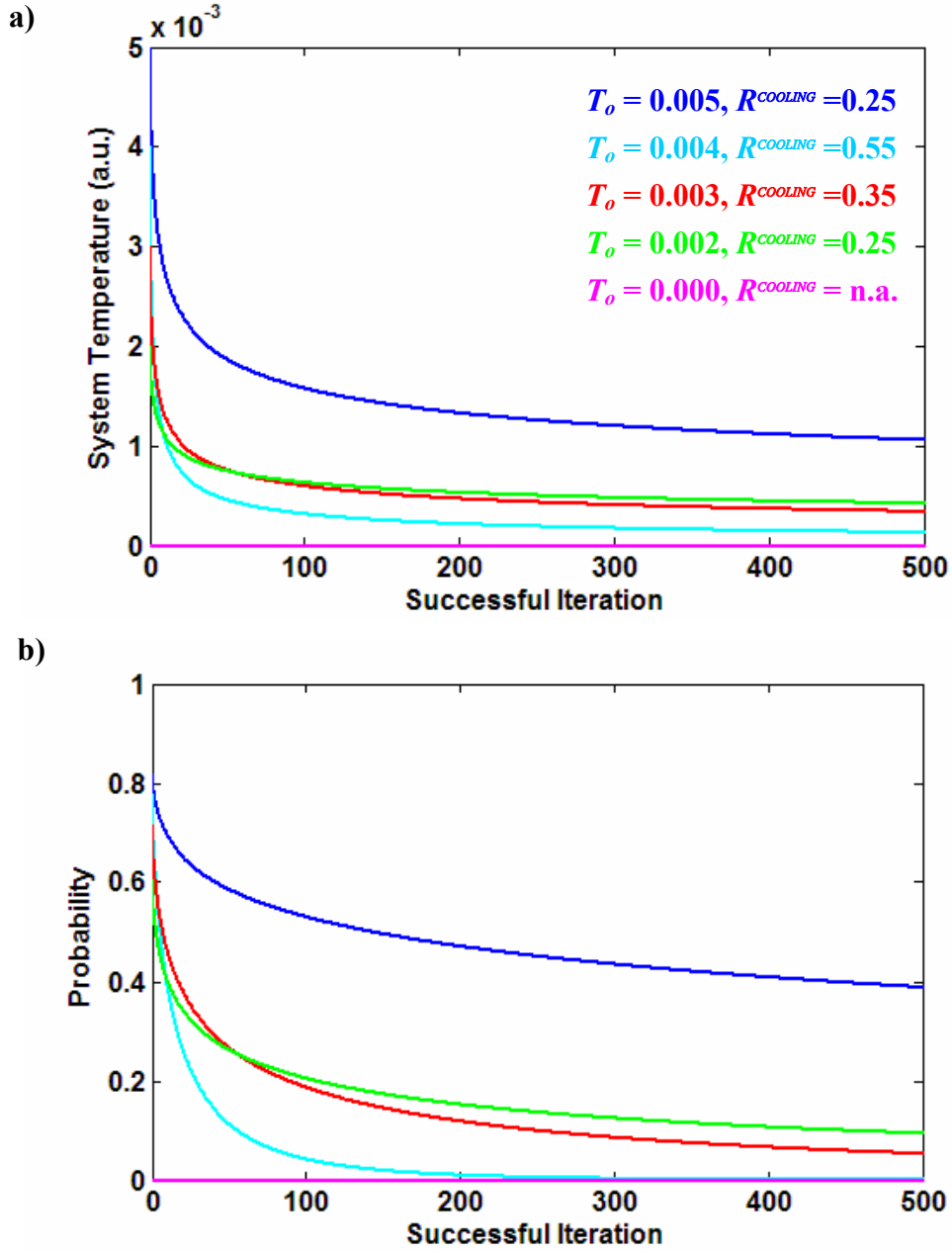


*Figure 4.9: a) System temperature as a function of the number of successful iterations for complete plan regeneration ($T_o$=0.005, $R^{COOLING}$=0.25) and accelerated plan adaptations (only four combinations of $T_o$ and $R^{COOLING}$ are shown for clarity). b) The corresponding probability of accepting a variable change which increases the cost value by 0.001, calculated via Equation 2.13 with $\Delta f$=0.001.*

In addition to reducing the system temperature, the search space was further reduced by modifying the manner in which the DAO algorithm cycled through the aperture weights and

MLC leaf positions. During the original plan optimization, the algorithm randomly selected apertures. Once an aperture was selected, a random selection of either MLC leaf or the aperture weight was performed. Once the variable to be changed was selected, a random change was applied to that variable. The main reason for having so many "random selections" incorporated into the original plan optimization was to broaden the search space. To reduce the search space during plan adaptation, the random component of the optimization was taken out. Instead, the algorithm cycled through all the apertures sequentially (Figure 4.10). Once the aperture was selected, the algorithm cycled through the aperture leaves and aperture weight in the order shown in Figure 4.10. Once the variable to be changed was selected, a change was applied to that variable. Instead of applying a random change as in the original optimization, the direction of the change was based on the success of the previous iteration. For example, if in the previous iteration it was found that MLC leaf should move inward for the objective function to decrease, then the current leaf was first moved inward. The reasoning behind this approach was that the MLC leaves that are adjacent to each other will more probably need to move in the same direction to compensate for the anatomical deformation.

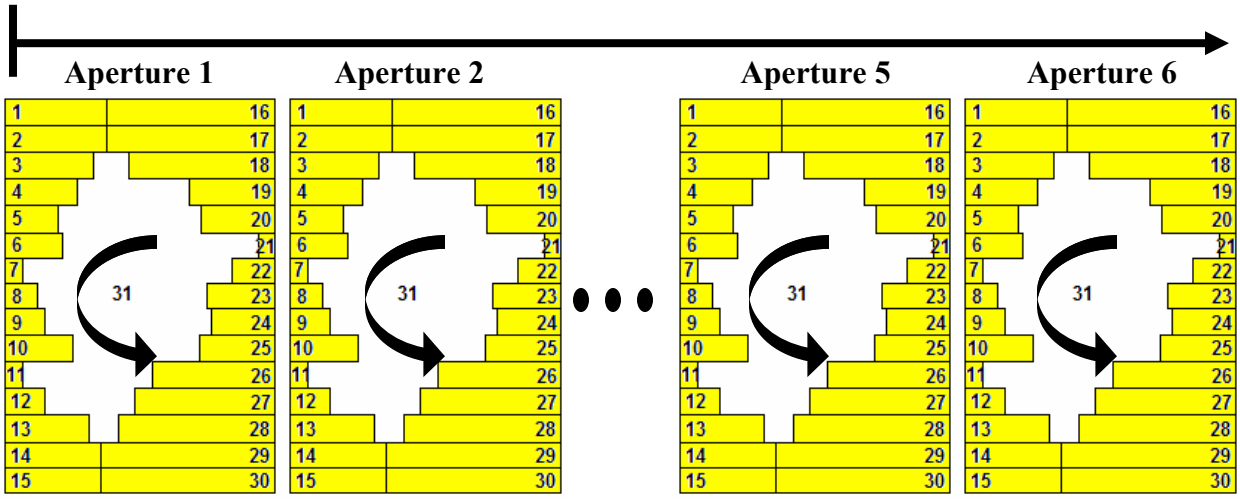


*Figure 4.10: To reduce the search space during plan adaptation, the random component of the optimization was taken out. The DAO algorithm cycled sequentially through all the apertures. Once the aperture was selected, the algorithm cycled through the aperture leaves and aperture weight in a predefined order, indicated by numbers.*

### 4.2.7 Statistical Analysis

Mean times to achieve clinically acceptable plan were compared using a two-tailed unpaired student t-test. Statistical analysis was performed in MATLAB (Mathworks Inc., Natick, MA).

## 4.3 Results and Discussion

### 4.3.1 Original Treatment Plan

The DVHs for the original treatment plan used for the original anatomy are shown in Figure 4.11. From Figure 4.11 it can be seen that the dose-volume constraints specified in Table 4.1 were satisfied. Also, the constraint on the maximum dose to the shell (80 Gy) was satisfied.

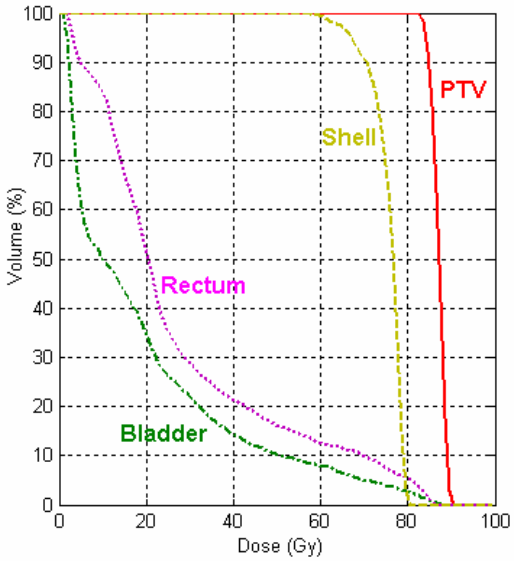


*Figure 4.11: DVHs for the original treatment plan used for the original anatomy. The shown DVHs are for the truncated rectum and the truncated bladder (as is the case for all the subsequent DVH plots presented in the thesis). Thus, 100% of the bladder volume corresponds to the total volume of the truncated bladder, and 100% of the rectum volume corresponds to the total volume of the truncated rectum.*

### 4.3.2 Non-Adapted Original Plan

Figure 4.12 displays DVHs for the non-adapted original plan used for the four deformed anatomies. As expected, the DVHs for critical structures and the PTV deteriorated with increased anatomy deformation. The minimum dose to PTV decreased from 83 Gy for the original anatomy to 80 Gy, 72 Gy, 61 Gy and 41 Gy for the 0.25 cm, 0.50 cm, 0.75 cm and 1.00 cm anatomical deformations, respectively. The overall dose to the rectum increased as the anatomy became more deformed, violating the dose-volume constraints specified in Table 4.1. Deformations of 0.50 cm and 0.75 cm violated the 5% and 15% volume constraints for rectum, and deformation of 1.00 cm violated the 5%, 15% and 25% volume constraints for rectum. The constraint on the maximum dose to the shell (80 Gy) was violated for all four deformed anatomies. The dose-volume constraints for the bladder were not violated for any deformation.

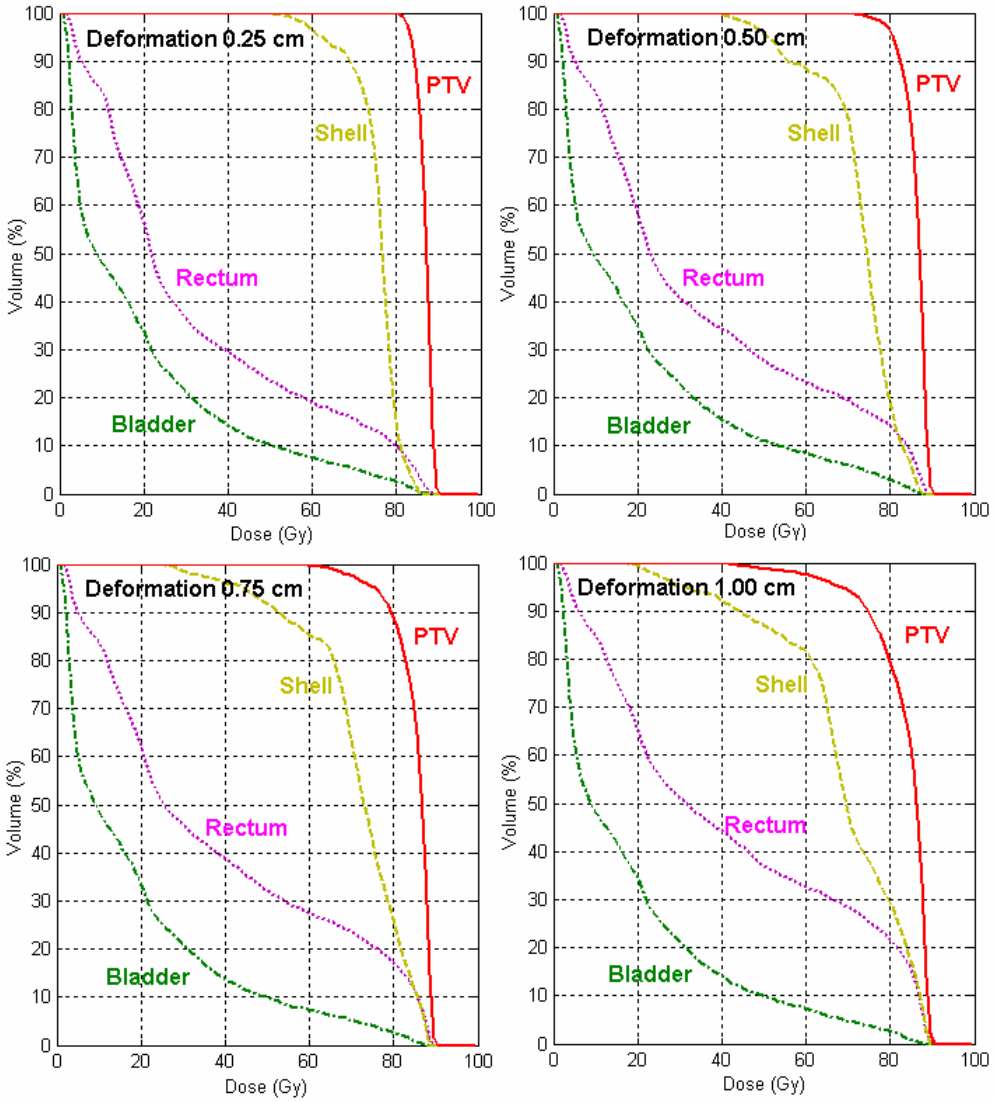


*Figure 4.12: DVHs for the non-adapted original treatment plan used for the four deformed anatomies.*

### 4.3.3 Original Plan Adaptation

To correct for the original plan quality deterioration caused by the anatomy deformations the original treatment plan was adapted. The original plan adaptation was compared to the complete plan regeneration in terms of the average time required to arrive at the clinically acceptable plan. Due to the random nature of the DAO algorithm, the original plan adaptation and the complete plan regeneration were repeated ten times, for each deformed anatomy (Figure 4.13).

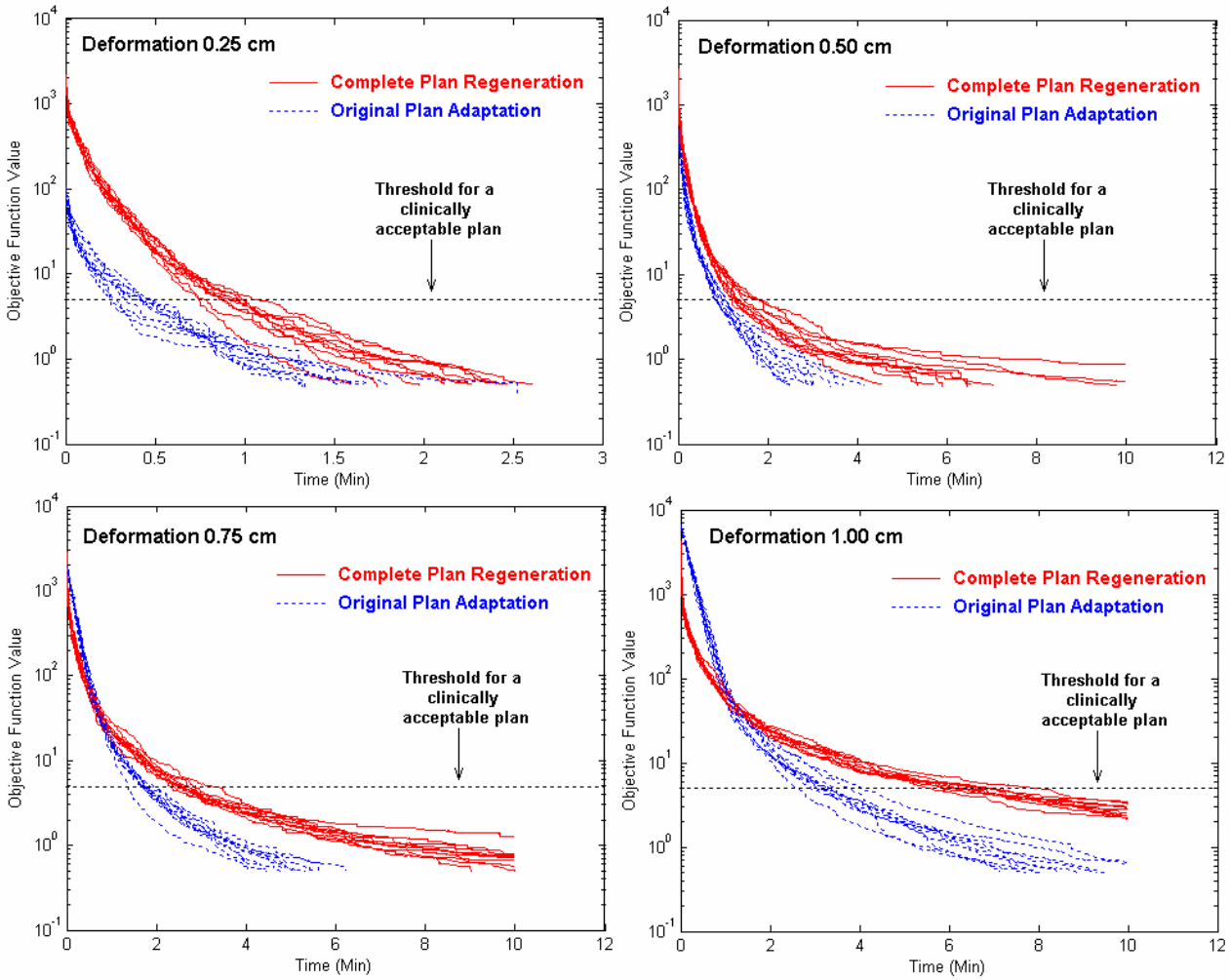


*Figure 4.13: The objective function value versus time for the original plan adaptation and the complete plan regeneration.*

For all four deformed anatomies, the original plan adaptation arrived at the clinically acceptable plan faster than the complete plan regeneration. The average time needed to arrive at the clinically acceptable plans is shown in Figure 4.14. The difference between the average time for the original plan adaptation and complete plan regeneration was statistically significant ($p < 0.001$) for all four deformed anatomies. The error bars shown represent the variations due to the stochastic nature of individual optimizations and are equal to one standard deviation. Figure 4.14 indicates that the average time required to arrive at the clinically acceptable plan by adapting the original plan was roughly half of the time needed for complete plan regeneration.

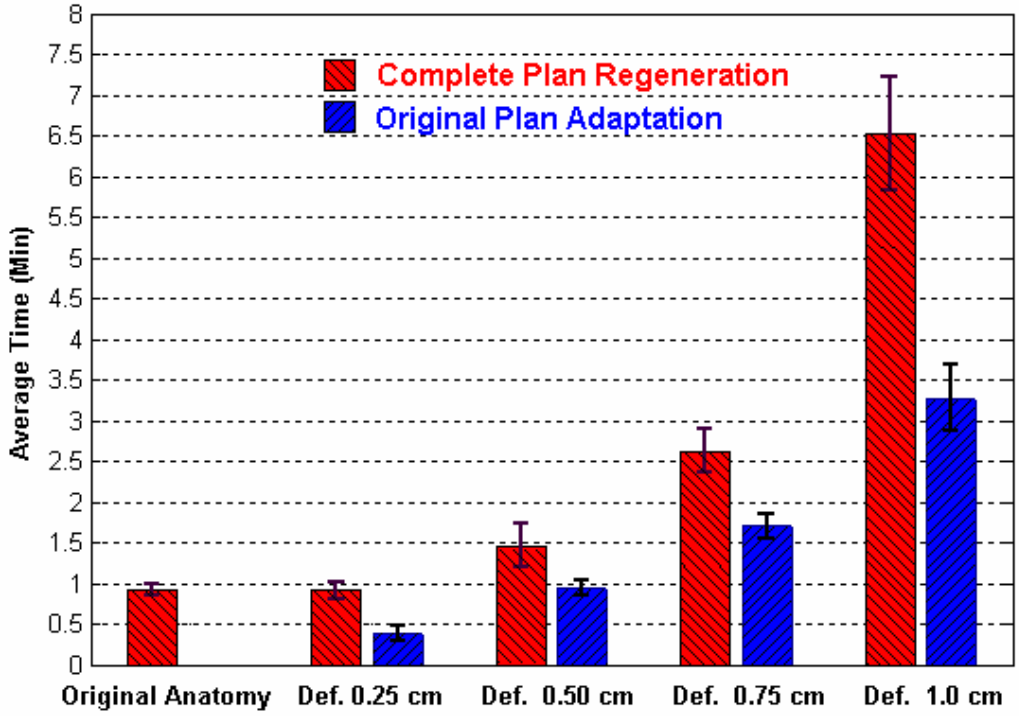


*Figure 4.14: The average times needed for the complete plan regeneration and the original plan adaptation to arrive at the clinically acceptable plan. The error bars shown represent one standard deviation.*

Figure 4.13 shows that as the anatomical deformation increased the initial objective function value (for t=0) for the original plan adaptation increased, as expected. It is interesting to note that for the anatomical deformations of 1.0 cm the initial objective function value was higher for the original plan adaptation than for the complete plan regeneration. The initial objective function values for the original plan adaptation and the complete plan regeneration were 5764 and 4655, respectively. This seems to indicate that the last anatomical deformation was so large that the original treatment plan was initially further away from the optimal solution than the initial plan for the complete plan regeneration, which had aperture shapes simply conformed to the BEV of the deformed PTV and equal aperture weights. However, even for the largest anatomical deformation the optimization curves 'crossed-over' (Figure 4.13) and the original plan adaptation still arrived at the clinically acceptable plan faster. One possible explanation for this phenomenon is illustrated in Figure 4.15.

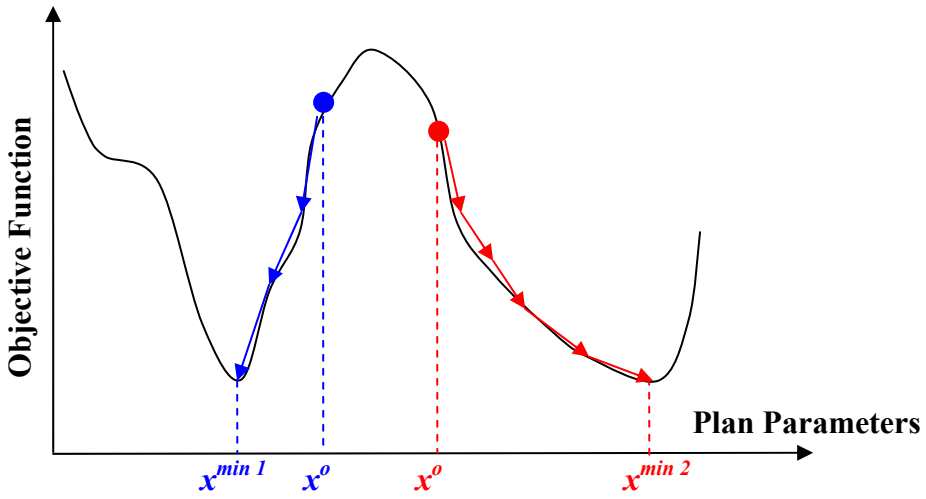


*Figure 4.15: Even though the original treatment plan (represented by a blue marker) has a higher initial objective function value than the initial plan for the complete plan regeneration (red marker), the original treatment plan is still "closer" to one of the local minima, $x^{min\ 1}$. This appeared to be the case for the last anatomical deformation for which the "cross-over" of the optimization curves occurred.*

Figure 4.15 implies that there exist multiple local minima within the search space with similar objective function values but different sets of treatment plan parameters (i.e. aperture shapes and aperture weights). The evidence for such multiple local minima was observed for all four anatomical deformations. The final aperture shapes and weights of different treatment plans with similar objective function values are given in Appendix B. Figure B.1 shows aperture shapes and weights of four different treatment plans for the anatomical deformation of 1.00 cm. Even though the objective function values are all close to each other (0.484, 0.406, 0.464 and 0.466, respectively) the aperture shapes and weights are significantly different between the four plans. Similar results were also observed for the other three anatomical deformations (not shown).

Figure 4.16 displays the DVHs that correspond to the objective function value of 5, a threshold for a plan to be considered clinically acceptable. From Figure 4.16, it can be seen that these plans indeed satisfied the dose-volume constraints specified in Table 4.1.

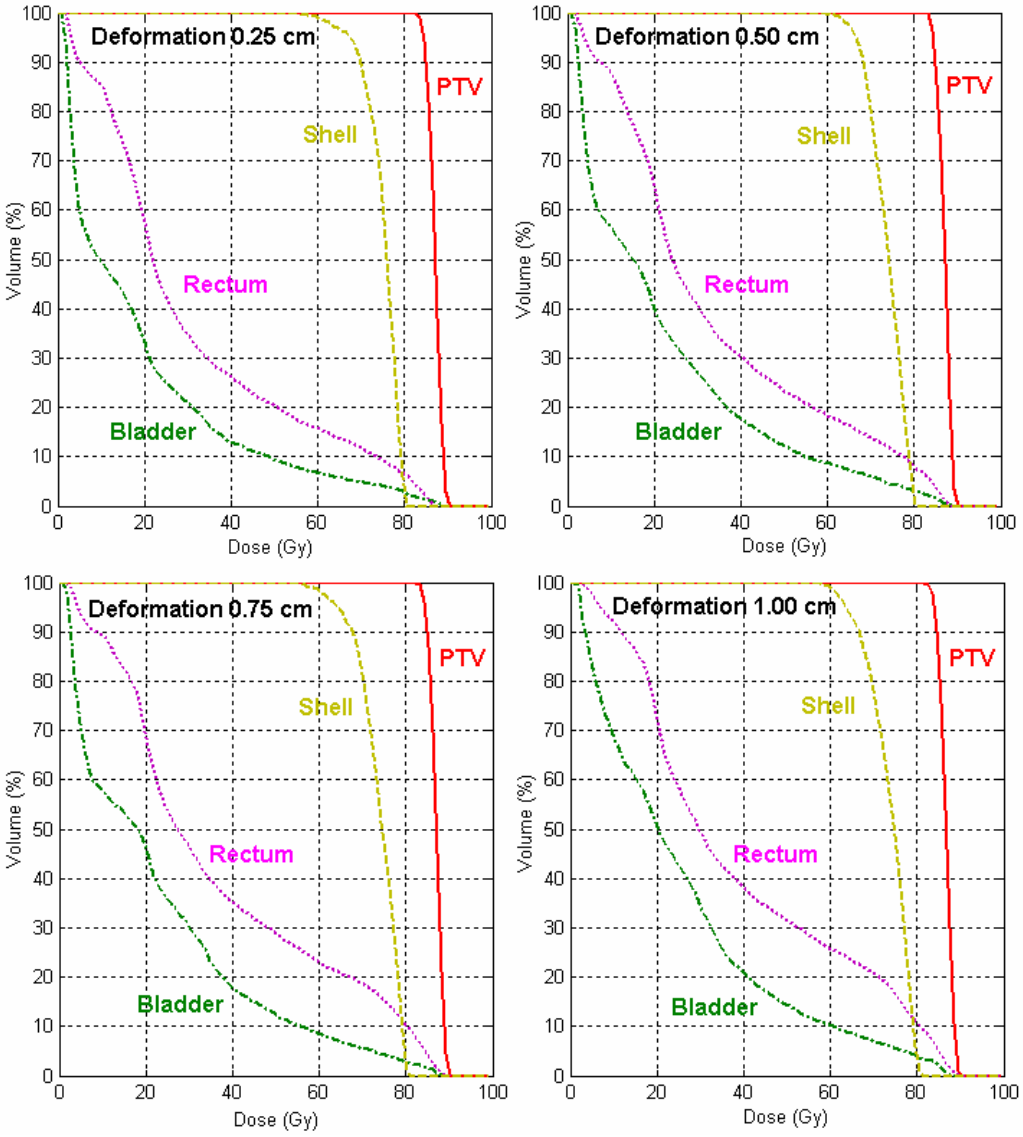


*Figure 4.16: DVHs that correspond to the objective function value of 5, a threshold for a plan to be considered clinically acceptable. The plans generated by the complete plan regeneration and the plans generated by the original plan adaptation have indistinguishable DVHs at this scale, so only the DVHs for original plan adaptation are shown.*

It should be noted that the times shown in Figure 4.14 are for the re-optimization process only. However, prior to the start of the re-optimization process the initial dose distribution needs to be determined, which also contributes to the total treatment time. For the deformed anatomies the initial dose distribution is the optimized (final) dose distribution for the original plan. Therefore, the dose distribution prior to the plan adaptation does not need to be explicitly re-

computed. The sampling points, points at which the dose is calculated, are simply reassigned based on the changes in the anatomy. For example, a voxel that was part of the PTV in the original anatomy might be reassigned to be a part of the rectum in the deformed anatomy. It should be noted that this reassignment of voxels, instead of dose re-computation, is possible due to the following two reasons. First, all the anatomy deformations are such that the displacement of the centroid of the PTV is negligible, so the position of the isocenter is kept the same as in the original plan. Second, the external contour is kept the same in the deformed anatomies as it was in the original anatomy. If the isocenter is shifted and/or the body contour is changed the dose prior to the adaptation would need to be re-computed. The dose re-computation takes approximately 10 seconds on a single 2.5 GHz personal computer, but could easily be parallelized and performed in a few seconds.

In this study, the PTV deformations were such that the displacement of the centroid of the PTV was less than 2 mm for all four deformed anatomies. The centroid displacement of the PTV was 0 mm, 0.5mm, 1 mm and 1mm for the four deformations, respectively. Given such a small displacement of the centroid of the PTV the patient would not be shifted to compensate for this displacement. However, for the cases where a displacement of the centroid of the PTV was significant, the patient would be first shifted and then the re-optimization of the apertures would commence. In addition to shifting the patient, another possible step prior to plan re-optimization could involve morphing the apertures of the original treatment plan, based on the daily anatomical deformations[51]. Using the morphed apertures as a starting point for plan adaptation could potentially bring the starting point even closer to the new objective function minimum and thus help speed-up plan adaptation. In this study, no initial aperture morphing was performed.

### 4.3.4 Accelerated Plan Adaptation

In the previous section, the adaptation of the original plan was performed using the same DAO algorithm that was used to generate the original treatment plan. In this section, the DAO algorithm was modified to further accelerate plan adaptation. The DAO parameters were modified in the three different methods to reduce the search space, as described previously. Due to the random nature of the DAO algorithm, the accelerated plan adaptation was repeated five times for each method to obtain the average time required to arrive at the clinically acceptable plan. The average times needed to arrive at the clinically acceptable plan using the original DAO

algorithm (original plan adaptation) and the modified DAO algorithm (accelerated plan adaptation) are shown in Figure 4.17. The average plan adaptation time is normalized by the complete plan regeneration average time for each deformed anatomy. This way the time efficiency of the plan adaptation can be directly compared to that of the complete plan regeneration.

Since the degree of patient anatomical deformation can be determined prior to the start of the plan adaptation, the optimal value for the DAO parameters (i.e. $M_o$, $R^{STEP}$, $T_o$, $R^{COOLING}$, constrained leaf range) can be selected prior to plan adaptation based on the anatomical deformation information. The amount of anatomical deformation can be determined by registering the planning image set (e.g. planning CT) with the daily image set (e.g. CBCT), obtained prior to the treatment. The patient would be 'categorized' based on the anatomical deformation prior to the start of the treatment plan adaptation and the optimal DAO parameters would be chosen for the accelerated plan adaptation. In consideration of this strategy for optimizing the adaptation, the average times for the accelerated plan adaptation displayed in Figure 4.17 correspond to the optimal values of the DAO parameters (given below) for each anatomical deformation.

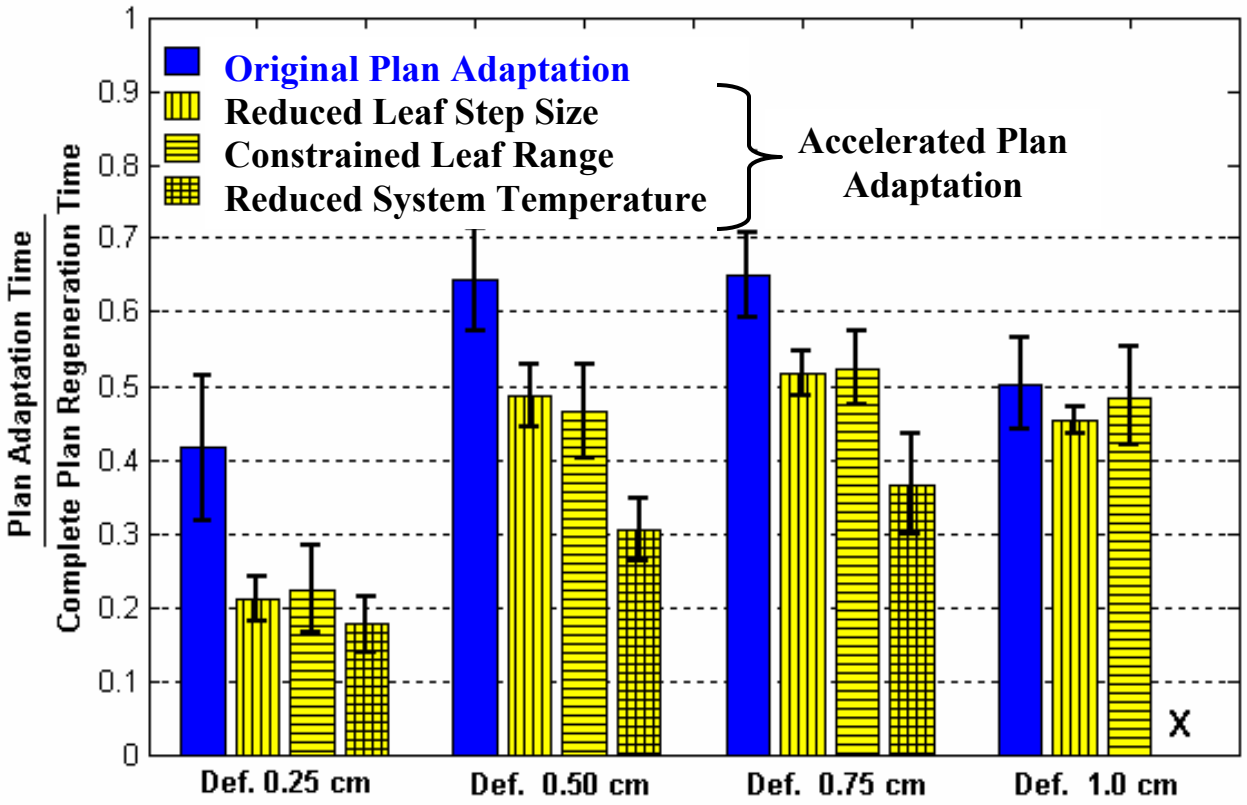


*Figure 4.17: The average normalized time needed for the Original Plan Adaptation and the Accelerated Plan Adaptation to arrive at the clinically acceptable plan. The average times for the Accelerated Plan Adaptation correspond to the optimal values of the DAO parameters for each anatomical deformation. The error bars shown represent one standard deviation. The "x" above the 1.00 cm deformation indicates that the Reduced System Temperature approach did not result in any time reduction over the Original Plan Adaptation. For this case, the Accelerated Plan Adaptation took longer than the Original Plan Adaptation to arrive at the clinically acceptable plan.*

### 1. Reduced Leaf Step Size

The optimal values for the initial maximum leaf step size, $M_o$, and the rate at which the maximum leaf step size decreased as a function of the number of successful changes, $R^{STEP}$, were:

1) $M_o$ = 1.5 cm, $R^{STEP}$ = 0.4 for 0.25 cm anatomical deformation
2) $M_o$ = 2.25 cm, $R^{STEP}$ = 0.4 for 0.50 cm anatomical deformation
3) $M_o$ = 3.0 cm, $R^{STEP}$ = 0.3 for 0.75 cm anatomical deformation
4) $M_o$ = 3.0 cm, $R^{STEP}$ = 0.2 for 1.00 cm anatomical deformation

The optimal value of $M_o$ increased and the optimal value of $R^{STEP}$ decreased as the degree of anatomical deformation increased. The average plan adaptation times with these optimal $M_o$ and $R^{STEP}$ values are shown in Figure 4.17. Using a two-tailed unpaired student t-test to compare the mean Reduced Leaf Step Size adaptation time to the mean Original Plan Adaptation time, it was found that for the first three deformed anatomies there was a significant improvement ($p < 0.001$). For the fourth deformed anatomy the improvement was not significant ($p = 0.11$).

### 2. Constrained Leaf Range

The optimal values for the constrained leaf range were:

1) Leaf range of 1.5 cm for 0.25 cm anatomical deformation
2) Leaf range of 3.5 cm for 0.50 cm anatomical deformation
3) Leaf range of 3.5 cm for 0.75 cm anatomical deformation
4) Leaf range of 5.0 cm for 1.00 cm anatomical deformation

The optimal value of the constrained leaf range increased as the degree of anatomical deformation increased. The average plan adaptation times with the optimal leaf range value are shown in Figure 4.17. Using a two-tailed unpaired student t-test to compare the mean Constrained Leaf Range adaptation time to the mean Original Plan Adaptation time, similar improvement was observed as for the previous approach. For the first three deformed anatomies there was a significant improvement ($p = 0.0016$ for anatomical deformation of 0.25 cm and $p <$

0.001 for anatomical deformations of 0.50 cm and 0.75 cm). For the 1.00 cm deformed anatomy the improvement was not significant (p = 0.63).

### 3. Reduced System Temperature

The optimal values for the initial system temperature, $T_o$, and the rate of cooling, $R^{COOLING}$, were:

1) $T_o = 0$, $R^{COOLING}$ = n.a for 0.25 cm anatomical deformation
2) $T_o = 0$, $R^{COOLING}$ = n.a for 0.50 cm anatomical deformation
3) $T_o = 0.002$, $R^{COOLING} = 0.35$ for 0.75 cm anatomical deformation
4) $T_o = 0.004$ cm, $R^{COOLING} = 0.35$ for 1.00 cm anatomical deformation

The optimal value of $T_o$ increased as the degree of anatomical deformation increased. The average plan adaptation times with these optimal $T_o$ and $R^{COOLING}$ values are shown in Figure 4.17. Using a two-tailed unpaired student t-test to compare the mean Reduced System Temperature adaptation time to the mean Original Plan Adaptation time, it was found that for the first three deformed anatomies there was a significant improvement ($p < 0.001$). Also, for the first three anatomical deformations this approach offered the largest improvement among the three approaches. For the anatomical deformation of 0.25 cm, the improvement between this approach and the first two approaches was not significant (p = 0.07 and p = 0.08, respectively). For the anatomical deformations of 0.50 cm and 0.75 cm, the improvement between this approach and the first two approaches was statistically significant ($p < 0.001$ for both approaches). For the anatomical deformation of 0.25 cm the optimization efficiency was increased by a factor of about 6 compared to the complete plan regeneration. For the anatomical deformations of 0.50 cm and 0.75 cm the optimization efficiency was increased by a factor of roughly 3 compared to the complete plan regeneration. However, for the last anatomical deformation, this approach did not result in any time reduction over the original plan adaptation. In fact, the accelerated adaptation took longer than the original plan adaptation to arrive at the clinically acceptable plan. One possible explanation is that the anatomical deformation of 1.00 cm was so great that the original plan was substantially displaced from the optimal solution. In this case it was not advantageous to reduce the search space during the plan adaptation. In other words, our original hypothesis did not hold for the 1.00 cm deformation. As mentioned earlier, the last deformation represents an

extreme case, which is rarely observed for the prostate, and demonstrates the limit at which this approach breaks down.

## 4.4 Conclusion

This study is the first investigation of applying DAO for on-line ART. An anatomical model representing a typical prostate anatomy was created. To simulate a range of clinically relevant anatomical deformations, four deformed anatomies were created by systematically deforming the original anatomy by various amounts (0.25, 0.50, 0.75 and 1.0 cm). A series of techniques was described in which the original DAO treatment plan was adapted in order to correct for the deterioration of dose distribution quality caused by the anatomical deformations. It was found that the average time needed for the original plan adaptation to arrive at the clinically acceptable plan was roughly half of the time needed for a complete plan regeneration, for all four anatomical deformations. Furthermore, through modification of the DAO algorithm the optimization search space was reduced and the plan adaptation was significantly accelerated. For the first anatomical deformation (0.25 cm) the optimization efficiency was increased by a factor of about 6 compared to the complete plan regeneration. For the 0.50 cm and 0.75 cm deformations the optimization efficiency was increased by a factor of roughly 3 compared to the complete plan regeneration. However, for the anatomical deformation of 1.0 cm the reduction of the search space during plan adaptation did not result in any efficiency improvement over the original (non-modified) plan adaptation. For our model, the anatomical deformation of 1.0 cm demonstrated the limit for which the plan adaptation could not be accelerated.

# Chapter 5
# Integrating Plan Adaptation and Radiation Delivery: Adaptive Radiation Delivery

## 5.1 Introduction

In the previous chapter, methods for accelerated plan adaptation (re-optimization) using the DAO technique were investigated. In an attempt to reduce the treatment time even further, a new approach to on-line ART is proposed and investigated in this chapter. The accelerated plan adaptation and radiation delivery are integrated together and performed concurrently. A fundamental advantage of combining plan adaptation and radiation delivery together is that radiation delivery can start almost immediately after daily imaging, based on the partially re-optimized plan. Most of the plan re-optimization is performed during the radiation delivery, so the time spent adapting the original plan does not significantly increase the overall treatment time (Figure 2.20c). This work was published together with the work from the previous chapter in the journal of Medical Physics[81].

## 5.2 Methods and Materials

A novel approach to on-line ART was investigated in which plan adaptation and radiation delivery were integrated together and performed concurrently. This approach is referred to as Adaptive Radiation Delivery (ARD), a term coined by our group. The work in this chapter directly follows the work from the previous chapter. As in Chapter 4, our DAO algorithm (Section 4.2.1) was used to create the original treatment plan (Section 4.2.3) for the original anatomy (Section 4.2.2). Four deformed anatomies were created by systematically deforming the original anatomy by various amounts (Section 4.2.2).

To accelerate plan adaptation for the 0.25, 0.50 and 0.75 cm deformed anatomies the original DAO algorithm was modified as described in Section 4.2.6.3 - Reduced System Temperature. The optimal values of the DAO parameters were used for each of the three deformed anatomies. The optimal values are given in Section 4.3.4.3. For the last anatomical deformation, the original (non-modified) DAO algorithm was used for plan adaptation since the reduction of the search space during the accelerated plan adaptation did not result in any significant efficiency improvement over the original plan adaptation (Figure 4.17). The plan adaptation procedure started as before. The DAO algorithm cycled through all the apertures sequentially. Once the aperture was selected, the algorithm optimized leaf positions and aperture weight in the order shown in Figure 4.10. For a treatment plan with seven beams and six apertures per beam there were 42 apertures in total. Instead of continuously cycling through all 42 apertures until the plan was completely adapted, the plan adaptation was performed for a pre-defined period of time, T. After time T the first aperture was 'fixed' and radiation delivery through the first aperture immediately started. While the radiation was being delivered through the first aperture, the DAO algorithm continued to optimize the remaining 41 apertures. After another time period T the second aperture was fixed and radiation delivery through the second aperture immediately commenced. While the radiation was being delivered through the second aperture, the DAO algorithm continued to optimize the remaining 40 apertures for another time period T, after which the third aperture was fixed. This process was continued to the last aperture. For a treatment plan with a total of N apertures, the number of delivered apertures and the number of remaining apertures is plotted as a function of time in Figure 5.1.

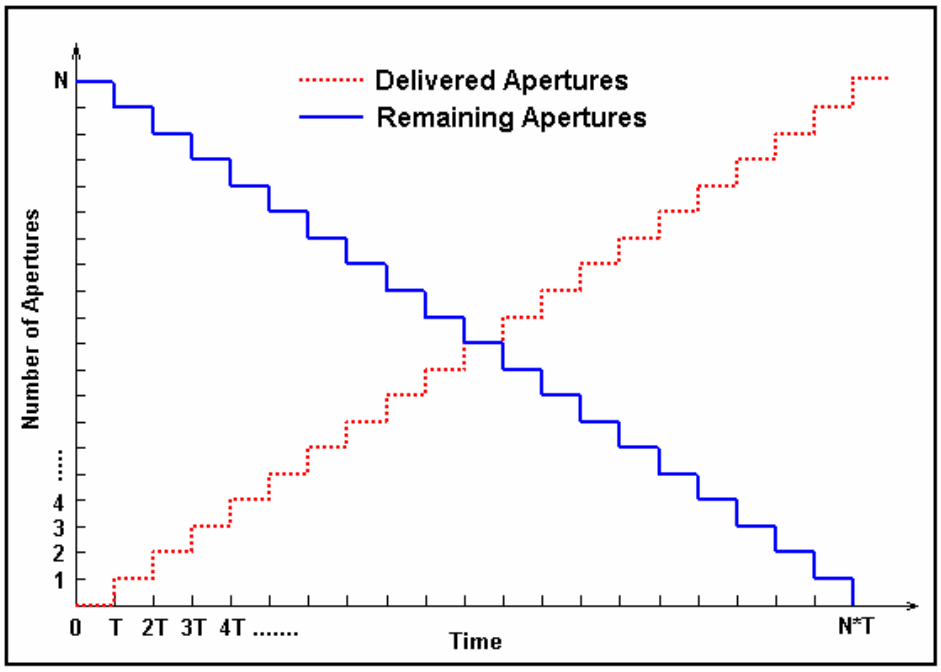


*Figure 5.1: Adaptive Radiation Delivery (ARD): Radiation delivery starts as soon as the first aperture is "fixed" (after time T). While the first aperture is being delivered the rest of the apertures are being optimized until the second aperture is fixed. This process is continued to the last aperture, N.*

During ARD, the apertures are delivered immediately one after the other, since the derivation of the next aperture is performed during the delivery time of the previous aperture. For a treatment plan with a total of 42 apertures, an average time required to deliver one aperture is about 5-10 seconds. To maximize the efficiency of ARD the next aperture should be derived before the delivery of the previous aperture is complete. This way the apertures can be delivered immediately one after the other, without any 'dead time' between the deliveries of the apertures. Therefore, to maximize the efficiency of the ARD process, the time interval T should be as short as possible. In ARD, the effective time allotted for plan adaptation is only time interval T, since the rest of the plan adaptation overlaps with the radiation delivery and does not prolong the treatment time, as shown in Figure 5.2.

To investigate the effect of using different values of T, the ARD optimization was repeated with a time T ranging from 2 seconds to 20 seconds (in increments of 2 seconds) for each deformed anatomy. The goal was to determine the minimum value of T for which the ARD optimization was able to arrive at the clinically acceptable plan.

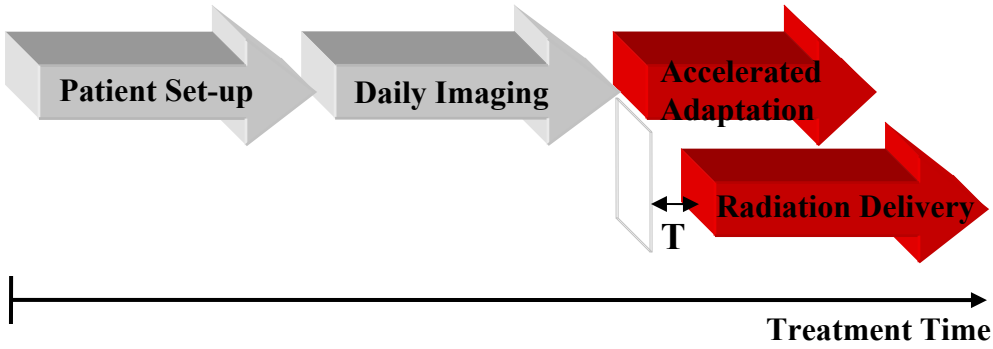


*Figure 5.2: In ARD, the effective time allotted for plan adaptation is only time interval T. The rest of the plan adaptation overlaps with the radiation delivery and does not prolong the treatment time.*

## 5.3 Results and Discussion

The ARD optimization was repeated for each deformed anatomy, with a parameter T ranging from 2 seconds to 20 seconds. Figure 5.3 displays objective function versus time for the four anatomical deformations (for clarity only selected values of T are shown). For the first deformed anatomy, the ARD optimization was able to arrive at an acceptable plan for all ten values of T (only T=2 s, 4 s and 6 s are shown). For the second deformed anatomy, the ARD optimization was able to arrive at an acceptable plan for all values of T, except T=2 s. For the third deformed anatomy, the ARD optimization was able to arrive at an acceptable plan for all values of T, except T=2 s and 4 s. Finally, for the last anatomical deformation the minimum value of T needed to reach the acceptable plan was 14 seconds. For all values of T larger than 14 seconds the ARD optimization was also able to reach the acceptable plan. The relatively large value for T needed for the last anatomical deformation is again most likely due to the fact that the anatomy deformation is so great that the original plan is substantially displaced from the global minimum and therefore substantially longer time is required for successful adaptation.

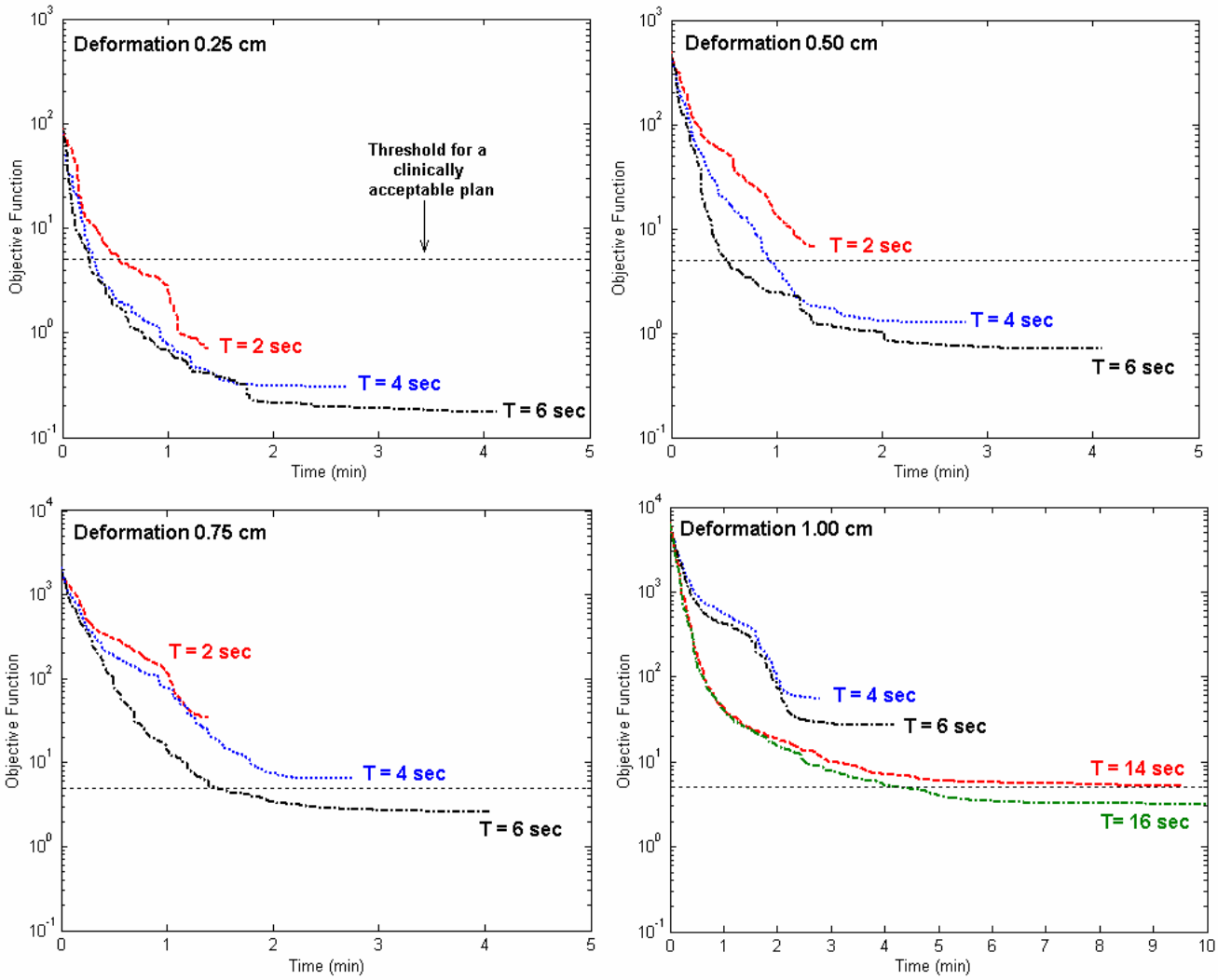


*Figure 5.3: To investigate the effect of using different values of T, time between the 'fixing' of the apertures, the ARD optimization was performed with T ranging from 2 seconds to 20 seconds (in increments of 2 seconds). The objective function versus time graphs are shown for the four anatomical deformations (for clarity only selected values of T are shown).*

An alternative way of displaying the ARD optimization results is to plot objective function value versus number of fixed apertures (Figure 5.4). Figure 5.4 shows that a clinically acceptable plan might be reached before all of the apertures are fixed. For example, for the 0.25 cm deformation a clinically acceptable plan was reached after only 16 apertures were fixed (T=2 seconds). Therefore, the plan adaptation could be stopped after the first 16 apertures were fixed, and the remaining apertures could be delivered without any further re-optimization. For the anatomical deformations 0.50 cm (T=4 sec) and 0.75 cm (T=6 sec), the clinically acceptable plan

was reached after about 15 apertures were fixed. For the last deformation, the clinically acceptable plan was reached only after all 42 apertures were fixed for T=14 seconds (or about 15 apertures for T=16 seconds).

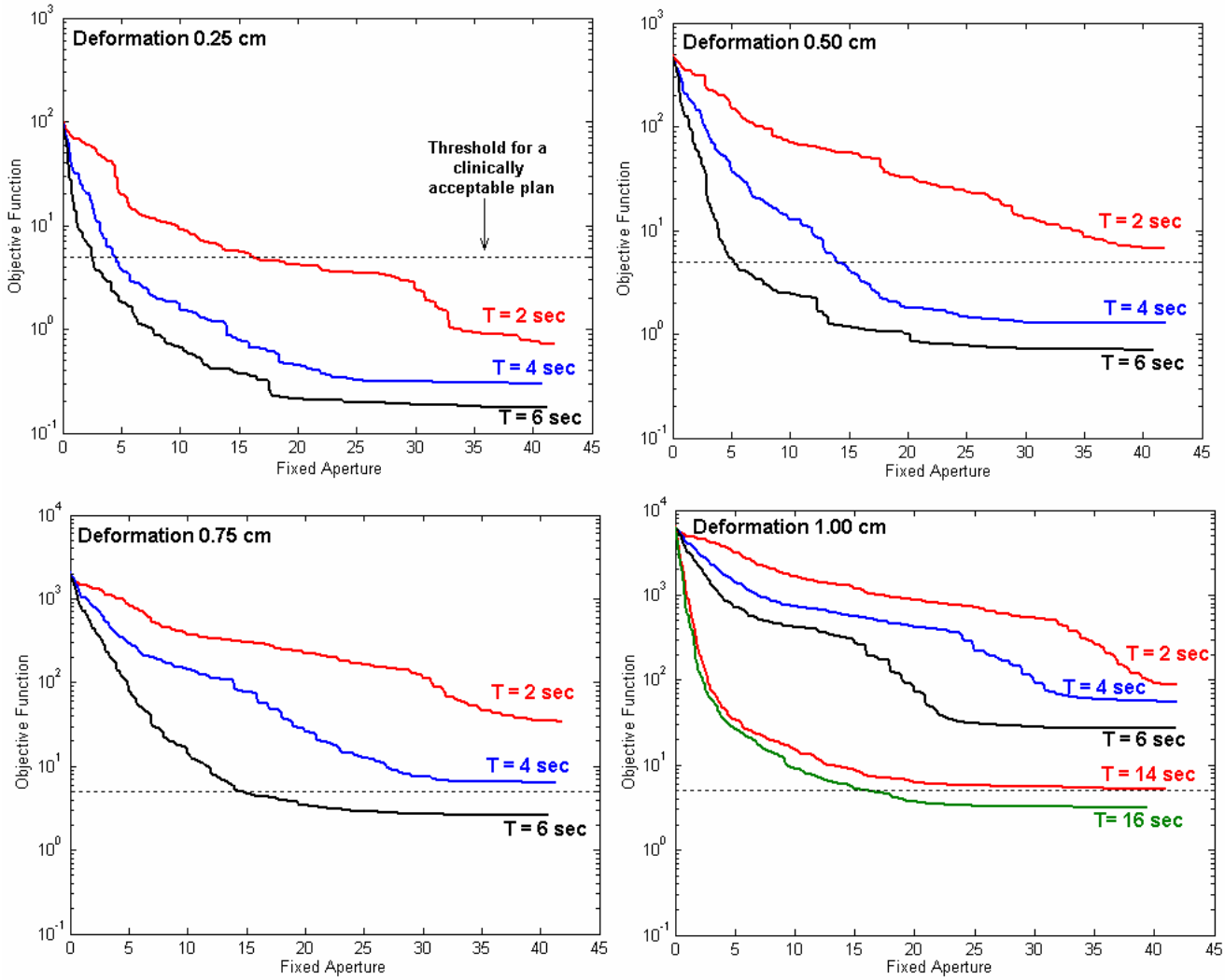


*Figure 5.4: The objective function versus number of fixed apertures for the ARD optimization (for clarity only selected values of T are shown).*

As mentioned earlier, to maximize the efficiency of ARD the next aperture should be derived before the delivery of the previous aperture is complete. This way the apertures can be delivered immediately one after the other, without any 'dead time' between the deliveries of the apertures. In this study, time interval T was set prior to the start of the adaptation and the same T was used throughout the adaptation process. However, to further maximize the efficiency of

ARD, a variable time interval T could be used. Time interval T could be varied for each aperture, based on the time required to deliver the previous aperture. T could simply be derived from the weight of the previous aperture (number of MUs that have to be delivered by that aperture) and the radiation delivery rate. This way the optimal value of T is used for each aperture and the efficiency of the process is maximized. However, there should always be a lower limit on T, since making T too small could result in failure to arrive at the acceptable plan. The lower limit on T should be chosen based on the degree of anatomical deformation, which can be determined by registering the planning image set (e.g. planning CT) with the daily image set (e.g. CBCT), obtained prior to the treatment. The patient anatomical deformation would be 'categorized' prior to the start of the treatment and the lower limit on T would be chosen. In our case, the appropriate lower limits for T would be 2 seconds, 4 seconds and 6 seconds for anatomical deformations of 0.25 cm, 0.5 cm and 0.75 cm, respectively. The anatomical deformation of 1.0 cm requires substantially longer minimum T and demonstrates the limit of this approach.

The concept of ARD is currently a topic of research rather than a technique that is ready for clinical implementation. However, one realistic clinical implementation of the ARD concept would be a combination of ARD with an additional off-line adaptation. As described in Section 2.5.3, in off-line ART the treatment plan adaptation is performed after the treatment fraction and the adapted plan is used for the subsequent fraction. Any dosimetric errors caused by the anatomical deformation in fraction n are attempted to be corrected in the subsequent fraction n+1. In a combined ARD and off-line ART approach, ARD would be performed, as described in this study, as a first step of a two-step adaptation process. If ARD alone is not able to completely adapt the plan so that it becomes clinically acceptable, another off-line step would be added, in which the next treatment fraction would compensate for the remaining error. For the cases in which ARD is able to successfully adapt the original plan, the second step would not be performed. In both cases, performing ARD as a first step would be beneficial since it would make the off-line correction easier, with minimal extension of treatment time.

## 5.4 Conclusion

A new approach to on-line ART was proposed and investigated, in which accelerated plan adaptation and radiation delivery are integrated together and performed concurrently. This approach is referred to as Adaptive Radiation Delivery (ARD). A fundamental advantage of ARD is the fact that radiation delivery can start almost immediately after daily imaging. Most of the plan re-optimization is performed during the radiation delivery, so the time spent adapting the original plan does not significantly increase the overall treatment time. As a consequence, the effective time allotted for plan adaptation is drastically reduced. For the 0.25 cm, 0.5 cm and 0.75 cm anatomical deformations, the treatment time was increased by only 2 sec, 4 sec and 6 sec, respectively, as compared to no plan adaptation. For the anatomical deformation of 1.0 cm the time increase was substantially larger. The anatomical deformation of 1.0 cm represents an extreme case, which is rarely observed for the prostate, and again demonstrates the limit of this approach.

# Chapter 6
# Integrating Daily Imaging, Plan Adaptation and Radiation Delivery

## 6.1 Introduction

In the previous chapter, accelerated plan adaptation and radiation delivery were integrated together and performed concurrently. An advantage of combining plan adaptation and radiation delivery together is that radiation delivery can start almost immediately after daily imaging, based on the partially re-optimized plan. Most of the plan re-optimization is performed during the radiation delivery, so the time spent adapting the original plan does not significantly increase the overall treatment time. In this chapter, the concept of combining plan adaptation and radiation delivery is extended by integrating it with daily image acquisition (Figure 2.20d). The advantages of this approach are:

1) Further reduction of treatment time compared to conventional on-line ART. Shorter treatment time increases patient throughput and decreases susceptibility to intra-fractional patient motion.

2) By continuously imaging the patient during the treatment, intra-fractional patient motion can be detected and accounted for.

Therefore, this approach has the ability to overcome the shortcomings and drawbacks of the current on-line ART procedure, discussed in Section 2.5.4. This study is the first investigation into the feasibility of performing on-line ART based on intra-fractionally acquired images. These ideas and concepts lead toward the ultimate goal of near real-time adaptive radiation therapy. This work has been submitted to the journal of Physics in Medicine and Biology[93].

## 6.2 Methods and Materials

The work in this chapter builds on the work from the previous two chapters. However, there is one important difference in this chapter with respect to modeling the anatomical deformations. In Chapters 4 and 5, the original anatomy and deformed anatomies were first created in the TPS and then were exported from the TPS into MATLAB using DICOM RT. In MATLAB, the DAO algorithm was used to create the original treatment plan based on the imported contours for the original anatomy. The original treatment plan was adapted based on the imported contours for the deformed anatomies. The contours of the structures for the deformed anatomies were assumed to be available prior to the start of plan adaptation. In other words, it was assumed that image acquisition, 3D image reconstruction and image segmentation (re-contouring), was performed prior to the start of the plan adaptation process.

In this chapter, on the other hand, image acquisition, 3D image reconstruction and image segmentation are integrated together with plan adaptation and radiation delivery. Therefore, the deformed anatomy contours are not available prior to the start of plan adaptation but are acquired during the treatment. To integrate image acquisition, image reconstruction and image segmentation together with plan adaptation and radiation delivery, 3D virtual patients with deformed anatomies were created directly in MATLAB.

### 6.2.1 3D Virtual Patients

Four virtual patients were created in MATLAB. The first virtual patient was created based on the original anatomy (Section 4.2.2). The other three virtual patients were created based on the 0.25 cm, 0.5 cm and 0.75 cm deformed anatomies (Section 4.2.2). The 0.25 cm, 0.50 cm and 0.75cm deformations represented small, medium and large clinical deformations, respectively. The extreme deformation of 1.00 cm was omitted from this part of the study. To create a virtual patient the prostate (CTV), rectum, bladder and phantom (body) contours were exported from the TPS into MATLAB using DICOM RT. In MATLAB, a 3D virtual patient was created by "filling" each contour with the material of an appropriate electron density. The electron densities relative to water assigned to prostate, rectum, bladder and body were 1.05, 0.97, 1.08 and 1.00, respectively[94]. Four virtual patient anatomies are shown in Figure 6.1.

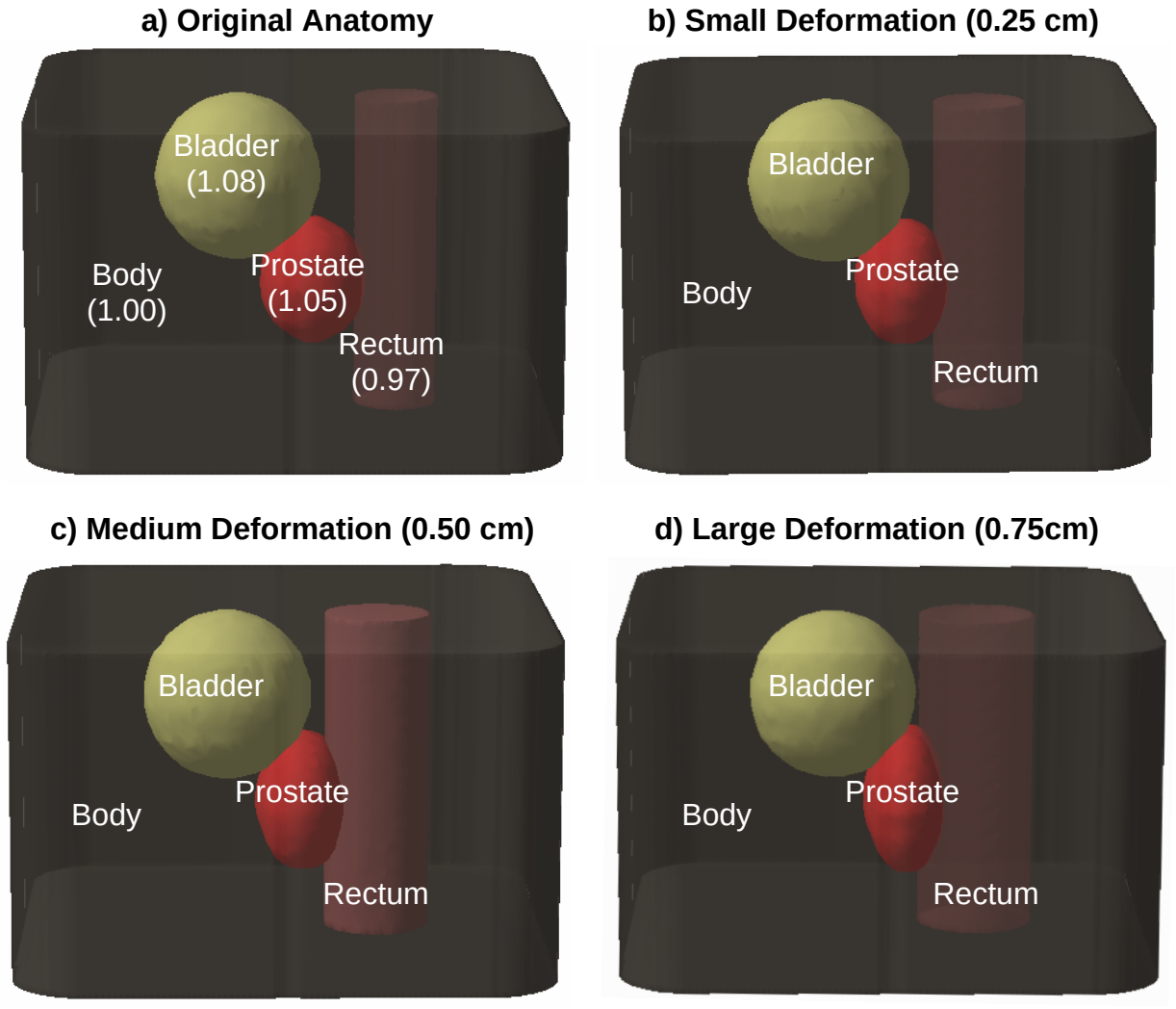


*Figure 6.1: Four 3D virtual patients were created in MATLAB based on: a) Original anatomy (parenthetical values specify electron density relative to water). b-d) 0.25 cm, 0.50 cm and 0.75 cm deformed anatomies.*

As in Chapter 4 and Chapter 5, our DAO algorithm (Section 4.2.1) was used to create the original treatment plan for the original anatomy. The dose-volume constraints for all the structures were the same as the ones specified in Section 4.2.3.

## 6.2.2 Daily Image Acquisition

To integrate daily image acquisition with treatment delivery an integrated Linac/CBCT system was simulated in MATLAB. As mentioned in Section 2.5.2, in an integrated Linac/CBCT system, the CBCT is mounted directly on the linac gantry and rotates together with the linac around a common axis of rotation. The common axis of rotation is parallel with the long axis of the treatment couch and goes through the isocenter. The CBCT consists of an x-ray source (kV x-ray tube) and a 2D flat-panel x-ray detector (Figure 6.2). CBCT operates on the same principle as conventional CT, with the exception that an entire volumetric image is acquired through a single rotation of the x-ray source and the x-ray detector. The ability of CBCT to acquire an entire volumetric image in a single rotation allows for natural integration of daily image acquisition with treatment delivery. As the primary MV beam rotates between treatment positions (beams) and delivers radiation, the CBCT can be used to continuously acquire 2D x-ray projections. Thus, daily image acquisition is integrated with treatment delivery in a single gantry rotation around the patient.

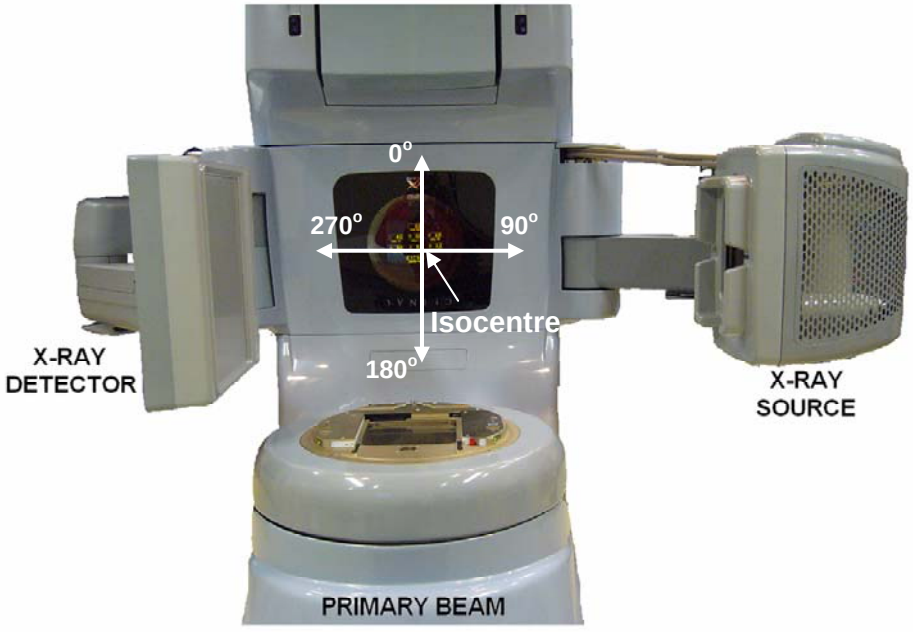


*Figure 6.2: In an integrated Linac/CBCT system, the CBCT is mounted directly on the linac gantry and rotates together with the linac around a common axis of rotation. The CBCT consists of an x-ray source and a 2D flat-panel x-ray detector.*

## 1. Simulated Linac/CBCT System

In this study, two different orientations of the CBCT system with respect to the primary MV beam were simulated: 1) CBCT mounted orthogonally to the MV beam (Figure 6.2) and 2) CBCT mounted at an angle of 180 degrees with respect to the MV beam (ie. in-line setup). Both of these orientations are commercially available: orthogonal set-up on Varian (Varian Medical Systems, Palo Alto, CA) and Elekta (Elekta Oncology Systems) linear accelerators and in-line set-up on Siemens (Siemens OCS, Concord, USA) linear accelerators.

Figure 6.3 shows the angular range through which the CBCT system rotated as the MV beam rotated counterclockwise from the initial position (gantry angle=180 deg) to the first treatment position (gantry angle=110 deg). This process was continued as the primary MV beam rotated from the first treatment position to the second treatment position, from the second treatment position to the third treatment position, etc. (not shown in Figure 6.3). Therefore, the total imaging range through which the CBCT system rotated progressively increased as the MV beam rotated between seven treatment positions (Table 6.1).

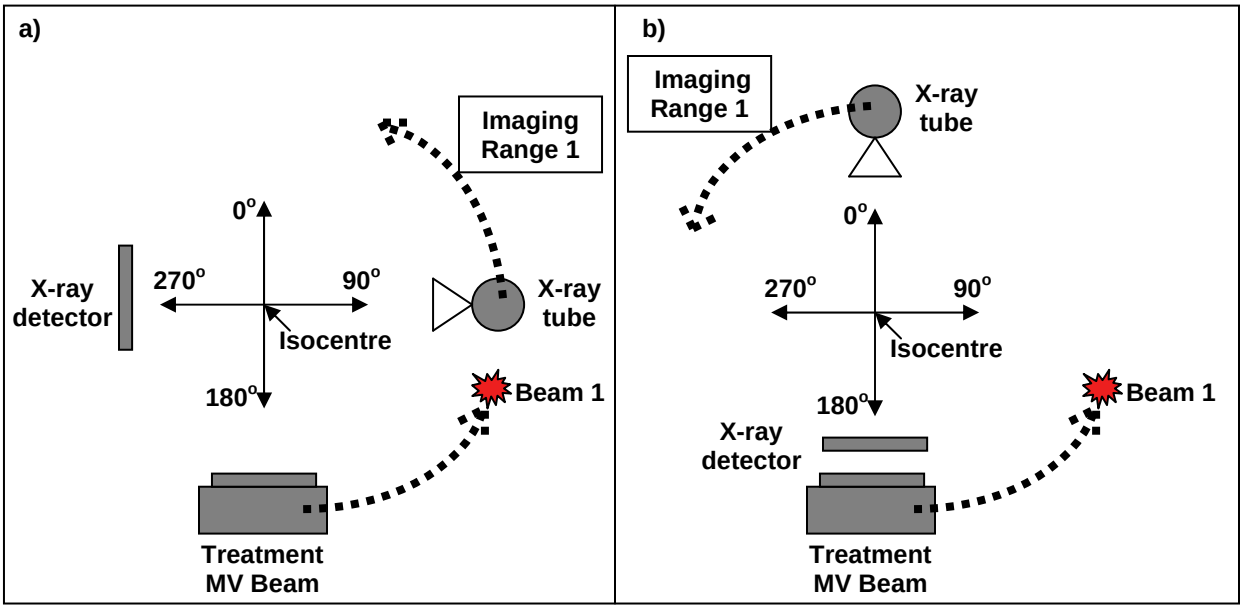


*Figure 6.3: The imaging range through which the CBCT system rotated as the primary MV beam rotated counterclockwise from the initial position (gantry angle=180 deg) to the first treatment position (gantry angle=110 deg) for: a) CBCT mounted orthogonally to the MV beam and b) CBCT mounted in-line with the MV beam.*

| Treatment positions (deg) | MV beam rotation (deg) | TOTAL imaging range (deg) | |
|---|---|---|---|
| | | Orthogonal set-up | In-line set-up |
| 110 | 180 to 110 | 90 to 20 | 0 to -70 |
| 80 | 180 to 80 | 90 to -10 | 0 to -100 |
| 40 | 180 to 40 | 90 to -50 | 0 to -140 |
| 355 (-5) | 180 to -5 | 90 to -95 | 0 to -185 |
| 310 (-50) | 180 to -50 | 90 to -140 | 0 to -230 |
| 280 (-80) | 180 to -80 | 90 to -170 | 0 to -260 |
| 250 (-110) | 180 to -110 | 90 to -200 | 0 to -290 |

*Table 6.1: The total imaging range through which the CBCT system rotated progressively increased as the primary MV beam rotated between seven treatment positions (beams).*

## 2. Simulated 2D X-ray Projection: Digitally Reconstructed Radiographs (DRRs)

For each virtual patient with deformed anatomy, DRRs were generated to simulate the 2D x-ray projections acquired for the imaging ranges listed in Table 6.1. To generate DRRs a virtual CBCT system was created in MATLAB. The DRRs were computed by forward projecting[95] a divergent radiation beam through the virtual patient from the x-ray source to the x-ray detector (Figure 6.4). The axial slice thickness of the virtual patient was 3 mm, corresponding to the slice thickness of the contours imported from the TPS. The beam intensity was exponentially attenuated based on the electron density of the voxels intersected along the beam's path. For low Z materials (e.g. human body) the electron density is directly proportional to the linear attenuation coefficient since the photons are predominantly attenuated by Compton interactions. This follows from the fact that the mass attenuation coefficient for Compton scattering ($\sigma_{incoherent}/\rho$) is independent of the atomic number of the medium (Section 1.2.3). The radiation beam was subdivided into divergent beamlets and the intensity attenuation along each beamlet was calculated using the following equation:

$$I = I_o \exp(-\sum_{i=1}^{N} \mu_i x_i) \tag{6.1}$$

Where:

$I$ = Attenuated beamlet intensity

$I_o$ = Initial beamlet intensity

$N$ = Number of voxels intersected along beamlet's path

$\mu_i$ = Linear attenuation coefficient of the i$^{th}$ voxel

$x_i$ = Path length through the i$^{th}$ voxel

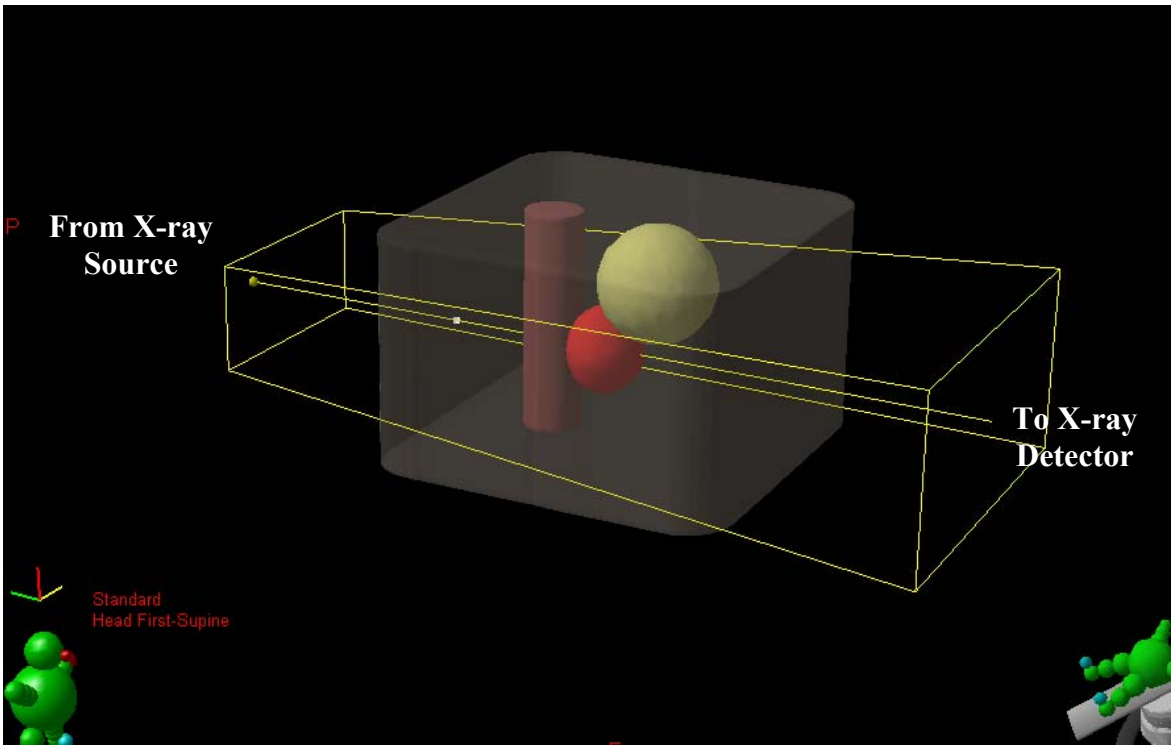


*Figure 6.4: A virtual CBCT system was created in MATLAB. The DRRs were computed by forward projecting a divergent x-ray beam through the virtual patient from the x-ray source to the x-ray detector.*

The DRRs were calculated with a beamlet resolution of 0.25 mm by 0.25 mm defined at the isocenter and angular spacing between DRRs of 0.50 deg (Figure 6.5). The source-to-axis distance (SAD) and source-to-detector distance (SDD) were 100 cm and 150 cm, respectively. These values were closely matched to the values used by the Varian (Varian Medical Systems, Palo Alto, CA) CBCT system: average angular spacing between x-ray projections of 0.57 deg,

detector resolution of 0.2587 mm by 0.2587 mm defined at the isocenter, SAD of 100 cm and SDD of 150 cm. The DRRs were normalized such that the pixel values ranged from 0 to 1.

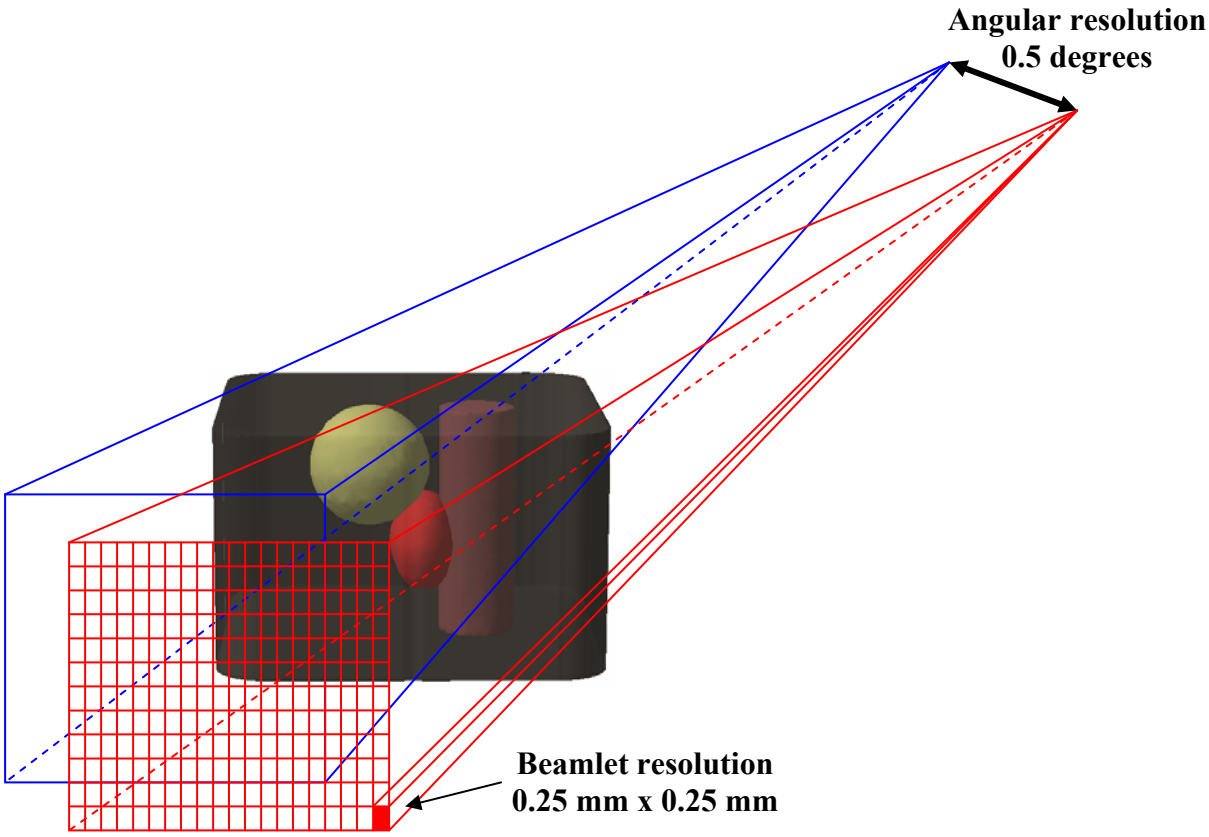


*Figure 6.5: The DRRs were computed with a beamlet resolution of 0.25 mm by 0.25 mm at the isocenter and angular spacing between DRRs of 0.5 deg.*

The DRRs calculated by forward projecting through the virtual patient do not account for various factors that can affect the CBCT imaging performance such as photon statistics (quantum noise), electronics noise, detector blurring, detector pixel size, etc. To account for image degradation due to these factors, detector blurring and system noise were experimentally measured and added to the DRRs prior to image reconstruction.

### 3. Detector Blurring

Detector blurring was modeled by a two dimensional Gaussian point spread function (PSF)

$$F(x,y) = \frac{1}{2\pi\sigma_b^{\,2}} \exp\left(\frac{-\left(x^2+y^2\right)}{2\sigma_b^{\,2}}\right) \tag{6.2}$$

where $\sigma_b$ was experimentally determined from line spread function (LSF) measurement[96]. To measure LSF of our CBCT system (Varian Medical Systems, Palo Alto, CA), a steel wire under tension suspended in a 20.5 cm diameter water bath was used (Figure 6.6). The diameter of the wire was 0.157 mm.

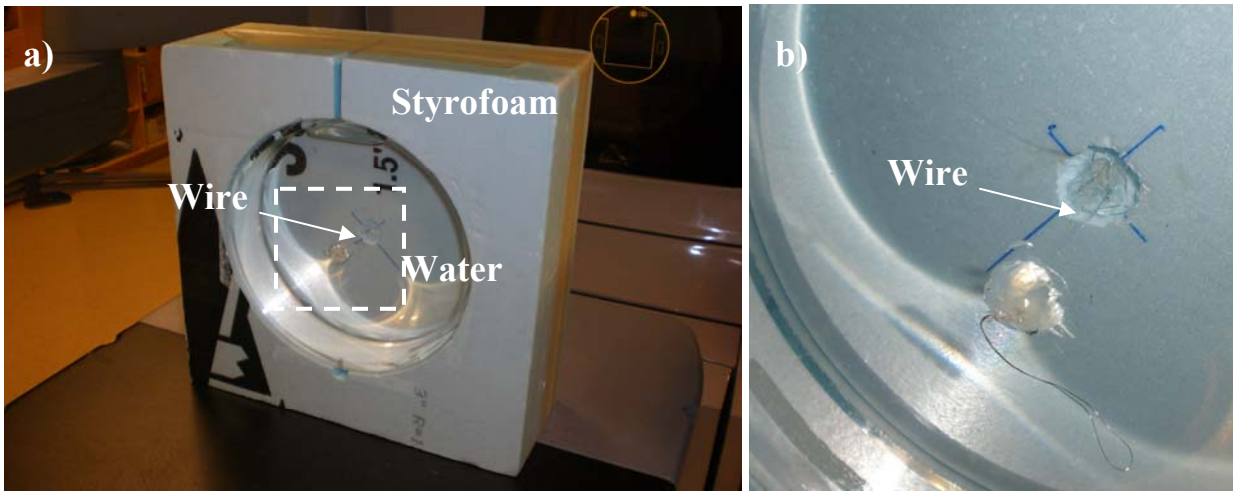


*Figure 6.6: a) A steel wire under tension suspended in a 20.5 cm diameter water bath was used to measure the line spread function (LSF) of our CBCT system. The diameter of the wire was 0.157 mm. b) The magnified view of the selected region in a).*

The phantom was positioned on the treatment couch such that the steel wire was close to the isocenter. Ten x-ray projection images were acquired using the CBCT system with the x-ray tube at 270 degrees and x-ray detector at 90 degrees (Figure 6.7). This particular CBCT orientation was arbitrarily chosen. The x-ray projection images were averaged to reduce the image noise and the LSF was determined from the averaged image. The line profile was taken perpendicular to the steel wire projection through the center of the image, as indicated by the dashed line on the detector in Figure 6.7.

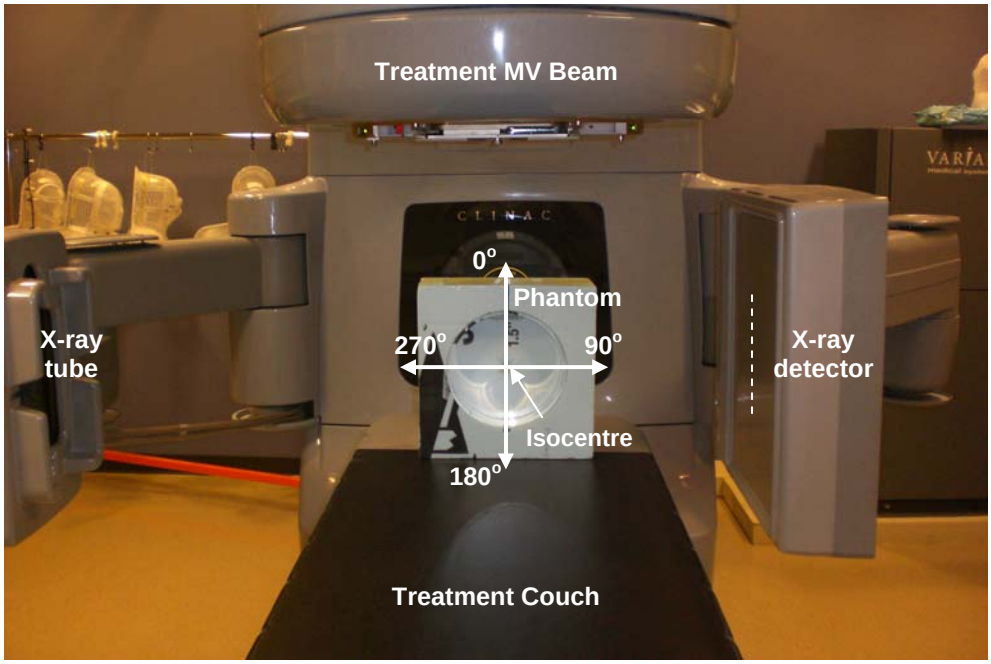


*Figure 6.7: To measure the line spread function (LSF) of the detector the phantom was positioned on the treatment couch with the steel wire at the isocenter. Ten x-ray projection images were acquired using the CBCT system. The line profile was taken perpendicular to the steel wire projection through the center of the averaged projection image, as indicated by the dashed line on the detector.*

Figure 6.8 displays the average pixel values along the line profile together with the associated error bars. The error bars correspond to one standard deviation. One dimensional Gaussian LSF was fitted to the data and the best fit is shown in Figure 6.8. The best fit was obtained with $\sigma_b = 0.75$ pixels or 0.3 mm, since the detector resolution was 0.388 mm/pixel. The value of 0.3 mm was in agreement with values reported by other groups[97]. This value was used to model the detector blurring for the rest of this chapter.

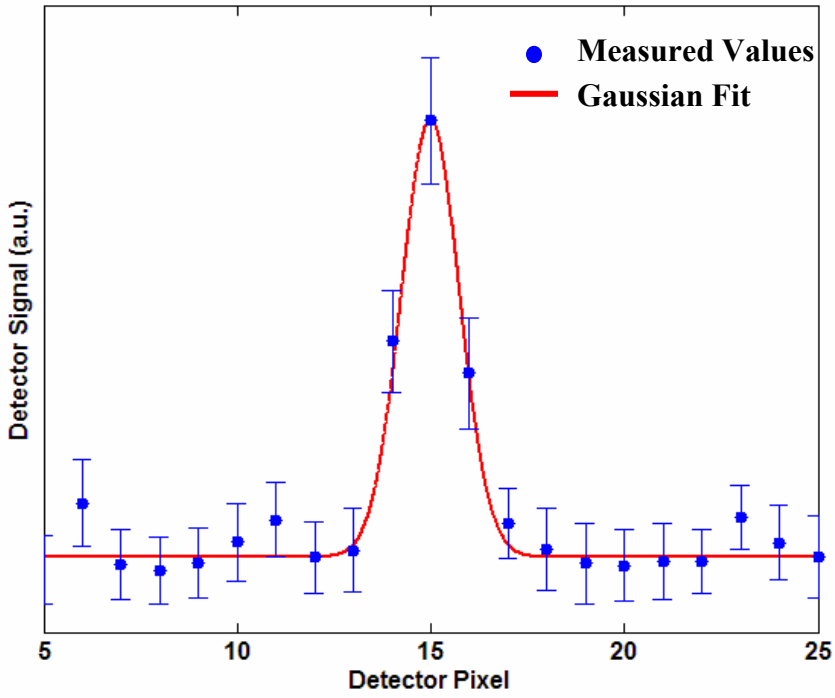


*Figure 6.8: The blue markers represent the average pixel values along the line profile across the detector, as indicated by a dashed line in Figure 6.7. The error bars correspond to one standard deviation. The red line represents the best fit 1D Gaussian line spread function (LSF). The best fit was obtained with $\sigma_b$ = 0.75 pixels or 0.3 mm, since the detector resolution was 0.388 mm/pixel.*

## 4. System Noise

System noise was sampled from the Gaussian probability density function

$$P(s) = \frac{1}{\sqrt{2\pi}\sigma_{noise}} \exp\left(\frac{-s^2}{2\sigma_{noise}^2}\right) \tag{6.3}$$

where noise variance $\sigma^2_{noise} = \sigma_n I_{norm}(x,y)$ was made proportional to the normalized x-ray fluence, $I_{norm}(x,y)$. The fluence was normalized such that the values ranged from 0 to 1. The value of $\sigma_n$ was experimentally determined from a noise power spectrum (NPS) analysis (next Section).

The system noise and detector blurring were incorporated into each simulated DRR to produce a more realistic DRR' to be used in image reconstruction, using the following equation[97]:

$$DRR'(x,y) = \left[DRR(x,y) + n_s(x,y)\right] \otimes F(x,y) \tag{6.4}$$

where $n_s(x,y)$ is sampled from the system noise probability density function (Equation 6.3) and *F(x,y)* is a two dimensional Gaussian PSF (Equation 6.2).

## 5. Noise Power Spectrum (NPS) Analysis

For the experimental measurement of the NPS, a water filled cylindrical phantom was used (Figure 6.9). The length and the diameter of the phantom were 28.5 cm and 21.5cm, respectively. The phantom was positioned on the treatment couch such that the central axis of the phantom coincided with the linac/CBCT axis of rotation. An x-ray projection image of a phantom was acquired using the CBCT system with the x-ray tube at 270 degrees and x-ray detector at 90 degrees (Figure 6.10). Again, this particular CBCT orientation was arbitrarily chosen.

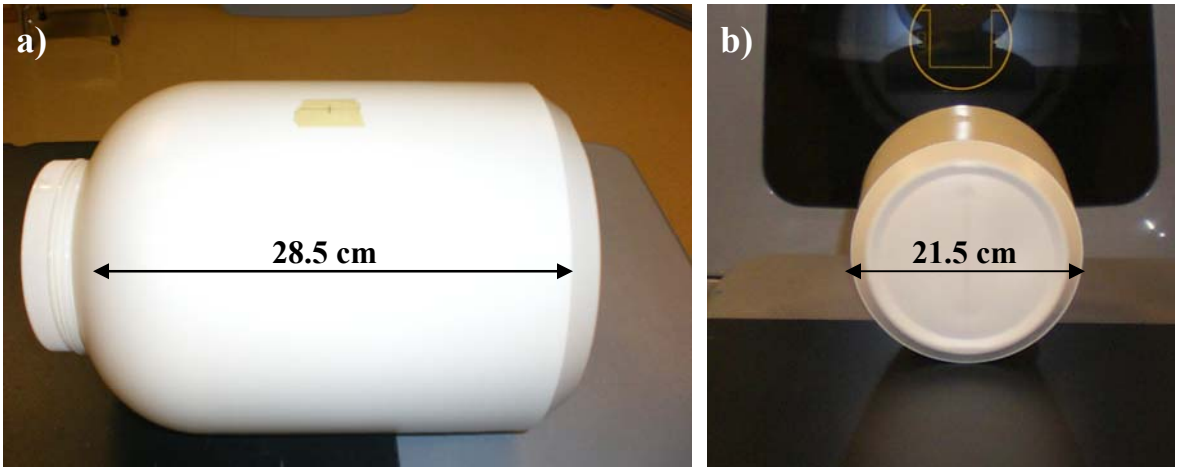


*Figure 6.9: Water filled cylindrical phantom was used to measure the noise power spectrum (NPS) of our CBCT system.*

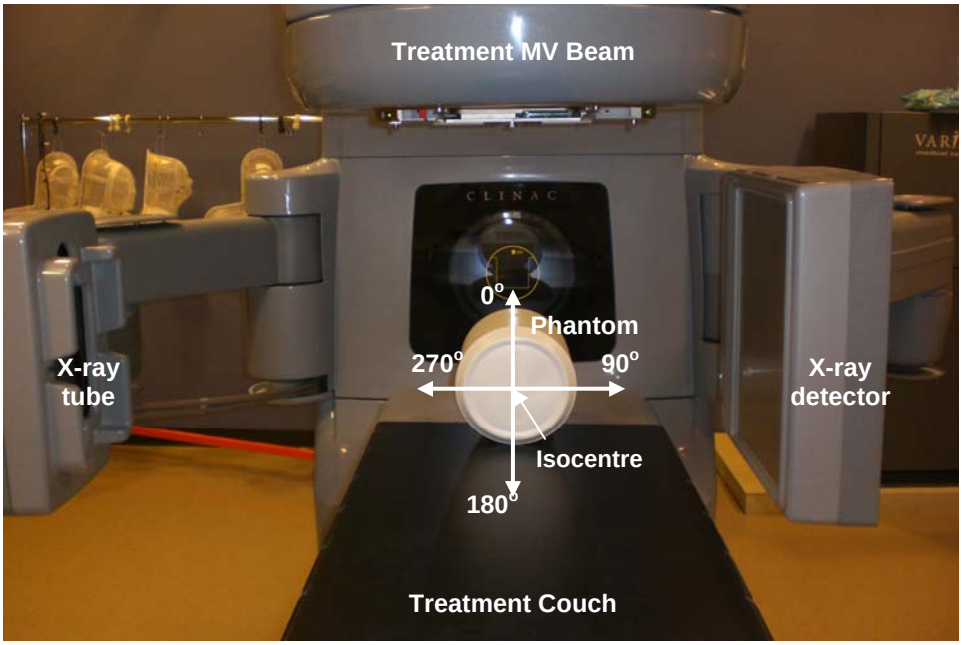


*Figure 6.10: To measure the noise power spectrum (NPS) of the CBCT system the phantom was positioned on the treatment couch such that the central axis of the phantom coincided with the linac/CBCT axis of rotation.*

The NPS was measured over 10 regions of interest (128 x 128 pixels in each region) from the central part of the acquired projection image. A 2D fast Fourier transform (FFT) was calculated for each 128 x 128 region and NPS was computed, as follows:

$$NPS(f_x, f_y) = \frac{\left|FFT\{I(x,y)\}\right|^2}{N_x N_y}\Delta x \Delta y \tag{6.5}$$

where $I(x,y)$ is the projection image data, $N_x$ and $N_y$ are the number of pixels in the $x$ and $y$ directions, $\Delta x$ and $\Delta y$ are the pixel sizes in the $x$ and $y$ directions, and $f_x$ and $f_y$ are the spatial frequencies in the $x$ and $y$ directions. Since the signal level may vary from one ROI to another ROI, NPS($f_x$,$f_y$) were normalized to the mean image signal squared $S^2$, computed for each ROI, to eliminate the direct effect of signal variation. The result, NPS($f_x$,$f_y$)/$S^2$, referred to as the normalized NPS, quantified the square of noise-to-signal ratio rather that the absolute noise values and offered a signal-independent measure of noise level in the image. The normalized NPSs were averaged over all the ROIs to reduce the fluctuations. Finally, the average from the

upper 4 ($y > 0$) and lower 4 ($y < 0$) horizontal lines parallel to the x-axis ($y = 0$) in the spatial frequency domain was taken to produce the one-dimensional normalized NPS [96, 97].

To determine the appropriate value for $\sigma_n$ in Equation 6.3, a mathematical model of an identical water filled cylindrical phantom was constructed in MATLAB. The DRR was calculated by forward projecting through this virtual cylindrical phantom from the x-ray source to the x-ray detector, as described earlier. The system noise and detector blurring were added to the DRR to produce DRR' using Equation 6.4. The amount of system noise added to the DRR was varied by selecting different values of $\sigma_n$ in Equation 6.3. For each selected value of $\sigma_n$, the normalized NPS was calculated for the resulting DRR' in exactly the same manner as the normalized NPS was calculated for the x-ray projection acquired with the CBCT system. Figure 6.11 shows experimentally measured normalized NPS and simulated normalized NPSs corresponding to different values of $\sigma_n$. By directly comparing the simulated normalized NPSs to the experimentally measured normalized NPS, the most appropriate value of $\sigma_n$ to be used in Equation 6.3 was determined to be $0.6 \times 10^{-5}$. This value was used for system noise simulation for the rest of this chapter.

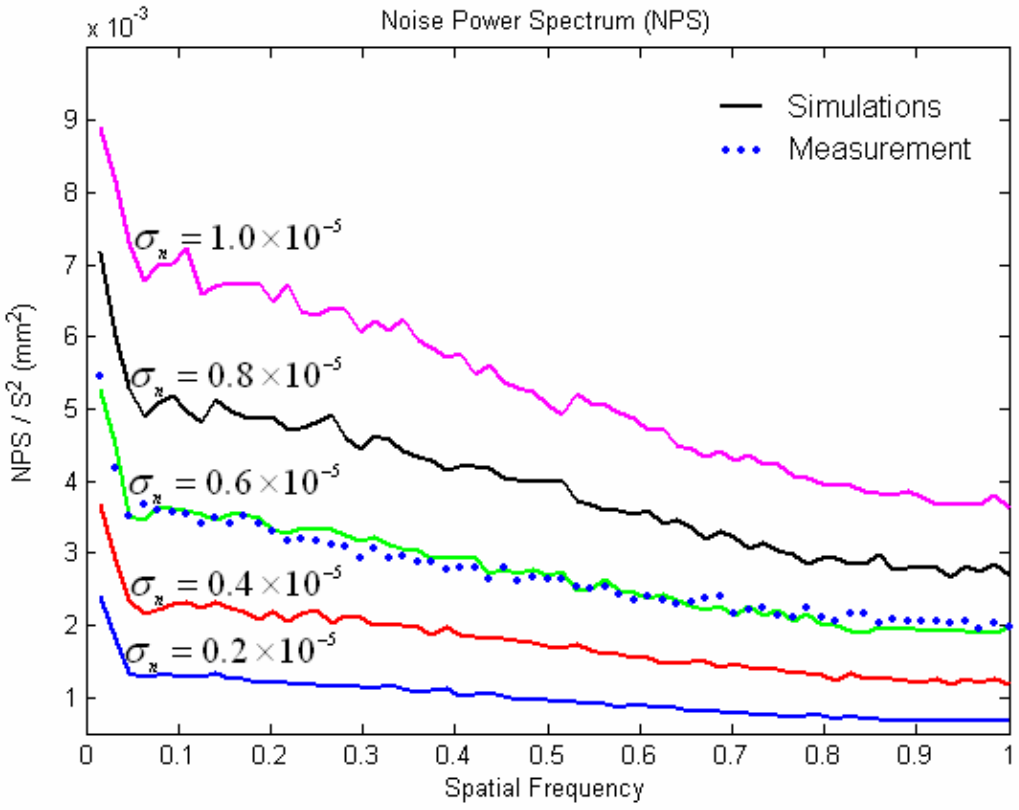


*Figure 6.11: Experimentally measured normalized NPS (blue markers) and simulated normalized NPSs corresponding to different values of $\sigma_n$ (solid lines) are shown together. The most appropriate value of $\sigma_n$ was determined to be $0.6\times10^{-5}$. Note that the spatial frequency on the x-axis has been rescaled so that Nyquist frequency is normalized to 1.*

## 6.2.3 Image Reconstruction

The modified Feldkamp filtered back-projection algorithm [98, 99] was used to reconstruct a 3D image from the x-ray projections. All of the image reconstruction was performed using the projections with the system noise and detector blurring incorporated via Equation 6.4. A process of 3D image reconstruction from the limited angle x-ray projection data (i.e. less than full 360 degrees) is known as Digital Tomosynthesis (DTS). A DTS image reconstruction algorithm was implemented in MATLAB. The following is an overview of our DTS algorithm implementation. A 3D representation of the attenuation coefficients $f(x,y,z)$, representing a 3D image of an object, was calculated using the following equation:

$$f(x,y,z)=\int_{\beta=\beta_{\min}}^{\beta=\beta_{\max}}\frac{D^2{}_{SAD}}{\left(D_{SAD}+x\sin\beta-y\cos\beta\right)^2}\times Q_\beta\left(\frac{D_{SAD}x\cos\beta+D_{SAD}y\sin\beta}{D_{SAD}+x\sin\beta-y\cos\beta},\frac{D_{SAD}z}{D_{SAD}+x\sin\beta-y\cos\beta}\right)d\beta \tag{6.6}$$

Where $D_{SAD}$ is the distance between the x-ray source and the isocenter, $\beta=[\beta_{\min},\beta_{\max}]$ is the limited angle x-ray projection range used in the reconstruction, and

$$Q_\beta(\rho,\xi)=\frac{D_{SAD}}{\sqrt{D^2{}_{SAD}+\rho^2+\xi^2}}R_\beta(\rho,\xi)*\frac{1}{2}h(\rho) \tag{6.7}$$

is the weighted filtered projection. In Equation 6.7, $\rho$ and $\xi$ are the detector's horizontal and vertical coordinates respectively, $h(\rho)$ denotes a one-dimensional ramp filter with a Hamming window and $R_\beta(\rho,\xi)$ is the projection of the attenuation values onto the detector:

$$R_\beta(\rho,\xi)=In\left[\frac{I_o(\rho,\xi)}{I(\rho,\xi)}\right] \tag{6.8}$$

In Equation 6.8, $I(\rho,\xi)$ is the projection data and $I_o(\rho,\xi)$ is the flood field, acquired with no object inside the field of view.

All image reconstruction was performed on the same axial set of planes on which the original anatomy contours were defined in the TPS. The spacing between the reconstructed planes was 3 mm and the in-slice resolution was set to 0.3418 mm by 0.3418 mm. The in-slice resolution was chosen to match the in-slice resolution of the axial slices on which the original anatomy contours were defined in the TPS. The advantage of reconstructing the DTS images on the same axial set of planes on which the original anatomy contours were defined is that the original contours can be directly used as a starting point in DTS image segmentation.

## 6.2.4 Image Segmentation

An edge detection algorithm was used to automatically segment the DTS image as the primary MV beam arrived at each treatment position (Table 6.1). The edge detection was performed based on the Interacting Multiple Model Probabilistic Data Association Filter (IMM/PDAF) technique[100]. The IMM/PDAF edge detection algorithm was originally implemented in MATLAB for Ultrasound image segmentation by Badiei et al.[101]. In this study, the original algorithm implementation was modified to be used for segmentation of 3D DTS images. The following is an overview of our IMM/PDAF edge detection algorithm for the segmentation of 3D DTS images. A seed point was placed at the center of the original contour for each DTS image slice. A number of equispaced rays, $N_r$, were radially projected from the seed point outwards. Based on the gradient of the image intensity along each ray, a number of candidate edge points, $N_c$, were selected for each ray. Only the pixels within the range [$L_{min}$ , $L_{max}$] around the original contour were considered (Figure 6.12). Using a modified Kalman filter recursively[100], one of the $N_c$ candidate edge points was selected as the new edge point for each ray, thus defining the new contour. The whole process was repeated for each DTS image slice and each structure.

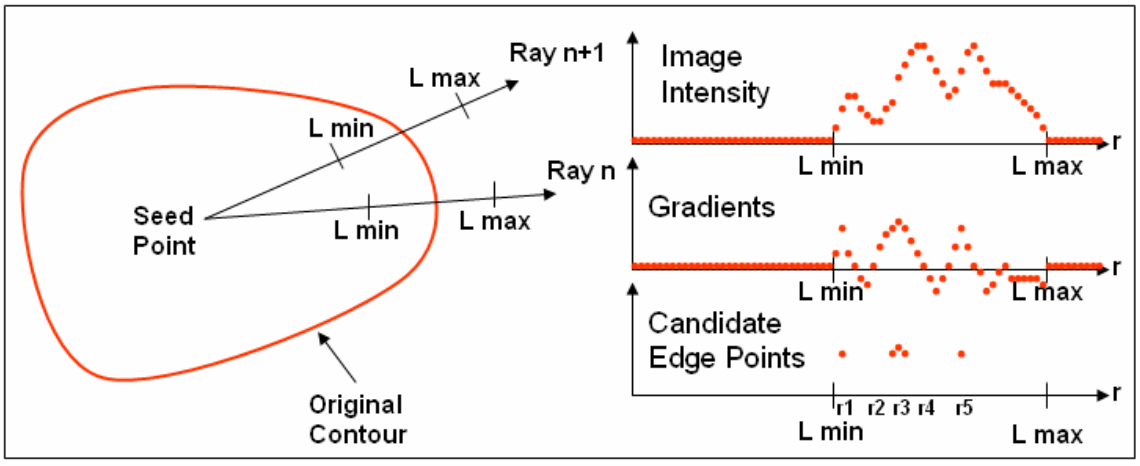


*Figure 6.12: A schematic diagram of the IMM/PDAF edge detection algorithm. Only two rays are shown for clarity.*

The IMM/PDAF edge detection algorithm was used to re-contour the prostate, rectum and bladder on the same axial set of planes on which the original anatomy contours were defined. The number of rays $N_r$ was set to 120 (i.e.3 degree angular increments) and the number of candidate edge points $N_c$ for each ray was set to 5. Only the pixels within 1.5 cm around the original contour were considered (i.e. $L_{min}$ = $L_{max}$ =1.5cm), setting the limit of 1.5 cm on the maximum deformation of the original contour. The total time required to re-contour the prostate, rectum and bladder on a 3D DTS image was about 5 seconds on a 2.8 GHz Pentium 4 PC.

Since the total angular range for which the x-ray projections were available for image reconstruction progressively increased as the primary MV beam rotated between treatment positions (Table 6.1), the reconstructed DTS image quality improved as the primary MV beam rotated between treatment positions. For the first few treatment positions the reconstructed DTS image quality was highly spatially non-isotropic: with high image quality in the direction perpendicular to the limited angle scan and lower image quality parallel to the scan direction[95]. When the image quality was such that the gradient of the image intensity along particular rays fell below some threshold value, no candidate edge points were defined for those rays. Instead, the difference between the original contour and the portions of the new contour, for which the edge was identifiable, was interpolated and used to "fill" in the missing gaps of the new contour (Figure 6.13).

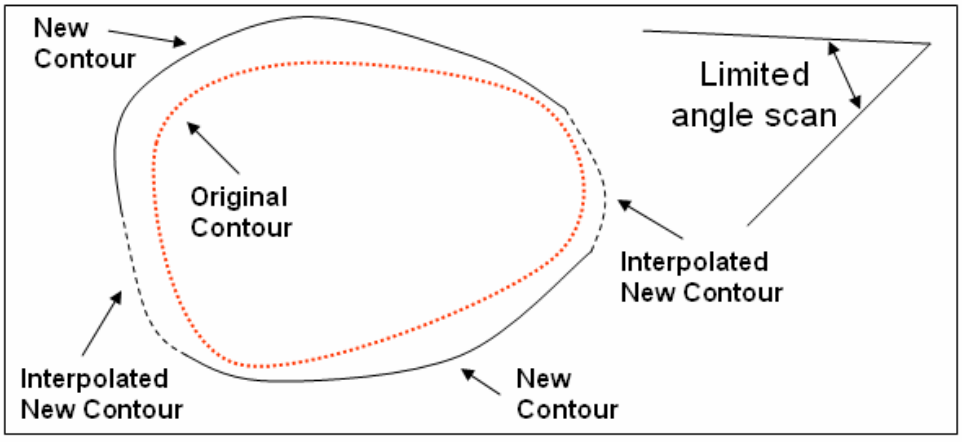


*Figure 6.13: When no candidate edge points could be identified in the direction parallel to the scan direction, the difference between the original contour and the portions of new contour, for which the edge was identifiable, was interpolated and used to "fill" in the missing gaps of the new contour.*

### 6.2.5 Plan Adaptation and Radiation Delivery

To illustrate how DTS imaging, plan adaptation and radiation delivery were integrated together consider the original DAO treatment plan with seven beams and with six apertures per beam, for a total of 42 apertures. As the primary MV beam arrived at the first treatment position (beam 1), the plan adaptation commenced based on the first DTS image contours. The leaf positions and aperture weights of all 42 apertures were re-optimized using the ARD technique (Chapter 5). Instead of continuing to re-optimize all 42 apertures until the plan was completely adapted, partial plan adaptation was performed for pre-defined periods of time T. After time T the first aperture of beam 1 was 'fixed' and radiation delivery through the first aperture immediately started. While the radiation was being delivered through the first aperture, the DAO algorithm continued to re-optimize the remaining 41 apertures. After another time period T the second aperture of beam 1 was fixed and radiation delivery through the second aperture immediately commenced. While the radiation was being delivered through the second aperture, the DAO algorithm continued to optimize the remaining 40 apertures for another time period T, after which the third aperture of beam 1 was fixed and delivered. This process was continued to the last aperture of beam 1. Once all the apertures of beam 1 were delivered, the primary MV beam moved to the second treatment position (beam 2), at which point plan adaptation continued based on the second DTS image contours. The plan adaptation of the remaining 36 apertures was again performed for a time period T. After time T the first aperture of beam 2 was 'fixed' and radiation delivery through this aperture immediately started. While the radiation was being delivered through this aperture, the DAO algorithm continued to re-optimize the remaining 35 apertures. This process was continued to the last aperture of beam 2. Once all the apertures of beam 2 were delivered, the primary MV beam moved to the third treatment position (beam 3), at which point plan adaptation continued in the same manner, based on the third DTS image contours. This process was continued to the last aperture of beam 7.

As mentioned in Chapter 5, to maximize the time efficiency of the ARD process the next aperture should be derived before the delivery of the previous aperture is complete. This way the apertures can be delivered immediately one after the other, without any 'dead time' between the deliveries of the apertures. Therefore, to maximize the time efficiency, time interval T should be equal to or less that the time required to deliver each aperture. For a treatment plan with a total of 42 apertures, an average time required to deliver one aperture is about 5-10 seconds. In this

study, T was set to 5 seconds. The effective time allotted for plan adaptation at each beam was T, the time required to re-optimize the first aperture of each beam. The adaptation of the rest of the apertures overlapped with the radiation delivery and did not prolong the treatment time. Therefore, for a treatment plan with seven beams the total effective time allotted for plan adaptation was 7T. With T set to 5 seconds, the total effective time allotted for plan adaptation was only 35 seconds. Adding this to the segmentation time of about 5 seconds (using 2.8 GHz Pentium 4 PC) for each 3D DTS image, the total treatment extension time was only 70 seconds. It should be mentioned that DTS image reconstruction did not add any extra treatment time, since filtered back-projection image reconstruction was performed concurrently with the x-ray projections acquisition, as the gantry rotated between treatment positions.

The hypothesis behind integrating patient imaging, plan adaptation and radiation delivery is that the dosimetric errors resulting from the initially incomplete knowledge of the anatomy can be compensated and corrected for during subsequent plan adaptation, as a more complete knowledge of the anatomy becomes available. To determine the feasibility of this approach, we used it to adapt the original treatment plan, optimized for the original anatomy, for each of the three anatomical deformations (i.e. small, medium and large), shown in Figure 6.1. The goal was to determine whether our integrated approach can adapt the original treatment plan so that it becomes "clinically acceptable" for each anatomical deformation. A treatment plan was considered to be clinically acceptable once the threshold objective function value of 10 was reached, as explained in Section 4.2.5. It should be mentioned that in Chapter 4 we decided to continue plan adaptation until a threshold objective function value of 5 was reached. As in Chapter 4, the overall findings of this chapter are insensitive to the exact value of the threshold used.

# 6.3 Results

## 6.3.1 Original Treatment Plan

The DVHs for the original treatment plan used for the original anatomy are shown in Figure 6.14. All of the treatment planning goals were met and the plan was considered clinically acceptable.

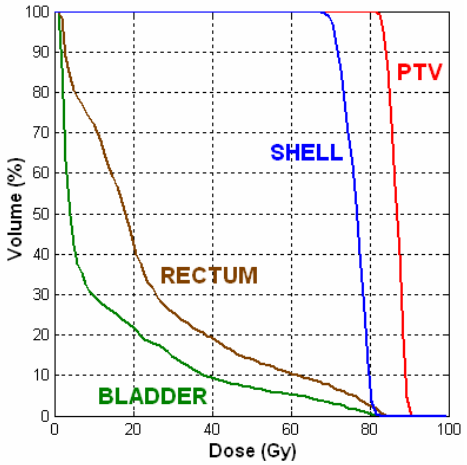


*Figure 6.14: DVHs for the original treatment plan used for the original anatomy. Shell (PTV + 5mm margin) eliminates hot spots outside the PTV.*

## 6.3.2 Non-Adapted Original Plan

Figure 6.15 shows DVHs for the non-adapted original plan used for the three deformed anatomies. As expected, the level of plan quality deterioration was related to the extent of the anatomy deformation. The overall dose to the rectum increased as the anatomy became more deformed. The inadequate PTV coverage made the original treatment plan clinically unacceptable for all three deformations.

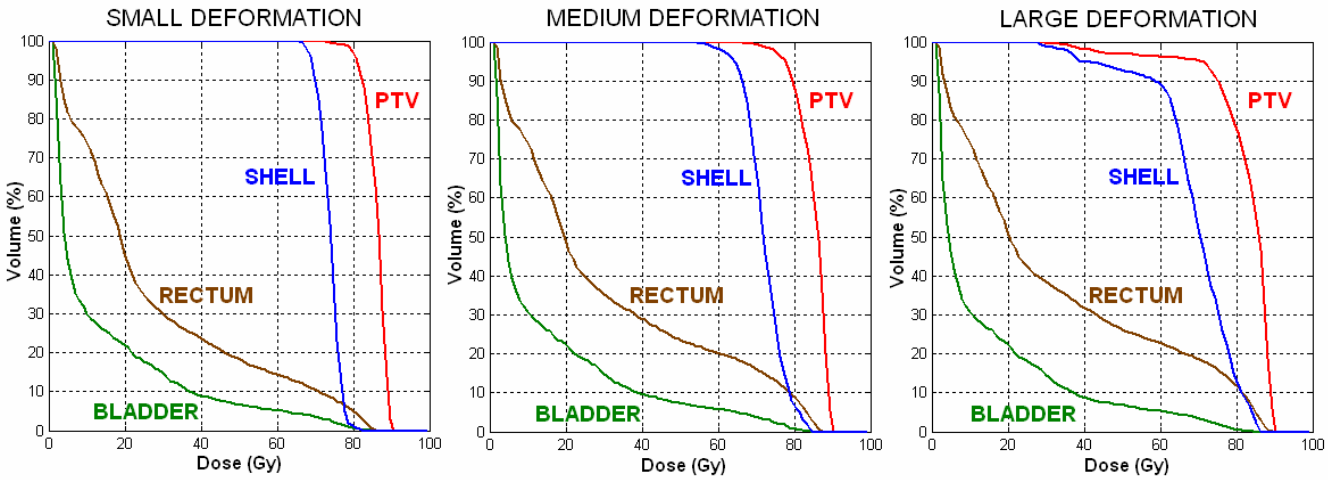


*Figure 6.15: DVHs for the non-adapted original treatment plan used for the deformed anatomies. The inadequate PTV coverage made the original treatment plan clinically unacceptable for all three deformations.*

### 6.3.3 Image Reconstruction and Segmentation

As mentioned earlier, the quality of the reconstructed DTS images improved as the gantry rotated between treatment positions, since more and more x-ray projections were available for DTS image reconstruction. Figure 6.16 shows DTS images obtained with the CBCT system mounted orthogonally to the primary beam (orthogonal set-up in Table 6.1). This figure shows the axial slice going through the isocentre of the reconstructed DTS image for the medium anatomical deformation at each treatment position. Figure 6.17 displays the same axial slice automatically segmented at each treatment position using the IMM/PDAF edge detection algorithm. The upper contour represents the re-segmented prostate and the lower contour represents the re-segmented rectum. The bladder is not present on the slice shown. The PTV contour (not shown) was obtained by adding a 5 mm margin to the prostate contour, as described in Section 4.2.3.

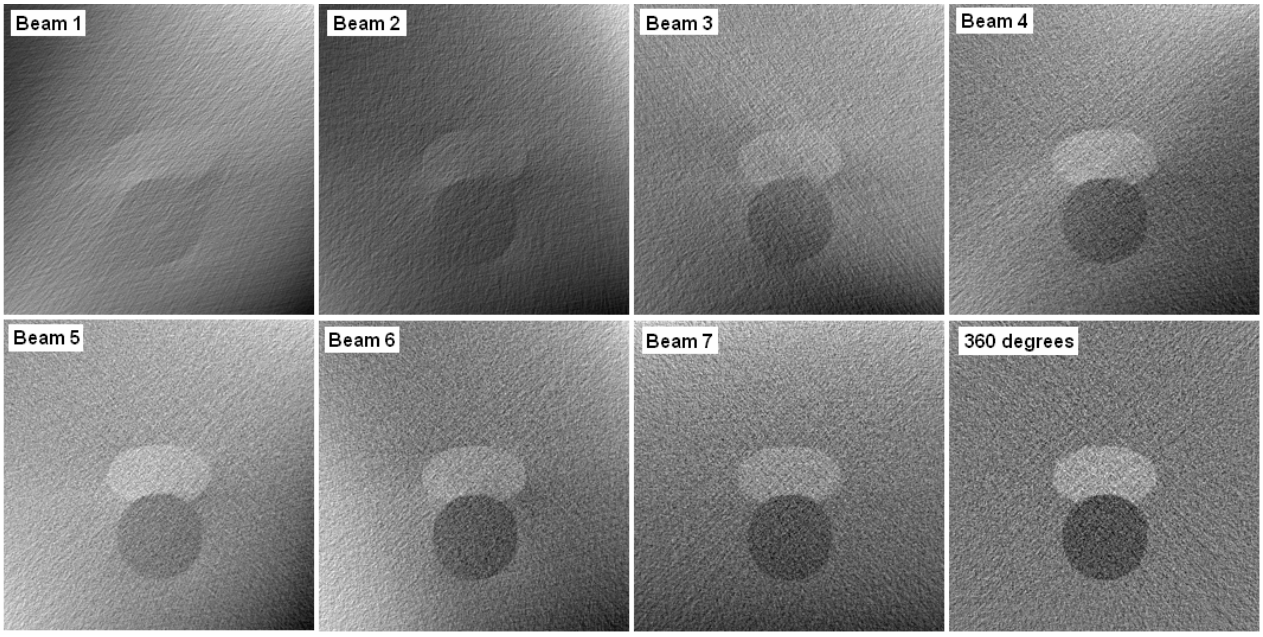


*Figure 6.16: DTS images for the medium anatomical deformation obtained with the CBCT system mounted ORTHOGONALLY to the treatment beam. The axial slice going through the isocentre of the reconstructed DTS image at each treatment position (beam 1 to 7) is shown. For comparison purposes, the reconstructed image with the full imaging range of 360 degrees is also shown.*

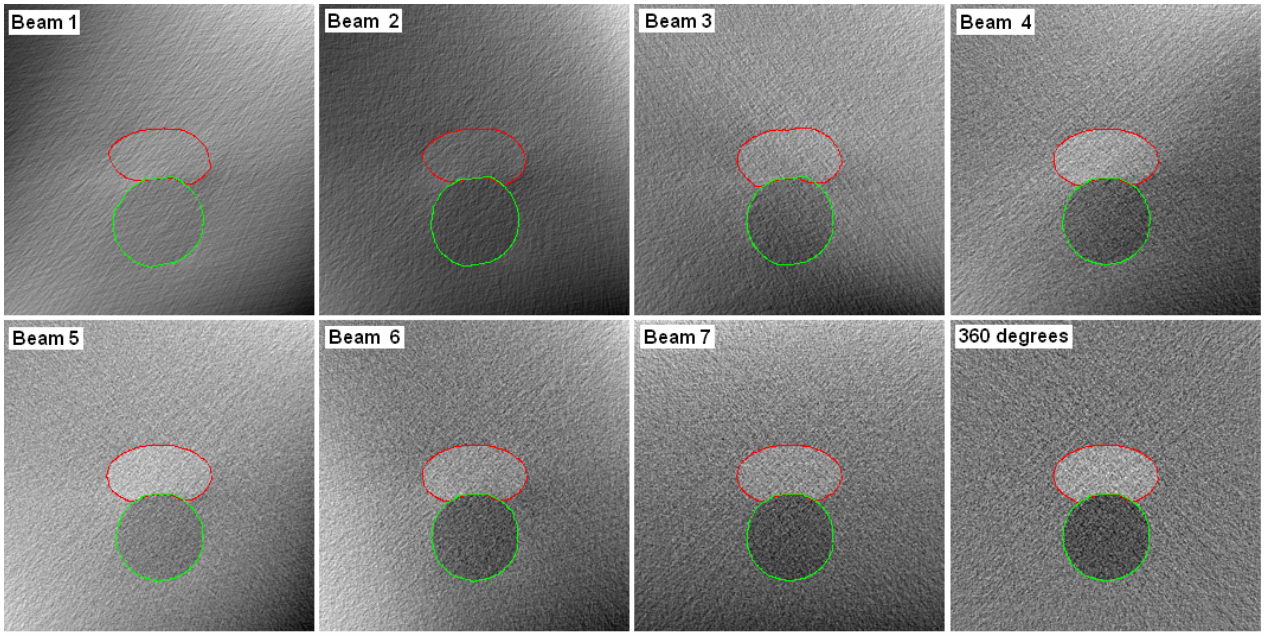


*Figure 6.17: The same axial slice from Figure 6.16 automatically segmented at each treatment position. The upper contour represents the re-segmented prostate and the lower contour represents the re-segmented rectum.*

Figure 6.18 shows DTS images obtained with the CBCT system mounted at an angle of 180 degrees with respect to the primary beam (in-line set-up in Table 6.1). Figure 6.19 displays the same axial slice automatically segmented at each treatment position using the IMM/PDAF edge detection algorithm.

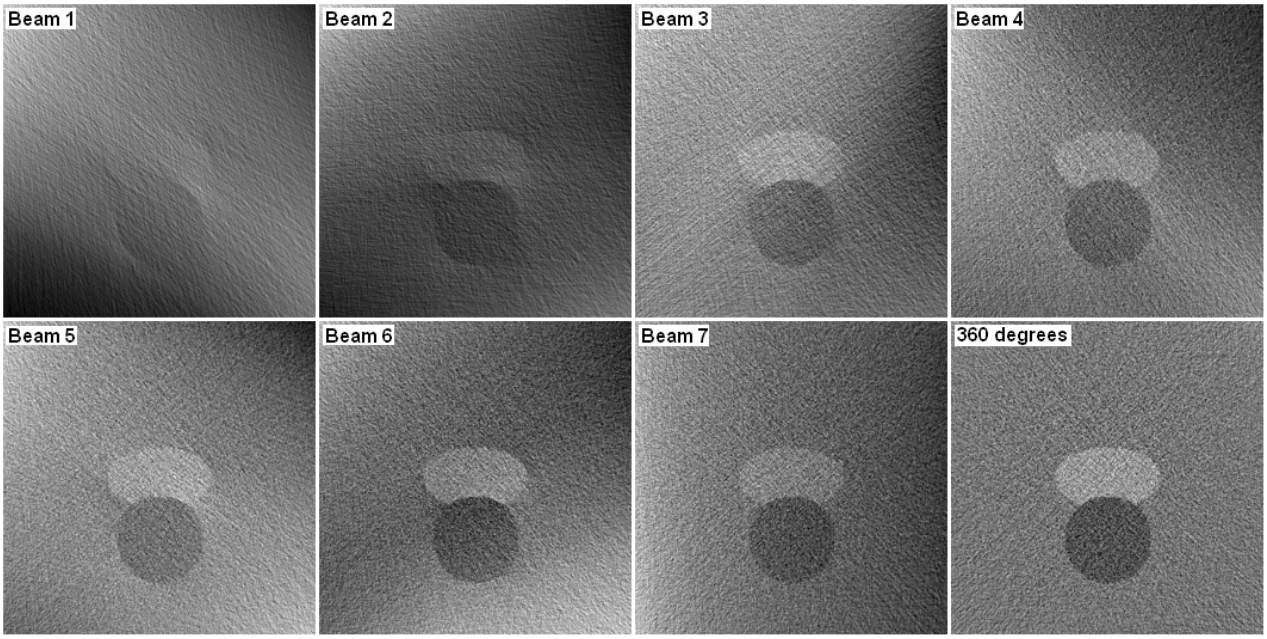


*Figure 6.18: DTS images for the medium anatomical deformation obtained with the CBCT system mounted IN-LINE (180 degrees) with respect to the primary beam. The axial slice going through the isocentre of the reconstructed DTS image at each treatment position (beam 1 to 7) is shown. For comparison purposes, the reconstructed image with the full imaging range of 360 degrees is also shown.*

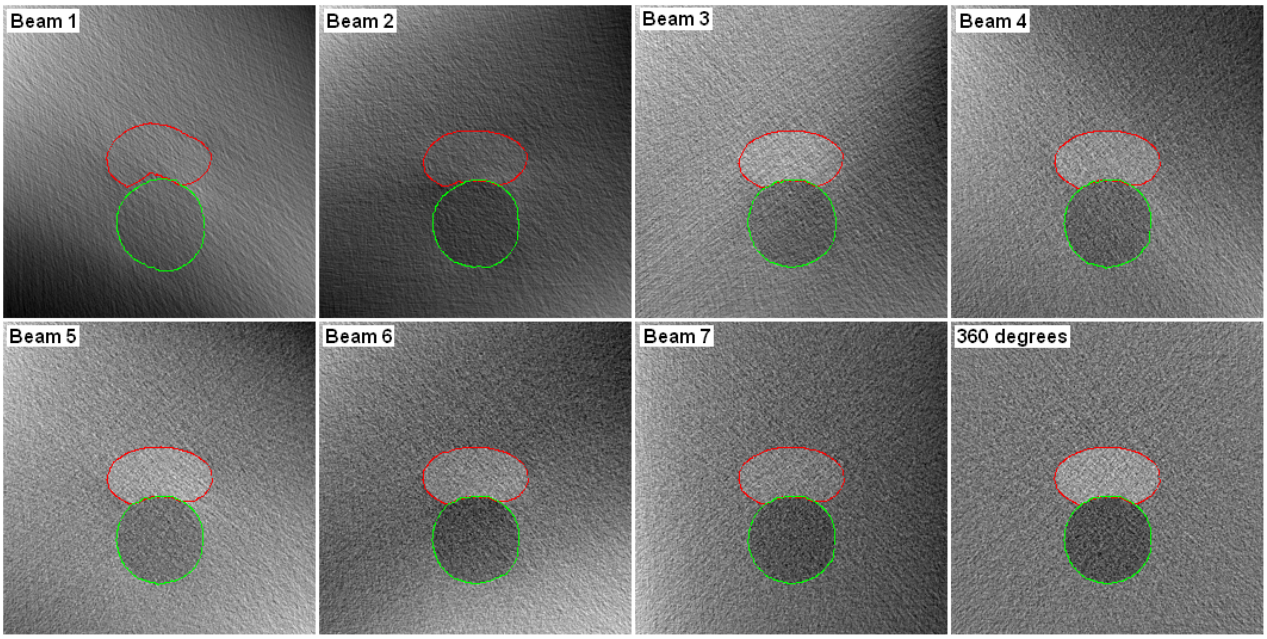


*Figure 6.19: The same axial slice from Figure 6.18 automatically segmented at each treatment position. The upper contour represents the re-segmented prostate and the lower contour represents the re-segmented rectum.*

### 6.3.4. Plan Adaptation and Radiation Delivery

To correct for deterioration of the quality of the original plan, caused by the anatomical deformations, the original treatment plan was adapted using our integrated approach. Figure 6.20 shows the objective value (cost) vs. delivered aperture during plan adaptation for the small anatomical deformation. Costs for both the orthogonal and in-line CBCT system set-ups are shown. The dashed line represents the cost value based on the current DTS image as it was used to guide the optimization. The discontinuities, marked by the arrows, correspond to the DTS image updates. They occur in between treatment positions: between aperture 6 and 7, aperture 12 and 13, aperture 18 and 19, etc. The solid line represents the actual cost value calculated based on the actual contours for the deformed anatomy. Obviously, these actual contours would not be available during the plan adaptation and therefore they were not used in the optimization. The actual cost is plotted only for comparison purposes. The horizontal dashed line in Figure 6.20 corresponds to a cost value of 10, which was the threshold for the plan to be considered clinically acceptable. Figure 6.21 shows the cost versus delivered aperture during plan adaptation for the

medium anatomical deformation. Finally, Figure 6.22 shows the cost versus delivered aperture during plan adaptation for the large anatomical deformation.

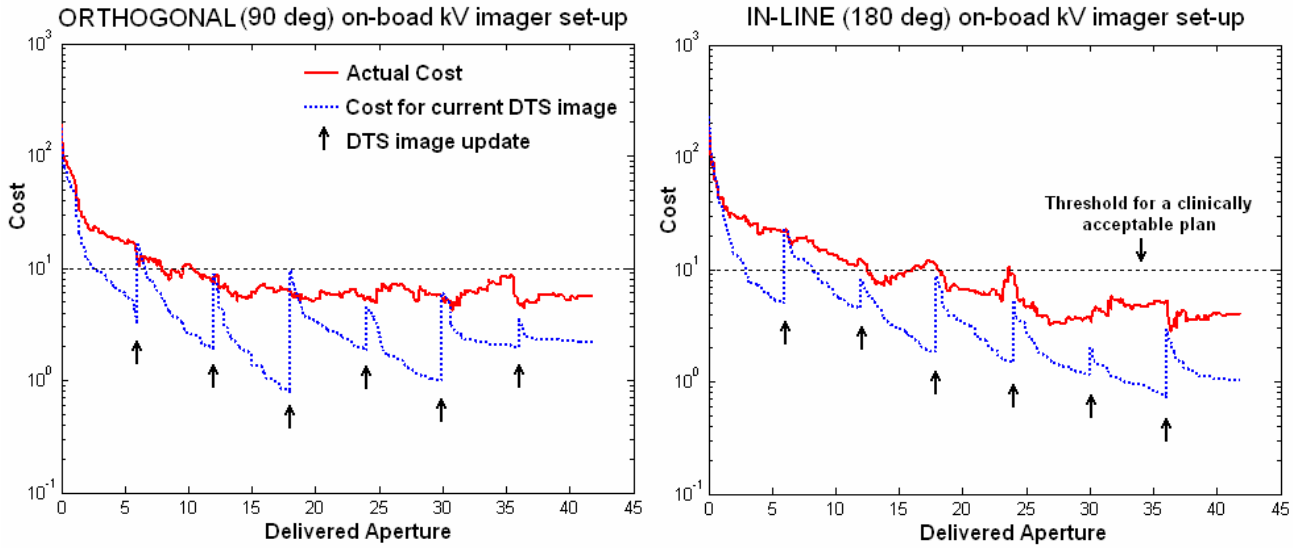


*Figure 6.20: Cost vs. delivered aperture for the SMALL anatomical deformation: a) Orthogonal CBCT system set-up and b) In-line CBCT system set-up.*

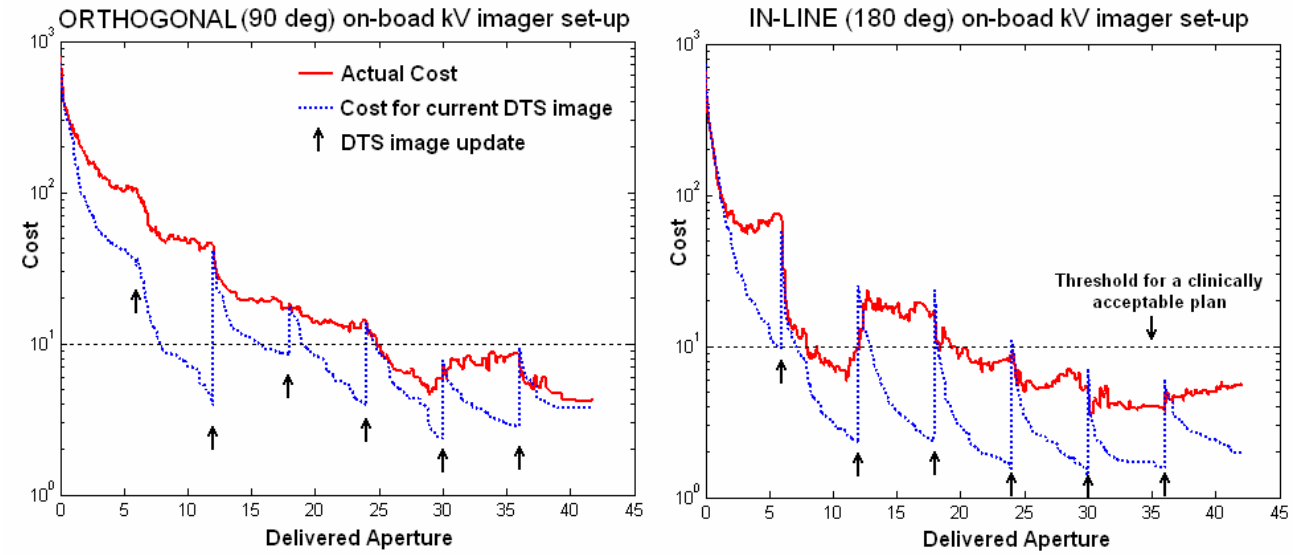


*Figure 6.21: Cost vs. delivered aperture for the MEDIUM anatomical deformation: a) Orthogonal CBCT system set-up and b) In-line CBCT system set-up.*

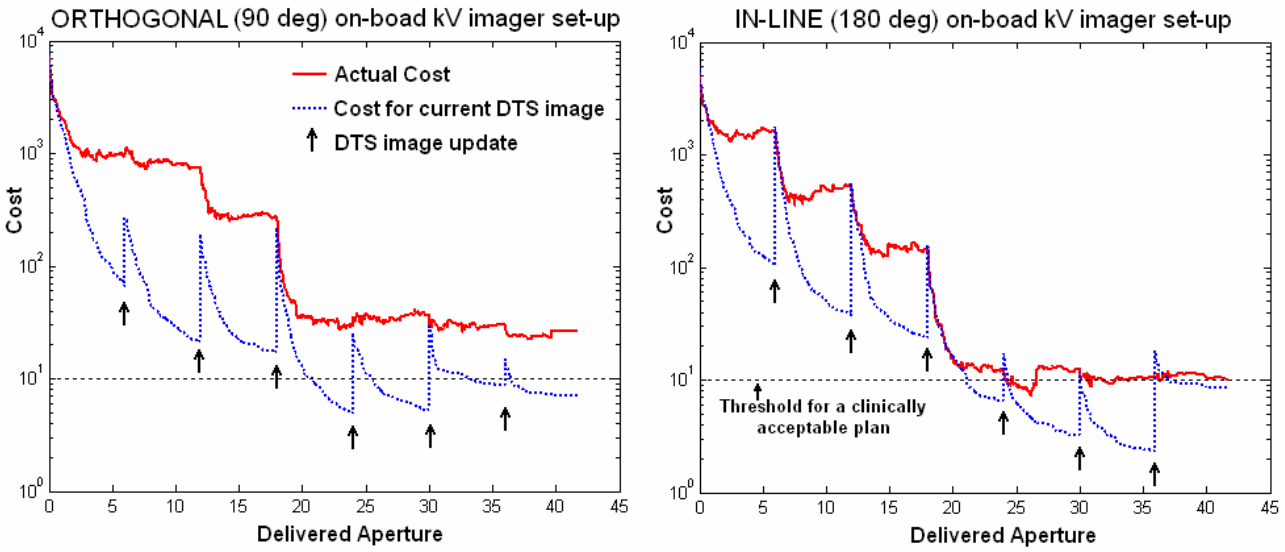


*Figure 6.22: Cost vs. delivered aperture for the LARGE anatomical deformation: a) Orthogonal CBCT system set-up and b) In-line CBCT system set-up.*

To illustrate the adapted plan quality in terms of DVHs, figure 6.23 and 6.24 display the DVHs for the adapted plans for all three anatomical deformations. Figure 6.23 shows the DVHs for the orthogonal CBCT system set-up and figure 6.24 shows the DVHs for in-line CBCT system set-up.

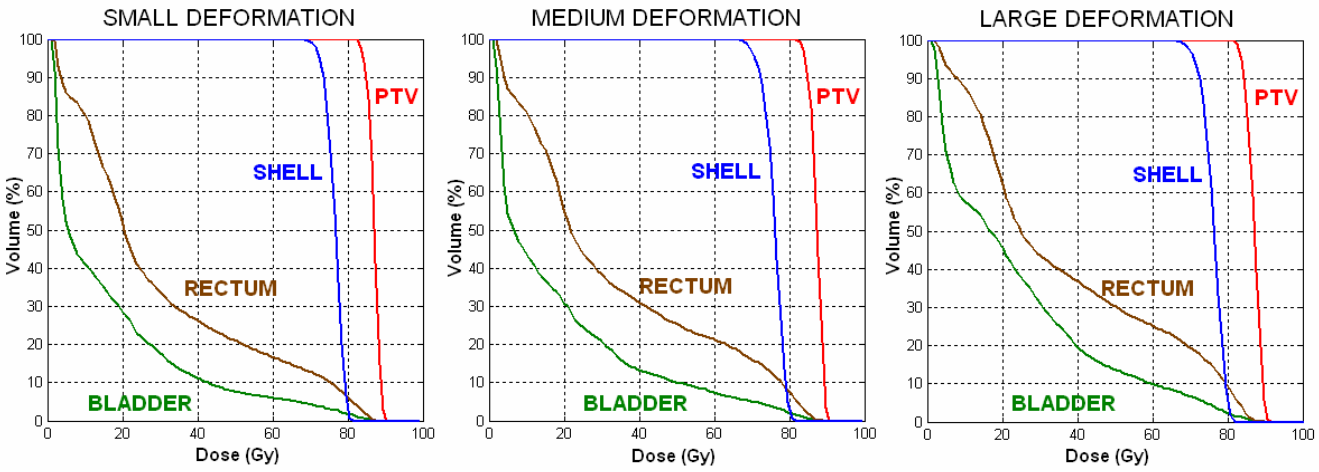


*Figure 6.23: DVHs for the adapted plans for the orthogonal CBCT system set-up for the three anatomical deformations.*

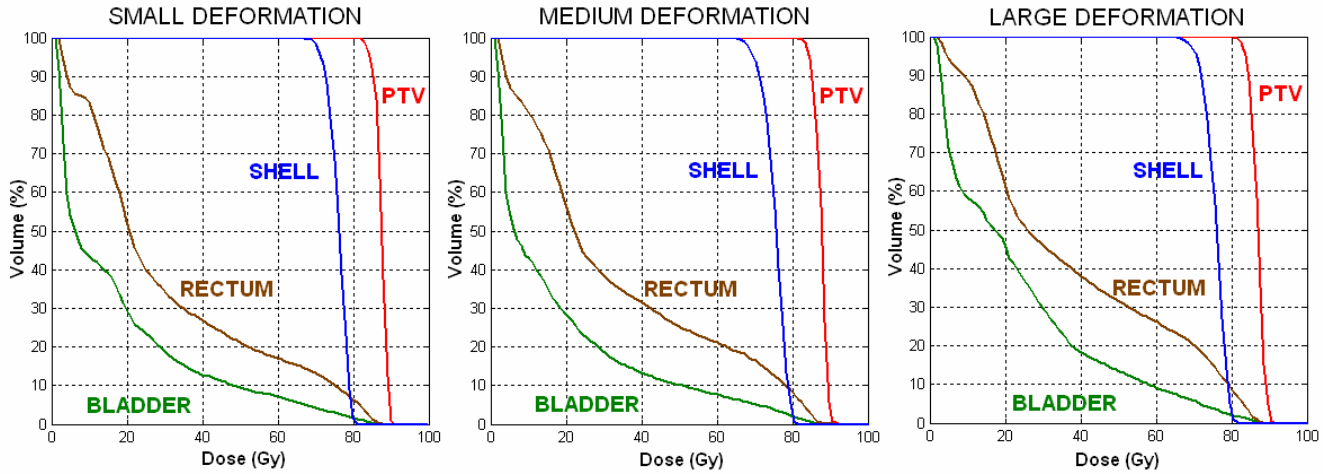


*Figure 6.24: DVHs for the adapted plans for the in-line CBCT system set-up for the three anatomical deformations.*

## 6.4 Discussion

A new approach to on-line ART is proposed in which daily image acquisition, plan adaptation and radiation delivery are integrated together. To integrate daily image acquisition with plan adaptation and radiation delivery we are proposing to acquire x-ray projections with the integrated kV imaging system, as the primary MV beam rotates between treatment positions and delivers radiation. As a result, image acquisition and radiation delivery are integrated in a single gantry rotation around the patient, instead of requiring two separate rotations, thus reducing treatment time. In this study, the concept of integrating daily imaging, plan adaptation and radiation delivery was illustrated using a Linac/CBCT system. However, the basic concepts and ideas presented here are extendable and applicable to other treatment/imaging systems. Various integrated treatment/imaging systems capable of acquiring intra-fractional images during treatment have been developed and continue to advance. For instance, Linac/MRI integration has been proposed by various research groups[102, 103] and prototype systems are currently under construction. Although these systems are not yet mature for clinical implementation, it is clear that these technologies will be available clinically in the not too distant future. Automatic segmentation is continuously improving and will also become a smaller limitation as research progresses in this area. Developments of reliable and robust automatic segmentation methods

will allow segmentation of intra-fractional images to be performed on-line. With the ability to acquire high-quality intra-fractional images and reliable automatic segmentation methods the concept of integrating patient imaging, plan adaptation and radiation delivery will become feasible in the near future.

As can be seen from Figures 6.17 and 6.19, the shapes of the re-segmented contours at different treatment positions gradually approach the shapes of the contours obtained with the full imaging range of 360 degrees. For the first couple of treatment positions the contours are clearly different from the final ones due to the insufficient number of x-ray projections available for DTS image reconstruction. This is particularly true in the direction parallel to the scan direction, where the interpolation was used to "fill" in the missing gaps of the new contour. However, enough information about the shape of the new contours appears to be available for the plan adaptation and radiation delivery to commence. It should be understood that even though the edge detection algorithm (IMM/PDAF) was used to automatically segment the DTS images in this study, any other type of segmentation method (eg. deformable registration) could be used to segment the DTS images. The concepts proposed in this study are independent of the exact nature of the segmentation method. Also, even though in this study the concept of integrating daily imaging, plan adaptation and radiation delivery was described for static gantry position IMRT, it could also be used with IMAT techniques (Section 2.3.3), such as volumetric modulated arc therapy (VMAT)[104]. One possible way of integrated the proposed approach with the VMAT delivery method is to acquire kV images in-between the MV treatment beam pulses, as the gantry rotates around the patient. This is currently under investigation.

From Figures 6.20 and 6.21, it can be seen that the final actual cost value was lower than the threshold value of 10 for the small and the medium anatomical deformations. Therefore, our integrated approach was able to adapt the original treatment plan so that it becomes clinically acceptable for the small and the medium anatomical deformations. This was true for both the orthogonal and in-line CBCT system set-ups. From Figure 6.22, the threshold value was just barely reached for the in-line CBCT system set-up. For the orthogonal CBCT system set-up, the final actual cost value was about 25, so the threshold value was not reached. Therefore, for the orthogonal CBCT system set-up our integrated approach was not able to adapt the original treatment plan so that it becomes clinically acceptable for the large anatomical deformation. Comparing the DVHs for the adapted plans from figures 6.23 and 6.24 to the DVHs for the non-adapted original plan used for the three deformed anatomies from Figure 6.15, we see that

adequate PTV coverage was regained for all of the adapted plans. The reason for not reaching the threshold cost value of 10 for the orthogonal CBCT system set-up for the large anatomical deformation was due to slight overdosing of the shell structure. The shell was created by expanding the PTV by a 5 mm margin and its purpose was to eliminate the hot spots outside the PTV. The maximum dose to the shell was constrained to 80 Gy, but it was approximately 82 Gy for the orthogonal CBCT system set-up for the large anatomical deformation. However, looking at the DVHs for this plan, even though the plan adaptation failed to reach the threshold cost value, the adapted plan was very close to being clinically acceptable even in this case.

In addition to the reduced treatment time, another advantage of imaging the patient during treatment is that the intra-fractional patient motion can be detected and accounted for. Patient motion can be detected based on the DTS image reconstructed using only the latest subset of the acquired projections (eg. last ~40 degrees)[95, 105, 106]. For example, only the projections acquired since the delivery of the last beam could be used to detect patient movement prior to the delivery of the next beam. If patient motion is detected, the patient can be moved into the right position (treatment couch translation) prior to the delivery of the next beam. This way any intra-fractional patient motion is corrected prior to the delivery of each beam. The intra-fractional patient repositioning would need to be accounted for in the subsequent DTS image reconstructions to appropriately combine the projection acquired after the patient movement with the projections acquired prior to the patient movement.

In this study, two different orientations of the CBCT system with respect to the primary beam were investigated: orthogonal set-up and in-line setup. Both of these orientations are commercially available on current linear accelerators. However, due to the spatially non-isotropic nature of the DTS images, with high image quality in the direction perpendicular to the scan direction and lower image quality parallel to the scan direction, there might be a different orientation of CBCT system with respect to the primary beam (between 0 and 90 degrees) which could be more optimal for intra-fractional patient imaging. This is currently being investigated.

The concept of integrating patient imaging, plan adaptation and radiation delivery is currently a topic of research rather than a technique that is ready for clinical implementation. However, as was the case with the ARD concept in Chapter 5, one realistic clinical implementation of this method would be to combine it with an additional off-line adaptation step. As described in Section 2.5.3, in off-line ART the treatment plan adaptation is performed after the treatment fraction and the adapted plan is used for the subsequent fraction. Any

dosimetric errors caused by the anatomical deformation in fraction n are attempted to be corrected in the subsequent fraction n+1. In this combined approach, on-line adaptation would be performed, as described in this study, as a first step of a two-step adaptation process. If on-line adaptation alone is not able to completely adapt the plan so that it becomes clinically acceptable, another off-line step would be added, in which the next treatment fraction would compensate for the remaining error. Performing on-line adaptation as a first step would be beneficial since it would make the off-line adaptation easier, with minimal extension of treatment time. For the cases in which on-line adaptation is able to successfully adapt the original plan, the off-line adaptation step would not be required.

## 6.5 Conclusion

We have shown that performing on-line ART based on intra-fractional DTS images is feasible. The original treatment plan was successfully adapted to arrive at a clinically acceptable plan for the small and the medium anatomy deformations for both orthogonal and in-line CBCT system setups. For the large anatomy deformation, the original treatment plan was successfully adapted to arrive at a clinically acceptable plan only for the in-line CBCT system setup. The advantage of integrating patient imaging and radiation delivery in a single gantry rotation around the patient is reduced treatment time and, as a consequence, increased patient throughput and decreased susceptibility to intra-fractional patient motion. In this study, the total treatment extension time due to patient imaging and plan adaptation was only 70 seconds. Another advantage is the ability to detect and correct for the intra-fractional patient motion prior to the delivery of each beam.

# Chapter 7
# Conclusions and Future Work

## 7.1 Conclusions

The primary objective of this thesis was to develop and implement a new approach to on-line ART, which has the potential to overcome the shortcomings and drawbacks of the current on-line ART procedure (Section 2.5.4). In this new approach to on-line ART, daily imaging, plan adaptation and radiation delivery were temporally integrated and performed concurrently (Figure 2.20d). The advantages of this approach are:

1) The reduction of treatment time compared to conventional on-line ART.

2) The ability to perform plan re-optimization during radiation delivery.

3) The ability to image a patient during a treatment fraction.

These advantages result in the following benefits:

1) Shorter treatment time increases patient throughput and decreases susceptibility to intra-fractional patient motion.

2) The ability to perform a complete plan re-optimization with minimal extension of treatment time.

3) The ability to detect and correct for intra-fractional patient motion.

This work was motivated by an initial study which compared four radiosurgery techniques: Circular Collimator Arcs, Dynamic Conformal Arcs, Static Conformal Fields and IMRS. This study was the first quantitative analysis of the correlation between the patient anatomy and the quality of the treatment plan produced by different radiosurgery techniques. Eighteen previously treated AVM patients were used as a sample group in this study. The patients exhibited a wide range of target size, shape and location. Multiple linear regression model was used to establish the correlation between the patient anatomy and the plan quality for different treatment techniques. The results from the model were used to compare the effectiveness of different treatment techniques based on the patient anatomy. A number of

different relationships, both qualitative and quantitative, were discovered between the patient anatomy and dosimetric parameters for different techniques. The model was employed to predict the overall plan quality of different radiosurgery techniques based on the individual patient anatomy. The predicted dosimetric parameters were in close agreement with the actual data, indicating the effectiveness of the model. The results were used to predetermine the optimal radiosurgery technique/techniques for a number of patients, without having to create multiple treatment plans with different techniques. Therefore, this method shows great potential to be a valuable tool for selecting the most effective treatment technique. Even though this method was developed for radiosurgery of AVMs, it is completely general and could be used for other types of radiation therapy treatments. A limitation of this study is the inherent assumption that the patient anatomy does not change during the entire treatment process. For radiosurgery treatments, this assumption is somewhat justified since all the radiation is delivered in only one fraction. However, for a majority of other radiation therapy treatments the radiation delivery is fractionated (30–40 fractions), and the assumption that the patient anatomy remains the same throughout the whole treatment process (several weeks) is no longer valid. This led to the second part of this thesis, on-line ART.

In Chapter 4, methods for accelerating plan adaptation (re-optimization) using the DAO technique were investigated (Figure 2.20b). This work was the first investigation of using the DAO for on-line ART. An anatomical model representing a typical prostate anatomy was created. To simulate a range of clinically relevant anatomical deformations, four deformed anatomies were created by systematically deforming the original anatomy by various amounts (0.25, 0.50, 0.75 and 1.0 cm). A series of techniques was described in which the original DAO treatment plan was adapted in order to correct for the deterioration of dose distribution quality caused by the anatomical deformations. It was found that the average time needed for the original plan adaptation to arrive at the clinically acceptable plan was roughly half of the time needed for a complete plan regeneration, for all four anatomical deformations. Furthermore, through modification of the DAO algorithm the optimization search space was reduced and the plan adaptation was significantly accelerated. For the first anatomical deformation (0.25 cm) the optimization efficiency was increased by a factor of about 6 compared to the complete plan regeneration. For the 0.50 cm and 0.75 cm deformations the optimization efficiency was increased by a factor of roughly 3 compared to the complete plan regeneration. However, for the anatomical deformation of 1.0 cm the reduction of the search space during plan adaptation did

not result in any efficiency improvement over the original (non-modified) plan adaptation. For our model, the anatomical deformation of 1.0 cm demonstrated the limit for which the plan adaptation could not be accelerated.

In Chapter 5, a new approach to on-line ART was proposed and investigated, in which accelerated plan adaptation and radiation delivery were integrated together and performed concurrently (Figure 2.20c). This approach is referred to as Adaptive Radiation Delivery (ARD). A fundamental advantage of ARD is the fact that radiation delivery can start almost immediately after daily imaging. Most of the plan re-optimization is performed during the radiation delivery, so the time spent adapting the original plan does not significantly increase the overall treatment time. As a consequence, the effective time allotted for plan adaptation is drastically reduced. For the 0.25 cm, 0.5 cm and 0.75 cm anatomical deformations, the treatment time was increased by only 2 sec, 4 sec and 6 sec, respectively, as compared to no plan adaptation. For the anatomical deformation of 1.0 cm the time increase was substantially larger. The anatomical deformation of 1.0 cm represents an extreme case, which is rarely observed for the prostate, and again demonstrates the limit of this approach.

Finally, in Chapter 6, daily patient imaging, accelerated plan adaptation and radiation delivery were all temporally integrated and performed concurrently (Figure 2.20d). Intra-fractional images were acquired using an integrated Linac/CBCT system. As the primary MV beam rotated between treatment positions and delivered radiation, x-ray projections were acquired with the CBCT system. As a result, image acquisition and radiation delivery were integrated in a single gantry rotation around the patient, instead of requiring two separate rotations. The intra-fractional DTS images were used to re-optimize the original treatment plan to account for anatomy deformations. The original treatment plan was successfully adapted to arrive at a clinically acceptable plan for the small and the medium anatomy deformations for both orthogonal and in-line CBCT system setups. For the large anatomy deformation, the original treatment plan was successfully adapted to arrive at a clinically acceptable plan only for the in-line CBCT system setup. The advantage of integrating patient imaging and radiation delivery in a single gantry rotation around the patient is reduced treatment time and, as a consequence, increased patient throughput and decreased susceptibility to intra-fractional patient motion. In this study, the total treatment extension time due to the patient imaging and plan adaptation was only 70 seconds. Another advantage is the ability to detect and correct for intra-fractional patient motion prior to the delivery of each beam. Therefore, this approach has the

ability to overcome the shortcomings and drawbacks of the current on-line ART procedure. This study was the first investigation into the feasibility of performing on-line ART based on intra-fractionally acquired images. These ideas and concepts lead toward the ultimate goal of near real-time ART.

## 7.2 Future Work

The primary goal of this thesis was to introduce, develop and implement ideas and concepts that lead toward real-time ART. This work guides to many different avenues of research to be further explored and investigated. These are some of the research projects that are currently under investigation, or will be investigated in the near future:

### Integrated Linac/CBCT Control System

Translating these ideas and concepts into clinical practice will require development and implementation of a new linac control system. The current linac control system is not designed to acquire images with the CBCT system, adapt the original treatment plan and deliver radiation based on the adapted plan during a single linac gantry rotation. To achieve this some major linac software upgrades will need to be implemented. Figure 7.1 schematically illustrates the integration of various components required to combine patient imaging, plan adaptation and radiation delivery. In Figure 7.1, x-ray projections acquired by the CBCT system are fed into the Image Analysis component, where a DTS image is reconstructed and segmented. The new contours are transferred to the Plan Adaptation component, where the original treatment plan is re-optimized based on the new contours. The re-optimized aperture shapes (MLC positions) and aperture weights (MUs) obtained from the Plan Adaptation component are fed into the Linac Control component, which controls the MLC leaf positions and number of MUs delivered by each aperture. Effectively orchestrating all of these components together in near-real time will require developments of both linac control software and hardware.

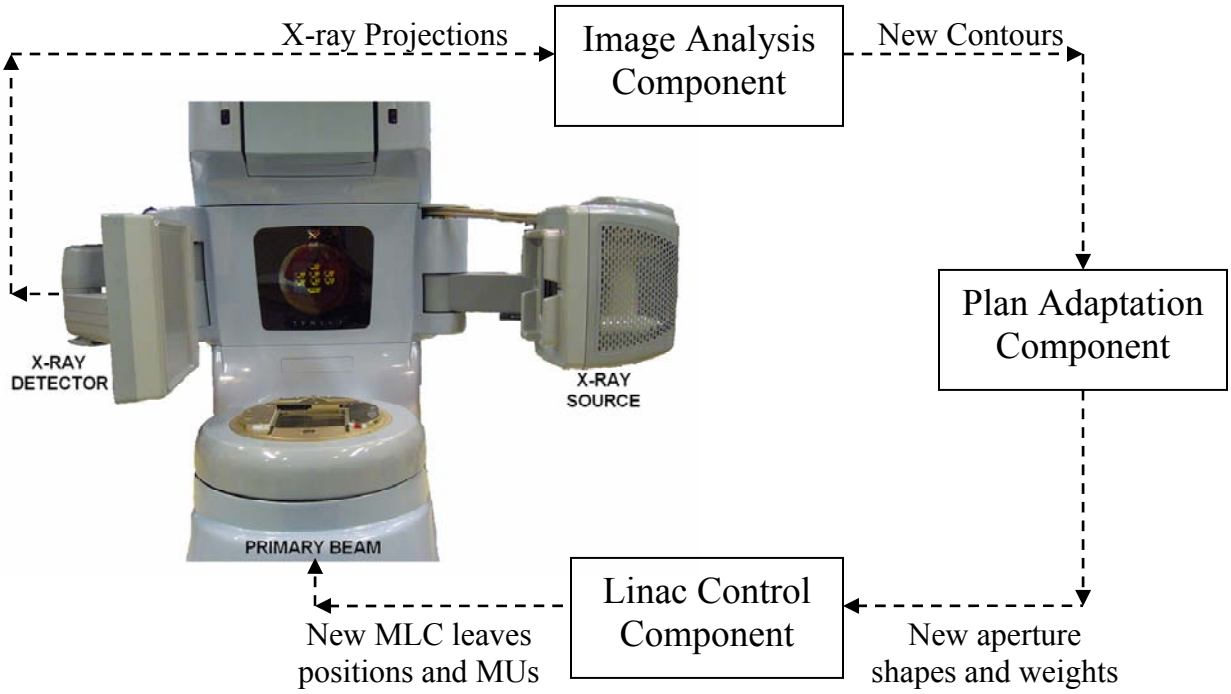


*Figure 7.1: Integration of various components required to combine patient imaging, plan adaptation and radiation delivery using a Linac/CBCT system.*

### Integration with VMAT Delivery

In this thesis, the concept of integrating daily imaging, plan adaptation and radiation delivery was described for static gantry IMRT. However, the concept could also be extended and used with IMAT techniques (Section 2.3.3), such as VMAT[104]. One possible way of integrated the proposed approach with the VMAT delivery method is to acquire kV images in-between the MV treatment beam pulses, as the gantry rotates around the patient and delivers radiation. The advantage of VMAT delivery is the significant reduction of radiation delivery time as compared to static gantry IMRT delivery. The shorter radiation delivery time further reduces the overall treatment time and decreases the susceptibility to intra-fractional patient motion.

### Alternative CBCT System Orientation

In this thesis, two different orientations of the CBCT system with respect to the primary beam were investigated: orthogonal set-up and in-line setup. Both of these orientations are

commercially available on current linear accelerators. However, due to the spatially non-isotropic nature of the DTS images, with high image quality in the direction perpendicular to the scan direction and lower image quality parallel to the scan direction, there might be a different orientation of the CBCT system with respect to the primary beam (between 0 and 90 degrees) which could be more optimal for intra-fractional patient imaging. This is currently under investigation.

### Combination with Off-line Adaptation

As mentioned before, one currently feasible clinical implementation of this concept would be to combine it with an additional off-line adaptation step. As described in Section 2.5.3, in off-line ART the treatment plan adaptation is performed after the treatment fraction and the adapted plan is used for the subsequent fraction. Any dosimetric errors caused by the anatomical deformation in fraction n are attempted to be corrected in the subsequent fraction n+1. In the combined approach, on-line adaptation would be performed, as described in this study, as a first step of a two-step adaptation process. If on-line adaptation alone is not able to completely adapt the plan so that it becomes clinically acceptable, another off-line step would be added, in which the next treatment fraction would compensate for the remaining error. Performing on-line adaptation as a first step would be beneficial since it would make off-line adaptation easier, with minimal extension of treatment time. For the cases in which on-line adaptation is able to successfully adapt the original plan, the off-line adaptation step would not be required. The advantage of this approach is that the off-line adaptation step would act as a "safety net" to be used when on-line adaptation alone is not able to adapt the plan so that it becomes clinically acceptable.

### Detection of Non-rigid Intra-fractional Anatomy Deformations

One of the advantages of imaging the patient during the treatment is the ability to detect and account for intra-fractional patient motion. As mentioned before, patient motion can be detected based on the DTS images reconstructed using only the latest subset of the acquired projections (eg. last ~40 degrees)[95, 105, 106]. For example, only the projections acquired since the delivery of the last beam could be used to detect patient movement prior to the delivery of the next beam. If patient motion is detected, the patient can be moved into the right position (treatment couch

translation) prior to the delivery of the next beam. This way any intra-fractional patient motion is corrected prior to the delivery of each beam.

In addition to intra-fractional patient motion, the non-rigid intra-fractional anatomy deformations could also in principle be detected based on the DTS image reconstructed using only the latest subset of the acquired projections. The number of projections (i.e. imaging range) required for successful characterization of non-rigid intra-fractional anatomical deformations needs to be investigated. Having more projections included into the reconstruction increases the spatial information incorporated into the DTS image, but decreases the temporal resolution of the DTS image (Figure 7.2). The optimal balance between spatial information and temporal resolution of the reconstructed DTS images will probably be highly dependent on the treatment site. For example, for treatment sites prone to intra-fractional anatomical deformations (e.g. lung, prostate) more emphasis will need to be placed on good temporal resolution, whereas for more anatomically stable sites (e.g. intracranial sites) increasing the spatial information, at the expense of a lower temporal resolution, will be more appropriate.

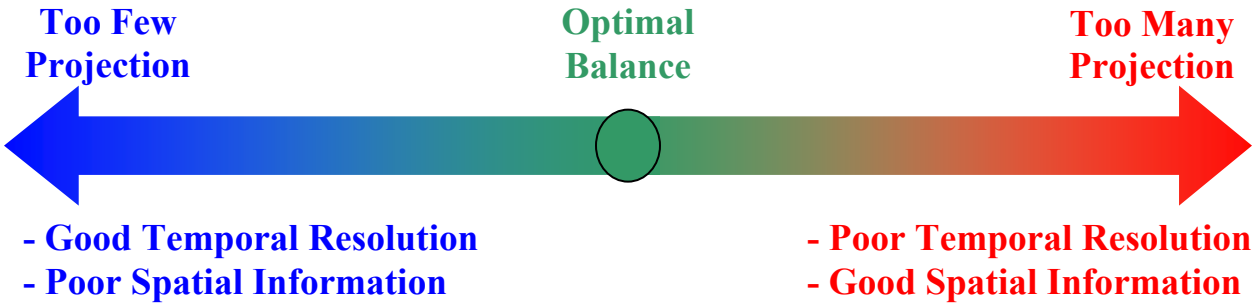


*Figure 7.2: Increasing the number of projections (i.e. imaging range) included into the DTS image reconstruction increases the spatial information incorporated into the DTS image but decreases the temporal resolution of the DTS image, and vice versa. The optimal balance between spatial information and temporal resolution of the reconstructed DTS images will probably be highly dependent on the treatment site.*

### Automatic On-line Segmentation for Real Patient DTS Images

Even though the edge detection algorithm (IMM/PDAF) was used to automatically segment the DTS images in this study, any other type of segmentation method (eg. deformable registration) could be used to segment the DTS images. The concepts proposed in this study are independent of the exact nature of the segmentation method. A very important component of a successful clinical implementation of near real-time ART is the development of a reliable and robust automatic segmentation method, which will allow on-line segmentation of intra-fractional DTS images to be performed for real patient DTS images. Currently, the commercially available segmentation methods are not able to reliably and consistently segment real patient DTS images. Therefore, before near real-time ART can be successfully put into clinical practice, developments and advances in automatic segmentation methods will be an essential prerequisite. As research progresses in this area and automatic segmentation methods improve and advance, they will become a smaller limitation for clinical implementation of near real-time ART.

### Real-time Quality Assurance (QA)

With the treatment plan re-optimized and radiation delivered to the patient in near real-time, the methods for real-time dose verification and treatment plan QA will need to be developed. This is already an active area of research. One technique that shows potential for real-time dose verification and treatment plan QA employs the use of an Electronic Portal Imaging Device (EPID). EPID is an amorphous silicon flat panel detector placed directly opposite to the primary MV beam (Figure 7.3). The benefit of the EPID is the ability to digitally capture images produced by the treatment MV beam prior to and during the treatment. These systems have been used in clinical practice for many years and come as standard equipment on every modern linac. Recently, EPID images acquired during the treatment have been used to reconstruct a 3D dose distribution in phantoms/patients[107]. With their ability to monitor the delivered dose to the patient during the treatment, these systems show potential to be used for real-time dose verification and treatment plan QA.

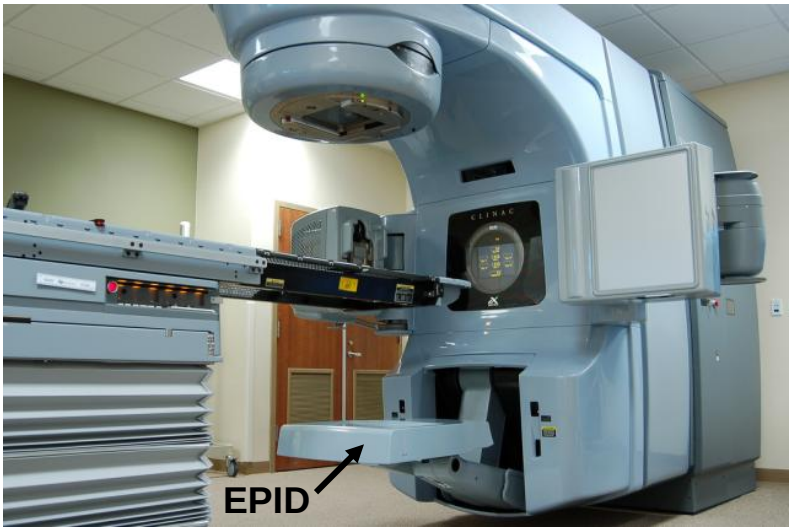


*Figure 7.3: Electronic Portal Imaging Device (EPID) is an amorphous silicon flat panel detector placed directly opposite to the primary MV beam. With the ability to digitally capture images produced by the treatment MV beam prior to and during the treatment, these systems show potential to be used for real-time dose verification and treatment plan quality assurance (QA).*

## Integrated Linac/MRI System

In this study, the concept of integrating daily imaging, plan adaptation and radiation delivery was illustrated using a Linac/CBCT system. However, basic concepts and ideas presented here are extendable and applicable to other treatment/imaging systems. Various integrated treatment/imaging systems capable of acquiring intra-fractional images during treatment have been developed and continue to advance. For instance, Linac/MRI integration has been proposed by various research groups[102, 103] and prototype systems are currently under construction. Although these systems are not yet mature for clinical implementation, it is clear that these technologies will be available clinically in the not too distant future. Use of MRI real-time imaging has the potential to provide higher quality images of tumors and organs than x-rays and CT, with no extra radiation given to the patient. With the ability to safely acquire high-quality intra-fractional images and reliable automatic segmentation methods, the concept of integrating patient imaging, plan adaptation and radiation delivery promises to become a part of clinical practice in the near future.

# Bibliography


1. *Intensity Modulated Radiation Therapy Collaborative Working Group*, "Intensity-modulated radiotherapy: current status and issues of interest," *Int.J.Radiat.Oncol.Biol.Phys.* 51 (4), 880-914 (2001).

2. Y. Xie, D. Djajaputra, C. R. King, S. Hossain, L. Ma and L. Xing, "Intrafractional motion of the prostate during hypofractionated radiotherapy," *Int.J.Radiat.Oncol.Biol.Phys.* 72 (1), 236-246 (2008).

3. H. D. Kogelnik, "The history and evolution of radiotherapy and radiation oncology in Austria," *Int.J.Radiat.Oncol.Biol.Phys.* 35 (2), 219-226 (1996).

4. M. Ritter, "Rationale, conduct, and outcome using hypofractionated radiotherapy in prostate cancer," *Semin.Radiat.Oncol.* 18 (4), 249-256 (2008).

5. Q. N. Nguyen and E. L. Chang, "Emerging role of proton beam radiation therapy for chordoma and chondrosarcoma of the skull base," *Curr.Oncol.Rep.* 10 (4), 338-343 (2008).

6. C. F. von Essen, M. A. Bagshaw, S. E. Bush, A. R. Smith and M. M. Kligerman, "Long-term results of pion therapy at Los Alamos," *Int.J.Radiat.Oncol.Biol.Phys.* 13 (9), 1389-1398 (1987).

7. L. D. Skarsgard, R. M. Henkelman and C. J. Eaves, "Pions for radiotherapy at TRIUMF," *J.Can.Assoc.Radiol.* 31 (1), 3-12 (1980).

8. G. E. Laramore, J. T. Griffith, M. Boespflug, J. G. Pelton, T. Griffin, B. R. Griffin, K. J. Russell, W. Koh, R. G. Parker and L. W. Davis, "Fast neutron radiotherapy for sarcomas of soft tissue, bone, and cartilage," *Am.J.Clin.Oncol.* 12 (4), 320-326 (1989).

9. L. Cohen, "Complications of fast neutron therapy," *Recent Results Cancer Res.* 150, 156-169 (1998).

10. D. Yan, D. Lockman, A. Martinez, J. Wong, D. Brabbins, F. Vicini, J. Liang and L. Kestin, "Computed tomography guided management of interfractional patient variation," *Semin.Radiat.Oncol.* 15 (3), 168-179 (2005).

11. M. Levivier, N. Massager, D. Wikler, D. Devriendt and S. Goldman, "Integration of functional imaging in radiosurgery: the example of PET scan," *Prog.Neurol.Surg.* 20, 68-81 (2007).

12. D. E. Heron, R. P. Smith and R. S. Andrade, "Advances in image-guided radiation therapy--the role of PET-CT," *Med.Dosim.* 31 (1), 3-11 (2006).

13. T. R. Mackie, J. Kapatoes, K. Ruchala, W. Lu, C. Wu, G. Olivera, L. Forrest, W. Tome, J. Welsh, R. Jeraj, P. Harari, P. Reckwerdt, B. Paliwal, M. Ritter, H. Keller, J. Fowler and M. Mehta, "Image guidance for precise conformal radiotherapy," *Int.J.Radiat.Oncol.Biol.Phys.* 56 (1), 89-105 (2003).

14. H. E. Johns and J. R. Cunningham, *The Physics of Radiology*, 4th ed., Charles C. Thomas, Springfield (Illinois), 1983.

15. F. H. Attix, *Introduction to Radiological Physics and Radiation Dosimetry*, John Wiley and Sons, Inc., Hoboken (NJ), 1986.

16. F. M. Khan, *The Physics of Radiation Therapy*, 3rd ed., Lippincott Williams and Wilkins, Philadelphia (PA), 2003.

17. T. Bortfeld, J. Stein and K. Preiser, "Clinically relevant intensity modulation optimization using physical criteria," *In Proceedings of the XII International Conference on the Use of Computers in Radiation Therapy, Salt Lake City, UT,* , 1-4 (1997).

18. Y. Watanabe, "Point dose calculations using an analytical pencil beam kernel for IMRT plan checking," *Phys.Med.Biol.* 46 (4), 1031-1038 (2001).

19. L. Dong, A. Shiu, S. Tung and K. Hogstrom, "A pencil-beam photon dose algorithm for stereotactic radiosurgery using a miniature multileaf collimator," *Med.Phys.* 25 (6), 841-850 (1998).

20. A. M. Bergman, K. Otto and C. Duzenli, "The use of modified single pencil beam dose kernels to improve IMRT dose calculation accuracy," *Med.Phys.* 31 (12), 3279-3287 (2004).

21. S. J. Thomas, "A modified power-law formula for inhomogeneity corrections in beams of high-energy x rays," *Med.Phys.* 18 (4), 719-723 (1991).

22. M. Engelsman, E. M. Damen, P. W. Koken, A. A. van 't Veld, K. M. van Ingen and B. J. Mijnheer, "Impact of simple tissue inhomogeneity correction algorithms on conformal radiotherapy of lung tumours," *Radiother.Oncol.* 60 (3), 299-309 (2001).

23. X. Zhang, H. Liu, X. Wang, L. Dong, Q. Wu and R. Mohan, "Speed and convergence properties of gradient algorithms for optimization of IMRT," *Med.Phys.* 31 (5), 1141-1152 (2004).

24. S. V. Spirou and C. S. Chui, "A gradient inverse planning algorithm with dose-volume constraints," *Med.Phys.* 25 (3), 321-333 (1998).

25. E. J. Hall and C. S. Wuu, "Radiation-induced second cancers: the impact of 3D-CRT and IMRT," *Int.J.Radiat.Oncol.Biol.Phys.* 56 (1), 83-88 (2003).

26. F. Umansky, Y. Shoshan, G. Rosenthal, S. Fraifeld and S. Spektor, "Radiation-induced meningioma," *Neurosurg.Focus.* 24 (5), E7 (2008).

27. E. Rapiti, G. Fioretta, H. M. Verkooijen, R. Zanetti, F. Schmidlin, H. Shubert, A. Merglen, R. Miralbell and C. Bouchardy, "Increased risk of colon cancer after external radiation therapy for prostate cancer," *Int.J.Cancer* 123 (5), 1141-1145 (2008).

28. W. De Gersem, F. Claus, C. De Wagter, B. Van Duyse and W. De Neve, "Leaf position optimization for step-and-shoot IMRT," *Int.J.Radiat.Oncol.Biol.Phys.* 51 (5), 1371-1388 (2001).

29. D. M. Shepard, M. A. Earl, X. A. Li, S. Naqvi and C. Yu, "Direct aperture optimization: a turnkey solution for step-and-shoot IMRT," *Med.Phys.* 29 (6), 1007-1018 (2002).

30. C. Cotrutz and L. Xing, "Segment-based dose optimization using a genetic algorithm," *Phys.Med.Biol.* 48 (18), 2987-2998 (2003).

31. J. L. Bedford and S. Webb, "Constrained segment shapes in direct-aperture optimization for step-and-shoot IMRT," *Med.Phys.* 33 (4), 944-958 (2006).

32. S. Kirkpatrick, C. Gellat and M. Vecchi, "Optimization by Simulated Annealing," *Science* 220, 671-680 (1983).

33. D. Yan, D. Lockman, A. Martinez, J. Wong, D. Brabbins, F. Vicini, J. Liang and L. Kestin, "Computed Tomography Guided Management of Interfractional Patient Variation," *Semin.Radiat.Oncol.* 15, 168-179 (2005).

34. T. Mackie, J. Kapatoes, K. Ruchala, W. Lu and C. Wu, "Image Guidance for Precise Conformal Radiotherapy," *Int.J.Radiat.Oncol.Biol.Phys.* 56 (1), 89-105 (2003).

35. J. R. Wong, Z. Gao, M. Uematsu, S. Merrick, N. P. Machernis, T. Chen and C. W. Cheng, "Interfractional Prostate Shifts: Review of 1870 Computed Tomography (CT) Scans Obtained During Image-Guided Radiotherapy Using CT-on-Rails for the Treatment of Prostate Cancer," *Int.J.Radiat.Oncol.Biol.Phys.* , (2008).

36. A. Y. Fung, J. R. Wong, C. W. Cheng, S. Lisa Grimm and M. Uematsu, "A comparison of two image fusion techniques in ct-on-rails localization of radiation delivery," *Phys.Med.* 21 (3), 113-119 (2005).

37. D. A. Jaffray and J. H. Siewerdsen, "Cone-beam computed tomography with a flat-panel imager: initial performance characterization," *Med.Phys.* 27 (6), 1311-1323 (2000).

38. D. Letourneau, J. W. Wong, M. Oldham, M. Gulam, L. Watt, D. A. Jaffray, J. H. Siewerdsen and A. A. Martinez, "Cone-beam-CT guided radiation therapy: technical implementation," *Radiother.Oncol.* 75 (3), 279-286 (2005).

39. D. A. Jaffray, J. H. Siewerdsen, J. W. Wong and A. A. Martinez, "Flat-panel cone-beam computed tomography for image-guided radiation therapy," *Int.J.Radiat.Oncol.Biol.Phys.* 53 (5), 1337-1349 (2002).

40. J. H. Siewerdsen and D. A. Jaffray, "Cone-beam computed tomography with a flat-panel imager: magnitude and effects of x-ray scatter," *Med.Phys.* 28 (2), 220-231 (2001).

41. J. H. Siewerdsen and D. A. Jaffray, "Optimization of x-ray imaging geometry (with specific application to flat-panel cone-beam computed tomography)," *Med.Phys.* 27 (8), 1903-1914 (2000).

42. M. Birkner, D. Yan, M. Alber, J. Liang and F. Nusslin, "Adapting inverse planning to patient and organ geometrical variation: algorithm and implementation," *Med.Phys.* 30 (10), 2822-2831 (2003).

43. H. Rehbinder, C. Forsgren and J. Lof, "Adaptive radiation therapy for compensation of errors in patient setup and treatment delivery," *Med.Phys.* 31 (12), 3363-3371 (2004).

44. D. Yan, F. Vicini, J. Wong and A. Martinez, "Adaptive radiation therapy," *Phys.Med.Biol.* 42 (1), 123-132 (1997).

45. C. Wu, R. Jeraj, G. H. Olivera and T. R. Mackie, "Re-optimization in adaptive radiotherapy," *Phys.Med.Biol.* 47 (17), 3181-3195 (2002).

46. C. Wu, R. Jeraj, W. Lu and T. R. Mackie, "Fast treatment plan modification with an over-relaxed Cimmino algorithm," *Med.Phys.* 31 (2), 191-200 (2004).

47. K. K. Brock, L. A. Dawson, M. B. Sharpe, D. J. Moseley and D. A. Jaffray, "Feasibility of a novel deformable image registration technique to facilitate classification, targeting, and monitoring of tumor and normal tissue," *Int.J.Radiat.Oncol.Biol.Phys.* 64 (4), 1245-1254 (2006).

48. B. C. Davis, M. Foskey, J. Rosenman, L. Goyal, S. Chang and S. Joshi, "Automatic segmentation of intra-treatment CT images for adaptive radiation therapy of the prostate," *Med Image Comput Comput Assist Interv Int Conf Med Image Comput Comput Assist Interv* 8, 442-450 (2005).

49. M. Foskey, B. Davis, L. Goyal, S. Chang, E. Chaney, N. Strehl, S. Tomei, J. Rosenman and S. Joshi, "Large deformation three-dimensional image registration in image-guided radiation therapy," *Phys.Med.Biol.* 50 (24), 5869-5892 (2005).

50. R. Munbodh, D. A. Jaffray, D. J. Moseley, Z. Chen, J. P. Knisely, P. Cathier and J. S. Duncan, "Automated 2D-3D registration of a radiograph and a cone beam CT using line-segment enhancement," *Med.Phys.* 33 (5), 1398-1411 (2006).

51. Y. Feng, C. Castro-Pareja, R. Shekhar and C. Yu, "Direct aperture deformation: an interfraction image guidance strategy," *Med.Phys.* 33 (12), 4490-4498 (2006).

52. R. Mohan, X. Zhang, H. Wang, Y. Kang, X. Wang, H. Liu, K. Ang, D. Kuban and L. Dong, "Use of Deformed Intensity Distributions for On-line Modification of Image-Guided IMRT to Account for Interfractional Anatomic Changes," *Int.J.Radiat.Oncol.Biol.Phys.* 61 (4), 1258-1266 (2005).

53. L. Court, L. Dong, A. Lee, R. Cheung, D. Bonnen, J. O`Daniel, H. Wang, R. Mohan and D. Kuban, "An Automatic CT-Guided Adaptive Radiation Therapy Technique by On-line Modification of MLC Leaf Positions for Prostate Cancer," *Int.J.Radiat.Oncol.Biol.Phys.* 62 (1), 154-163 (2005).

54. L. E. Court, R. B. Tishler, J. Petit, R. Cormack and L. Chin, "Automatic online adaptive radiation therapy techniques for targets with significant shape change: a feasibility study," *Phys.Med.Biol.* 51 (10), 2493-2501 (2006).

55. W. Song, B. Schaly, G. Bauman, J. Battista and J. Van Dyk, "Image-guided adaptive radiation therapy (IGART): Radiobiological and dose escalation considerations for localized carcinoma of the prostate," *Med.Phys.* 32 (7), 2193-2203 (2005).

56. A. M. Amer, J. Mott, R. I. Mackay, P. C. Williams, J. Livsey, J. P. Logue and J. H. Hendry, "Prediction of the benefits from dose-escalated hypofractionated intensity-modulated radiotherapy for prostate cancer," *Int.J.Radiat.Oncol.Biol.Phys.* 56 (1), 199-207 (2003).

57. C. M. Nutting, C. M. Corbishley, B. Sanchez-Nieto, V. P. Cosgrove, S. Webb and D. P. Dearnaley, "Potential improvements in the therapeutic ratio of prostate cancer irradiation: dose escalation of pathologically identified tumour nodules using intensity modulated radiotherapy," *Br.J.Radiol.* 75 (890), 151-161 (2002).

58. J. L. Bedford, L. Viviers, Z. Guzel, P. J. Childs, S. Webb and D. M. Tait, "A quantitative treatment planning study evaluating the potential of dose escalation in conformal radiotherapy of the oesophagus," *Radiother.Oncol.* 57 (2), 183-193 (2000).

59. S. H. Benedict, F. J. Bova, B. Clark, S. J. Goetsch, W. H. Hinson, D. D. Leavitt, D. J. Schlesinger and K. M. Yenice, "Anniversary Paper: the role of medical physicists in developing stereotactic radiosurgery," *Med.Phys.* 35 (9), 4262-4277 (2008).

60. H. Kano, A. Niranjan, S. Mongia, D. Kondziolka, J. C. Flickinger and L. D. Lunsford, "The role of stereotactic radiosurgery for intracranial hemangioblastomas," *Neurosurgery* 63 (3), 443-50; discussion 450-1 (2008).

61. H. Kano, A. Niranjan, D. Kondziolka, J. C. Flickinger and L. D. Lunsford, "Stereotactic radiosurgery for pituitary metastases," *Surg.Neurol.* , (2008).

62. F. Yamane, M. Takeshita, M. Izawa, M. Kagawa, K. Sato and K. Takakura, "Natural history of arteriovenous malformations: analysis of non-radically treated patients," *J.Clin.Neurosci.* 5 Suppl, 26-29 (1998).

63. S. Fukuoka, M. Takanashi, Y. Seo, K. Suematsu and J. Nakamura, "Radiosurgery for arteriovenous malformations with gamma-knife: a multivariate analysis of factors influencing the complete obliteration rate," *J.Clin.Neurosci.* 5 Suppl, 68-71 (1998).

64. Y. Kida, "Radiosurgery for intracranial dural arteriovenous fistulas--indications for limitations for gamma knife treatment," *Brain Nerve* 60 (8), 915-921 (2008).

65. H. Kikuchi, S. Miyamoto, I. Nagata, M. Ishikawa, W. Taki and I. Nakahara, "Surgical treatment of pericentral arteriovenous malformation," *J.Clin.Neurosci.* 5 Suppl, 87-90 (1998).

66. D. R. Buis, F. J. Lagerwaard, F. Barkhof, C. M. Dirven, G. J. Lycklama, O. W. Meijer, R. van den Berg, H. A. Langendijk, B. J. Slotman and W. P. Vandertop, "Stereotactic radiosurgery for brain AVMs: role of interobserver variation in target definition on digital subtraction angiography," *Int.J.Radiat.Oncol.Biol.Phys.* 62 (1), 246-252 (2005).

67. A. Jawahar, L. L. Jawahar, A. Nanda, C. D. Sharp, A. Warren, J. W. Elrod, M. Jennings, J. S. Alexander and A. Minagar, "Stereotactic radiosurgery using the Leksell Gamma Knife: current trends and future directives," *Front.Biosci.* 9, 932-938 (2004).

68. G. Grebe, M. Pfaender, M. Roll, L. Luedemann and R. E. Wurm, "Dynamic arc radiosurgery and radiotherapy: commissioning and verification of dose distributions," *Int.J.Radiat.Oncol.Biol.Phys.* 49 (5), 1451-1460 (2001).

69. B. G. Clark, J. L. Robar and A. M. Nichol, "Analysis of treatment parameters for conformal shaped field stereotactic irradiation: comparison with non-coplanar arcs," *Phys.Med.Biol.* 46 (12), 3089-3103 (2001).

70. S. H. Benedict, R. M. Cardinale, Q. Wu, R. D. Zwicker, W. C. Broaddus and R. Mohan, "Intensity-modulated stereotactic radiosurgery using dynamic micro-multileaf collimation," *Int.J.Radiat.Oncol.Biol.Phys.* 50 (3), 751-758 (2001).

71. M. Schwartz, P. O'Brien, P. Davey, C. Young, R. Willinsky and C. Catton, "Current status of radiosurgery for arteriovenous malformations," *Can.J.Neurol.Sci.* 18 (4), 499-502 (1991).

72. R. M. Cardinale, S. H. Benedict, Q. Wu, R. D. Zwicker, H. E. Gaballa and R. Mohan, "A comparison of three stereotactic radiotherapy techniques; ARCS vs. noncoplanar fixed fields vs. intensity modulation," *Int.J.Radiat.Oncol.Biol.Phys.* 42 (2), 431-436 (1998).

73. B. A. Kramer, D. E. Wazer, M. J. Engler, J. S. Tsai and M. N. Ling, "Dosimetric comparison of stereotactic radiosurgery to intensity modulated radiotherapy," *Radiat.Oncol.Investig.* 6 (1), 18-25 (1998).

74. C. Yu, G. Luxton, G. Jozsef, M. L. Apuzzo and Z. Petrovich, "Dosimetric comparison of three photon radiosurgery techniques for an elongated ellipsoid target," *Int.J.Radiat.Oncol.Biol.Phys.* 45 (3), 817-826 (1999).

75. T. D. Solberg, K. L. Boedeker, R. Fogg, M. T. Selch and A. A. DeSalles, "Dynamic arc radiosurgery field shaping: a comparison with static field conformal and noncoplanar circular arcs," *Int.J.Radiat.Oncol.Biol.Phys.* 49 (5), 1481-1491 (2001).

76. D. Huang, P. Xia, P. Akazawa, C. Akazawa, J. M. Quivey, L. J. Verhey, M. Kaplan and N. Lee, "Comparison of treatment plans using intensity-modulated radiotherapy and three-dimensional conformal radiotherapy for paranasal sinus carcinoma," *Int.J.Radiat.Oncol.Biol.Phys.* 56 (1), 158-168 (2003).

77. A. Mestrovic and B. G. Clark, "Geometric parameter analysis to predetermine optimal radiosurgery technique for the treatment of arteriovenous malformation," *Int.J.Radiat.Oncol.Biol.Phys.* 63 (3), 940-951 (2005).

78. G. Herman, J. Zheng and C. Bucholtz, "Shape-based Interpolation," *IEEE Computer Graphics & Applications* 12 (3), 69-79 (1992).

79. M. Younger, *A First Course in Linear Regression*, PWS Publishers, Boston (Massachusetts), 1985.

80. A. Afifi and S. Azen, *Statistical Analysis: A Computer Oriented Approach*, Academic Press, Inc, London (England), 1979.

81. A. Mestrovic, M. P. Milette, A. Nichol, B. G. Clark and K. Otto, "Direct aperture optimization for online adaptive radiation therapy," *Med.Phys.* 34 (5), 1631-1646 (2007).

82. A. Niemierko and M. Goitein, "Random sampling for evaluating treatment plans," *Med.Phys.* 17 (5), 753-762 (1990).

83. C. S. Chui, T. LoSasso and S. Spirou, "Dose calculation for photon beams with intensity modulation generated by dynamic jaw or multileaf collimations," *Med.Phys.* 21 (8), 1237-1244 (1994).

84. M. J. Ghilezan, D. A. Jaffray, J. H. Siewerdsen, M. Van Herk, A. Shetty, M. B. Sharpe, S. Zafar Jafri, F. A. Vicini, R. C. Matter, D. S. Brabbins and A. A. Martinez, "Prostate gland motion assessed with cine-magnetic resonance imaging (cine-MRI)," *Int.J.Radiat.Oncol.Biol.Phys.* 62 (2), 406-417 (2005).

85. A. M. Nichol, T. Rosewall, C. N. Catton, D. A. Jaffray and P. R. Warde, "Intra-prostatic fiducial markers and concurrent androgen deprivation," *Clin.Oncol.(R.Coll.Radiol.)* 17 (6), 465-468 (2005).

86. R. C. Zellars, P. L. Roberson, M. Strawderman, D. Zhang, H. M. Sandler, R. K. Ten Haken, D. Osher and P. W. McLaughlin, "Prostate position late in the course of external beam therapy: patterns and predictors," *Int.J.Radiat.Oncol.Biol.Phys.* 47 (3), 655-660 (2000).

87. G. Sanguineti, M. Marcenaro, P. Franzone, F. Foppiano and V. Vitale, "Neoadjuvant androgen deprivation and prostate gland shrinkage during conformal radiotherapy," *Radiother.Oncol.* 66 (2), 151-157 (2003).

88. A. M. Nichol, K. K. Brock, G. A. Lockwood, D. J. Moseley, T. Rosewall, P. R. Warde, C. N. Catton and D. A. Jaffray, "A magnetic resonance imaging study of prostate deformation relative to implanted gold fiducial markers," *Int.J.Radiat.Oncol.Biol.Phys.* 67 (1), 48-56 (2007).

89. R. T. O. G. 0415, "A Phase III Randomized Study of Hypofractionated 3D - CRT/IMRT Versus Conventionally Fractionated 3D - CRT/IMRT in Patients with Favorable-Risk Prostate Cancer," *(www.RTOG.org accessed on May 2008)* , (2008).

90. D. Yan, B. Xu, D. Lockman, K. Kota, D. S. Brabbins, J. Wong and A. A. Martinez, "The influence of interpatient and intrapatient rectum variation on external beam treatment of prostate cancer," *Int.J.Radiat.Oncol.Biol.Phys.* 51 (4), 1111-1119 (2001).

91. M. S. Hoogeman, M. van Herk, D. Yan, L. J. Boersma, P. C. Koper and J. V. Lebesque, "A model to simulate day-to-day variations in rectum shape," *Int.J.Radiat.Oncol.Biol.Phys.* 54 (2), 615-625 (2002).

92. Z. Jiang, M. A. Earl, G. W. Zhang, C. X. Yu and D. M. Shepard, "An examination of the number of required apertures for step-and-shoot IMRT," *Phys.Med.Biol.* 50 (23), 5653-5663 (2005).

93. A. Mestrovic, A. Nichol, B. G. Clark and K. Otto, "On-line adaptive radiation therapy based on intra-fractional digital tomosynthesis images," *Phys.Med.Biol.* , (submitted).

94. B. R. Paliwal, M. A. Ritter, T. R. McNutt, T. R. Mackie, B. R. Thomadsen, J. A. Purdy and T. J. Kinsella, "A solid water pelvic and prostate phantom for imaging, volume rendering, treatment planning, and dosimetry for an RTOG multi-institutional, 3-D dose escalation study. Radiation Therapy Oncology Group," *Int.J.Radiat.Oncol.Biol.Phys.* 42 (1), 205-211 (1998).

95. D. J. Godfrey, F. F. Yin, M. Oldham, S. Yoo and C. Willett, "Digital tomosynthesis with an on-board kilovoltage imaging device," *Int.J.Radiat.Oncol.Biol.Phys.* 65 (1), 8-15 (2006).

96. X. Liu and C. C. Shaw, "A-Si:H/CsI(Tl) flat-panel versus computed radiography for chest imaging applications: image quality metrics measurement," *Med.Phys.* 31 (1), 98-110 (2004).

97. S. J. Tu, C. C. Shaw and L. Chen, "Noise simulation in cone beam CT imaging with parallel computing," *Phys.Med.Biol.* 51 (5), 1283-1297 (2006).

98. L. A. Feldkamp, L. C. Davis and K. J. W., "Practical cone-beam algorithm," *J.Opt.Soc.Am.* 1 (6), 612-619 (1984).

99. G. Wang, T. H. Lin, P. Cheng and D. M. Shinozaki, "A general cone-beam reconstruction algorithm," *IEEE Trans.Med.Imaging* 12 (3), 486-496 (1993).

100. P. Abolmaesumi and M. R. Sirouspour, "An interacting multiple model probabilistic data association filter for cavity boundary extraction from ultrasound images," *IEEE Trans.Med.Imaging* 23 (6), 772-784 (2004).

101. S. Badiei, S. E. Salcudean, J. Varah and W. J. Morris, "Prostate segmentation in 2D ultrasound images using image warping and ellipse fitting," *Med.Image Comput.Comput.Assist.Interv.Int.Conf.Med.Image Comput.Comput.Assist.Interv.* 9 (Pt 2), 17-24 (2006).

102. J. J. Lagendijk, B. W. Raaymakers, A. J. Raaijmakers, J. Overweg, K. J. Brown, E. M. Kerkhof, R. W. van der Put, B. Hardemark, M. van Vulpen and U. A. van der Heide, "MRI/linac integration," *Radiother.Oncol.* 86 (1), 25-29 (2008).

103. C. Kirkby, T. Stanescu, S. Rathee, M. Carlone, B. Murray and B. G. Fallone, "Patient dosimetry for hybrid MRI-radiotherapy systems," *Med.Phys.* 35 (3), 1019-1027 (2008).

104. K. Otto, "Volumetric modulated arc therapy: IMRT in a single gantry arc," *Med.Phys.* 35 (1), 310-317 (2008).

105. D. J. Godfrey, L. Ren, H. Yan, Q. Wu, S. Yoo, M. Oldham and F. F. Yin, "Evaluation of three types of reference image data for external beam radiotherapy target localization using digital tomosynthesis (DTS)," *Med.Phys.* 34 (8), 3374-3384 (2007).

106. L. Ren, D. J. Godfrey, H. Yan, Q. J. Wu and F. F. Yin, "Automatic registration between reference and on-board digital tomosynthesis images for positioning verification," *Med.Phys.* 35 (2), 664-672 (2008).

107. W. Ansbacher, "Three-dimensional portal image-based dose reconstruction in a virtual phantom for rapid evaluation of IMRT plans," *Med.Phys.* 33 (9), 3369-3382 (2006).

# Appendix A

# Multiple Linear Regression Correlation Results

| Dosimetric Parameters | Significant Predictors ($p<0.05$) for different techniques | | | |
|---|---|---|---|---|
| | **Circular Arcs** | **Conformal** | **Dynamic Arcs** | **IMRS** |
| MDPD | none | $b_3 = -0.13 \pm 0.05$<br>$b_5 = .010 \pm .002$<br>$a = 1.42$ RMSE = .013 | $b_1 = .003 \pm .001$<br>$b_8 = .011 \pm .003$<br>$a = 1.24$ RMSE = .010 | none |
| PITV | none | $b_1 = -.022 \pm .009$<br>$b_4 = .027 \pm .004$<br>$a = 1.47$ RMSE = .150 | $b_4 = .030 \pm .002$<br>$b_5 = -.070 \pm .016$<br>$a = 1.33$ RMSE = .072 | $b_1 = -.039 \pm .008$<br>$b_4 = .017 \pm .003$<br>$a = 2.04$ RMSE = .137 |
| Normal Tissue $D_{max}$ | none | $b_9 = 2.84 \pm 0.60$<br>$a = 12.9$ RMSE = 2.13 | $b_9 = 3.11 \pm 0.60$<br>$a = 11.7$ RMSE = 2.13 | $b_9 = 3.02 \pm 0.54$<br>$a = 10.7$ RMSE = 1.92 |
| Normal Tissue $V_{24}$ | none | none | none | none |
| Normal Tissue $V_{18}$ | $b_1 = 1.63 \pm 0.17$<br>$b_6 = -2.52 \pm 1.05$<br>$a = 2.68$ RMSE = 2.91 | $b_2 = -7.90 \pm 1.87$<br>$b_5 = 4.00 \pm 0.37$<br>$b_9 = 1.92 \pm 0.47$<br>$a = 4.15$ RMSE = 1.64 | $b_1 = 0.92 \pm 0.09$<br>$b_4 = 0.10 \pm 0.04$<br>$b_8 = 2.57 \pm 0.52$<br>$a = -4.40$ RMSE = 1.60 | $b_1 = 1.14 \pm 0.17$<br>$b_8 = 3.15 \pm 0.94$<br>$a = -2.37$ RMSE = 2.92 |
| Normal Tissue $V_{12}$ | $b_1 = 3.21 \pm 0.25$<br>$a = 3.67$ RMSE = 4.30 | $b_1 = 1.34 \pm 0.18$<br>$b_5 = 4.14 \pm 0.55$<br>$b_9 = 2.42 \pm 0.62$<br>$a = -7.67$ RMSE = 2.18 | $b_1 = 1.60 \pm 0.19$<br>$b_5 = 2.25 \pm 0.58$<br>$b_8 = 2.74 \pm 0.74$<br>$a = -0.50$ RMSE = 2.29 | $b_1 = 1.85 \pm 0.24$<br>$b_5 = 3.50 \pm 0.73$<br>$b_9 = 3.22 \pm 0.83$<br>$a = -9.46$ RMSE = 2.90 |
| Normal Tissue $V_6$ | $b_1 = 9.33 \pm 0.67$<br>$a = 9.45$ RMSE = 11.4 | $b_1 = 9.42 \pm 1.69$<br>$a = 22.7$ RMSE = 28.9 | $b_1 = 5.68 \pm 0.58$<br>$b_5 = 6.06 \pm 1.77$<br>$b_9 = 4.87 \pm 2.01$<br>$a = -11.4$ RMSE = 7.0 | $b_1 = 8.14 \pm 1.72$<br>$b_5 = 11.63 \pm 5.25$<br>$a = 16.5$ RMSE = 21.0 |
| Normal Tissue $V_3$ | $b_1 = 25.2 \pm 2.0$<br>$b_7 = -1.84 \pm 0.51$<br>$b_8 = -32.0 \pm 11.1$<br>$a = 120$ RMSE = 31.1 | $b_1 = 25.0 \pm 3.6$<br>$a = 146$ RMSE = 62.1 | $b_1 = 21.1 \pm 1.4$<br>$a = 36.9$ RMSE = 23.5 | $b_1 = 30.2 \pm 3.7$<br>$a = 136$ RMSE = 62.9 |

*Table A.1: Multiple linear regression model correlation results for* ***the target and normal tissue*** *dosimetric parameters. All non-zero regression coefficients ($b_1$, $b_2$, $b_3$,...$b_k$), labelled according to the first column of Table 3.1, and the regression constant (a) are listed for the predictors that*

*showed statistically significant correlation ($p<0.05$) with different dosimetric parameters. Abbreviations: MDPD = ratio of the maximum dose to the prescription dose, PITV = ratio of the volume within the prescription isodose surface to the target volume, $D_{max}$ = the maximum dose, $V_n$ = the volume treated to n Gy, RMSE = root mean squared errors.*

| Dosimetric Parameters | Significant Predictors ($p<0.05$) for different techniques | | | |
|---|---|---|---|---|
| | **Circular Arcs** | **Conformal** | **Dynamic Arcs** | **IMRS** |
| Critical Structure $D_{max}$ | $b_6 = 9.00 \pm 1.81$<br>$b_8 = -5.57 \pm 0.43$<br>$a = 18.8$ RMSE = 2.69 | $b_6 = 6.02 \pm 1.52$<br>$b_8 = -4.51 \pm 0.36$<br>$a = 16.9$ RMSE = 2.77 | $b_6 = 6.43 \pm 1.45$<br>$b_8 = -4.89 \pm 0.35$<br>$a = 16.4$ RMSE =2.56 | $b_6 = 5.02 \pm 1.34$<br>$b_8 = -4.23 \pm 0.32$<br>$a = 16.6$ RMSE = 2.23 |
| Critical Structure $V_{24}$ | none | $b_6 = 0.63 \pm 0.05$<br>$a = .001$ RMSE = 0.18 | $b_6 = 0.66 \pm 0.05$<br>$a = -.005$ RMSE = 0.18 | none |
| Critical Structure $V_{18}$ | $b_6 = 2.98 \pm 0.24$<br>$b_7 = -0.12 \pm 0.02$<br>$b_9 = -0.14 \pm 0.04$<br>$a = 0.76$ RMSE = 0.49 | $b_6 = 2.49 \pm 0.22$<br>$b_7 = -.087 \pm .023$<br>$b_9 = -.099 \pm .035$<br>$a = 0.52$ RMSE = 0.47 | $b_6 = 2.37 \pm 0.20$<br>$b_7 = -.080 \pm .020$<br>$b_9 = -.089 \pm .032$<br>$a = 0.47$ RMSE = 0.42 | $b_6 = 2.71 \pm 0.24$<br>$b_7 = -0.11 \pm 0.02$<br>$b_9 = -0.10 \pm 0.04$<br>$a = 0.54$ RMSE = 0.51 |
| Critical Structure $V_{12}$ | $b_6 = 4.47 \pm 0.45$<br>$b_7 = -0.22 \pm 0.05$<br>$b_9 = -0.37 \pm 0.07$<br>$a = 1.99$ RMSE = 0.73 | $b_6 = 3.83 \pm 0.45$<br>$b_7 = -0.17 \pm 0.05$<br>$b_9 = -0.31 \pm 0.07$<br>$a = 1.62$ RMSE = 0.74 | $b_6 = 3.66 \pm 0.43$<br>$b_7 = -0.16 \pm 0.04$<br>$b_9 = -0.28 \pm 0.07$<br>$a = 1.47$ RMSE = 0.71 | $b_6 = 4.57 \pm 0.51$<br>$b_7 = -0.22 \pm 0.05$<br>$b_9 = -0.34 \pm 0.08$<br>$a = 1.82$ RMSE = 0.86 |
| Critical Structure $V_6$ | $b_5 = .009 \pm .004$<br>$b_6 = 6.89 \pm 1.10$<br>$b_7 = -0.45 \pm 0.11$<br>$b_9 = -1.14 \pm 0.17$<br>$a = 5.54$ RMSE = 1.20 | $b_1 = .063 \pm .017$<br>$b_6 = 5.54 \pm 1.13$<br>$b_7 = -0.35 \pm 0.12$<br>$b_9 = -1.16 \pm 0.18$<br>$a = 4.93$ RMSE = 1.32 | $b_1 = .039 \pm .016$<br>$b_6 = 6.30 \pm 1.06$<br>$b_7 = -0.39 \pm 0.11$<br>$b_9 = -1.02 \pm 0.17$<br>$a = 4.47$ RMSE = 1.18 | $b_1 = .068 \pm .018$<br>$b_6 = 6.64 \pm 1.19$<br>$b_7 = -0.44 \pm 0.12$<br>$b_9 = -1.31 \pm 0.19$<br>$a = 5.64$ RMSE = 1.43 |
| Critical Structure $V_3$ | $b_1 = 0.17 \pm 0.03$<br>$b_6 = 7.57 \pm 2.03$<br>$b_7 = -0.61 \pm 0.21$<br>$b_9 = -2.35 \pm 0.32$<br>$a = 10.2$ RMSE = 2.19 | $b_1 = 0.31 \pm 0.03$<br>$b_9 = -2.68 \pm 0.32$<br>$a = 9.60$ RMSE = 2.46 | $b_1 = 0.12 \pm 0.03$<br>$b_6 = 8.49 \pm 1.97$<br>$b_7 = -0.72 \pm 0.20$<br>$b_9 = -2.53 \pm 0.31$<br>$a = 11.3$ RMSE = 2.06 | $b_1 = 0.32 \pm 0.03$<br>$b_9 = -2.74 \pm 0.33$<br>$a = 10.0$ RMSE = 2.59 |

*Table A.2: Multiple linear regression model correlation results for* ***the critical structure*** *dosimetric parameters. All non-zero regression coefficients ($b_1$, $b_2$, $b_3$,…$b_k$), labelled according to the third column of Table 3.1, and the regression constant (a) are listed for the predictors that showed statistically significant correlation ($p<0.05$) with different dosimetric parameters. Abbreviations: $D_{max}$ = the maximum dose, $V_n$ = the volume treated to n Gy, RMSE = root mean squared errors.*

# Appendix B

# Examples of Final Aperture Shapes and Weights

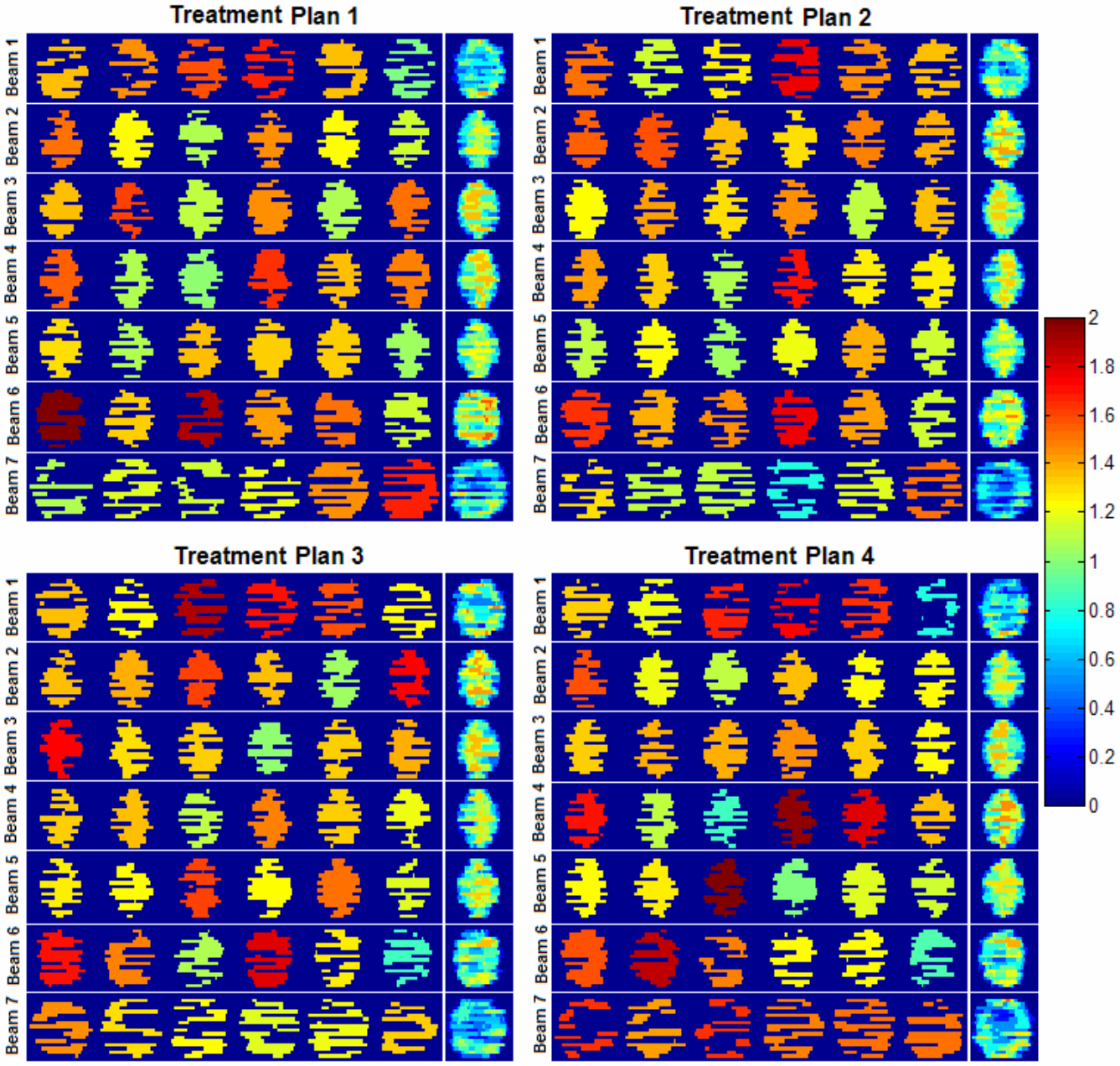


*Figure B.1: The final aperture shapes and weights of four different treatment plans for the anatomical deformation of 1.00 cm. Even though the final objective function values for different plans are all close to each other (0.484, 0.406, 0.464 and 0.466, respectively) the aperture shapes and weights are significantly different between the four plans. This provides evidence that multiple local minima with similar cost values exist within the search space.*