# Atlas Selection for Automated Segmentation of Pelvic CT for Prostate Radiotherapy

# Atlas Selection for Automated Segmentation of Pelvic CT for Prostate Radiotherapy

By ABRAR A. MALLAWI, B.Sc.

Medical Physics and Applied Radiation Sciences
McMaster University
Hamilton, Canada
April 2015

A Thesis Submitted to the School of Graduate Studies in Partial Fulfillment of the Requirements for the Degree Master of Science

McMaster University MASTER OF SCIENCE (2015) Hamilton, Ontario (Medical Physics and Applied Radiation Sciences)

TITLE: Atlas Selection for Automated Segmentation of Pelvic CT for Prostate Radiotherapy
AUTHOR: Abrar A. Mallawi, B.Sc. SUPERVISOR: Dr. Marcin Wierzbicki
NUMBER OF PAGES: iii, 74

# Abstract

Radiotherapy has become a standard modality for treating prostate cancer. Typically, intensity modulated radiotherapy (IMRT) is employed. Accurate delineation is important to ensure that the clinical target volume (prostate) is sufficiently irradiated and that the organs at risk (OARs) are appropriately spared. A recent technological development in radiotherapy treatment planning is the employment of atlas-based segmentation to automate target volume and OAR delineation. Atlas based-segmentation utilizes the spatial relationship between a pre-contoured atlas subject and a new patient image to derive the segmentation result. The typical approach has three steps: many atlas subject images are globally registered to the target image, an atlas subject image that is the most similar to the target is selected, and the chosen atlas image and contours are aligned with the target image space using deformable registration. The purpose of this work was to design an atlas selection strategy and evaluate its impact on the final atlas-based segmentation outcome. Segmentation accuracy was mainly quantified using the Dice Similarity Coefficient (DSC), which was used to score the overlap between automatic and manual contours on a 0 to 1 scale.

An alternative atlas selection approach was proposed that identified the most similar atlas subject based on several anatomical measurements that were chosen to indicate the overall prostate and body shape. A brute force procedure was first performed for a training dataset of 20 patients using image registration to pair subjects with similar contours based on DSC. For the identified best matches, anatomical measurements were compared. An atlas selection procedure was designed; relying on the computation of a similarity score defined as a weighted sum of

differences between the target and atlas subject anatomical measurements. Finally, an optimization procedure was performed to obtain the weights that gave the highest DSC between automatic and manual contours for the training set. The mean DSCs obtained using brute force were 0.78±0.07 and 0.90±0.02 for the prostate and either femoral head. The proposed atlas selection method achieved 0.72±0.11 and 0.87±0.03 for the prostate and either femoral head. Clearly, the algorithm was able to identify the best matching atlas subject for any target subject in the training set of data.

The key point of this work was to also validate the atlas selection strategy. Thus, the optimized atlas selection procedure was tested on images of 10 additional subjects. Again, the algorithm's ability to predict the most similar atlas subject was excellent. A brute force search for the set of 10 images achieved mean DSCs of 0.76±0.03 and 0.88±0.03 for the prostate and either femoral head. The proposed method yielded DSCs of 0.64±0.09 and 0.86±0.04 for the prostate and either femoral head. The difference in mean DSCs between the proposed method and the brute force was statically significant ($p < 0.05$). Overall, the brute force results demonstrated that atlas-based segmentation can reproduce a similar level of accuracy as manual recontouring (Granberg *et a*l., 2011). More importantly, the same level of accuracy was achieved with the proposed atlas selection method as with more computationally intensive techniques. These results indicate that atlas-based segmentation is a promising technique for prostate radiation therapy.

***This thesis is dedicated to***
***My parents...***
***My sisters...***
***My brother...***
*For their endless love, support and encouragement*

# Acknowledgments

It is my privilege to release my gratitude to several persons who helped me directly and indirectly during this project.

I would like to start by thanking my supervisor, **Dr. Marcin Wierzbicki** for his invaluable support throughout the master journey. His guidance, patience, encouragement, and vast knowledge at all the stages of this project made it a wonderful experience. My eternal gratitude to **Dr. Wierzbicki;** this thesis could not be completed without his efforts.

I also sincerely thank **Dr. Tom Farrell** and **Dr. Kevin Diamond** for their valuable comments and suggestions.

I would like to express my gratitude to my friends for their comments, moral support.

I would like also to thank my family; without their love, support and motivation, none of this would have been done. My wonderful little sister for being supportive, my grandmother wholehearted thanks as well.

Thank you to the National Guard and Health Affairs (NGHA) for the financial support I received for this project.

# Contents

**Chapter 2: Materials and Methods**



**Chapter 3: Results and Discussion**



**Chapter 4:**

# List of Tables

# List of Figures

# List of Abbreviation

| | |
|---|---|
| **linac** | **Linear Accelerator** |
| **CT** | **Computed Tomography** |
| **EBRT** | **External Beam Radiation Therapy** |
| **3DCRT** | **Three Dimensional Conformal Radiotherapy** |
| **IMRT** | **Intensity Modulated Radiation Therapy** |
| **JCC** | **Juravinski Cancer Center** |
| **VMAT** | **Volumetric Modulated Radiation Arc Therapy** |
| **IGRT** | **Image Guided Radiotherapy** |
| **GTV** | **Gross Target Volume** |
| **CTV** | **Clinical Target Volume** |
| **PTV** | **Planning Target Volume** |
| **DVH** | **Dose Volume Histogram** |
| **GI** | **Gastrointestinal** |
| **HDR** | **High Dose Rate** |
| **HIFU** | **High intensity focused ultrasound** |
| **LDR** | **Low Dose Rate** |
| **SA** | **Simulated Annealing** |
| **MI** | **Mutual Information** |
| **MLC** | **Multi Leaf Collimator** |

| | |
|---|---|
| **OAR** | **Organ at Risk** |
| **ROI** | **Region of Interest** |
| **TPS** | **Treatment Planning System** |
| **DOF** | **Degree of Freedom** |
| **ABS** | **Atlas Based Segmentation** |
| **AP** | **Anterior-Posterior** |
| **LR** | **Left-Right (lateral)** |
| **RFHD** | **Right Femoral Head Diameter** |
| **LFHD** | **Light Femoral Head Diameter** |
| **PL** | **Prostate Length** |
| **RF** | **Right Femoral** |
| **LF** | **Left Femoral** |
| **RW** | **Rectal Wall** |
| **BW** | **Bladder Wall** |
| **DSC** | **Dice Similarity Coefficient** |
| **ASM** | **Average Shape Model** |
| **PCA** | **Principle Component Analysis** |

Chapter 1

# Introduction

## 1.1. Prostate Cancer

The human body is made up of millions of cells that organize forming tissues and organs. Normally, genes inside the cell are responsible for cell growth, division, and cell death. When normal cells are damaged, they enter the apoptosis process (cell death) and are often replaced by new cells in a programmed and controlled way. When this process fails, new cells are produced without control leading to the formation of a tumor. Cell growth within the prostate gland can be either benign (non-cancerous) or malignant (cancerous), with almost all men experiencing prostate enlargement by the age of 70 (Canadian Cancer Society, 2014). Benign prostate enlargement is common but is not life threatening. The growth of the prostate compresses the urethra preventing urine from flowing normally. Malignant cell growth, on the other hand, is life threatening since cancer cells have the potential to spread to, and damage, different parts of the body. Men of different race, family history, diet, and lifestyle have a varying risk of developing prostate cancer. For example, African men have about 60% higher rate of prostate cancer than

Caucasian men, while Asian men have a reduced rate of developing prostate cancer (Canadian Cancer Society, 2014).

Prostate cancer is one of the most frequently diagnosed cancers in males worldwide (Globocan, 2012). In Canada, prostate cancer is the most commonly diagnosed cancer after skin cancer with 23,600 new cases estimated in 2014 (Canadian Cancer Society, 2014). Currently, prostate cancer is one of the leading causes of cancer-related death among Canadians with 4,000 deaths in 2014, representing 10% of all cancer deaths (Canadian Cancer Society, 2014).

## 1.2. Regional Anatomy

The prostate is a part of the male reproductive system. It is a fibromuscular gland, which surrounds the prostatic urethra. The adult prostate is about the size of a walnut and weights 20-25 g. As shown Figure 1.1, the prostate gland is located between the pubic symphysis and the anterior rectal wall and it is near the bladder neck and seminal vesicles (Snell, 2012). The prostate gland is composed of a mixture of smooth muscle and glandular tissue, with openings into the urethra. The prostate itself is divided into the anterior, middle, posterior, right, and left lobes. The function of the prostate gland is to produce a milky fluid composed of citric acid and acid phosphatase. This fluid is added to semen at the time of ejaculation to neutralize the acidity in the vagina (Snell, 2012).

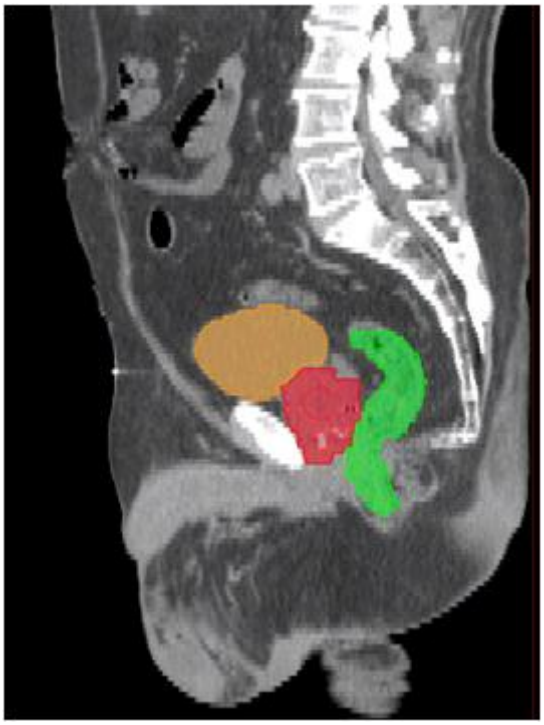


**Figure 1.1:** Sagittal section of a contoured CT image. The red area is the prostate gland, orange corresponds to the bladder, and green represents the rectum.

## 1.3. Prostate Cancer Radiation Therapy

Treatment options for prostate cancer patients depend on the stage of the disease and factors such as age, health condition, and personal preference. The available options include radical prostatectomy usually with pelvic lymph node dissection (lymphadenectomy) (Boxer *et al*., 1977), hormonal therapy (Galbraith and Duchesne, 1997), chemotherapy (Crawford and Flaig, 2012), external-beam radiation therapy (EBRT) (Abdel-Wahab *et al*., 2012), and high dose rate (HDR) and low dose rate (LDR) brachytherapy (Sylvester *et al*., 2011). Moreover, various treatment modalities are commonly combined with hormonal therapy. In addition, new types of treatments are being tested in clinical trials including proton beam radiation therapy, high intensity focused ultrasound (HIFU), and ultrasound–guided cryosurgery (National Cancer Institute, 2014). In this work, the treatment option of interest is EBRT.

A typical dose prescription for early stage prostate cancer is 78 Gy delivered in 2 Gy daily fractions. Furthermore, there are indications that using a higher dose can lead to better treatment outcomes (Zelefsky *et al*., 1996, Vicini *et al*., 2001). Dose escalation requires the use of highly advanced treatment modalities such as intensity-modulated radiotherapy (IMRT) or volume-modulated arc-therapy (VMAT) where dose to the rectum, bladder, and femoral heads may be limited despite increased dose to the tumor (De Meerleer et al., 2004). These technologies deliver dose distributions with steep dose gradients so it becomes increasingly important to define the tumor and surrounding organs at risk (OAR) for successful targeting.

### 1.3.1 External Photon Beam Radiotherapy

EBRT is typically delivered using a linear accelerator (linac) where the radiation source is mounted on a gantry that rotates about a fixed point in space. The head of the linac is fitted with a multi-leaf collimator (MLC) that allows beam shaping via independent motion of many tungsten leaves. In three-dimensional conformal radiotherapy (3DCRT), multiple fields from different gantry angles around the patient are used to concentrate the radiation dose in the tumor and spare the surrounding normal tissue. The MLC is used to define the shape of each treatment field to fit the projection of the target. Currently, 3DCRT is considered to be a conventional technique. IMRT is an advanced modality in which each field is divided into many "beamlets", each carrying a different intensity of radiation. This is achieved by superimposing many MLC shapes that are often smaller than the target to deliver various beam intensities. One variant of IMRT is step-and-shoot, where the radiation beam is off during gantry rotation and when the

MLC is moving to shape the field. The beam is turned on once the field shaping is complete. IMRT offers many degrees of freedom to alter the dose distribution, improving the ability to conform the dose to the target volume. Figure 1.2 shows an example MLC configuration used for prostate IMRT, while Figure 1.3 shows a typical seven-beam arrangement.

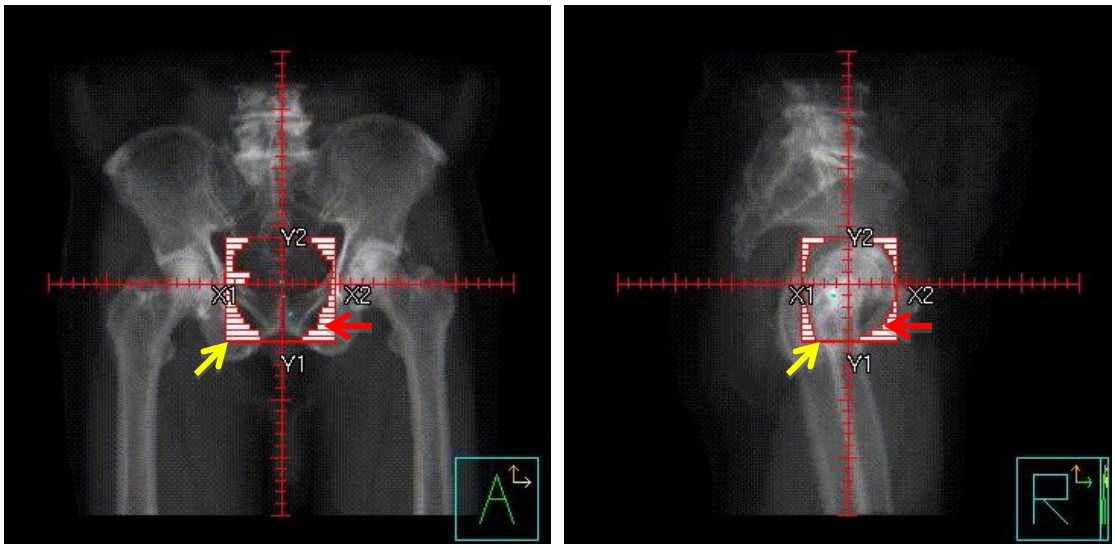



**Figure 1.2:** MLC configuration of anterior (left) and right anterior oblique (right) treatment fields. The red arrows represent MLC leafs and yellow arrows represent the jaws.

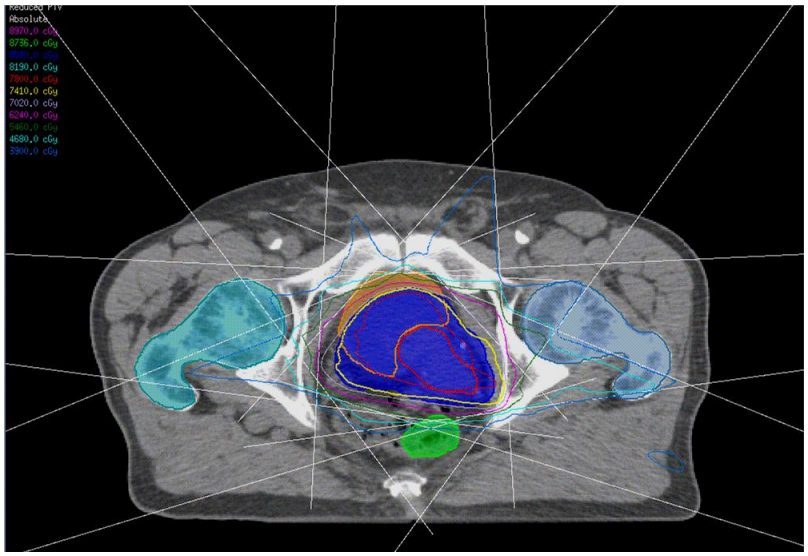



**Figure 1.3**: Axial view of a CT image for a prostate cancer case showing seven IMRT beams and the resulting dose distribution. The blue area is PTV that includes the prostate gland with additional geometric margin, orange is the bladder, green is the rectum, and turquoise and light blue are the femoral heads. The red, yellow, purple isodose lines represent the 78, 74 and 70 Gy respectively. The legend in the top left corner represents different isodose lines.

It has been shown that IMRT reduces gastrointestinal morbidity and the chance of hip fracture for prostate patients compared with 3DCRT due to reductions in bowel and femoral head doses (Sheets *et al*., 2012). A further development of IMRT is VMAT, where radiation is delivered as the gantry travels in an arc around the patient. In VMAT, a continuous beam is shaped by a dynamic MLC with simultaneous gantry rotation and dose rate control. This increases the number of beam angles available compared with IMRT, potentially leading to higher dose conformality, improved normal tissue sparing and faster delivery.

## 1.3.2 Treatment Planning

Prior to radiotherapy a treatment planning system (TPS) is typically used to determine treatment parameters such as beam direction, intensity and shape that will ensure the prescribed dose is delivered to the target with minimal dose to normal tissue. A typical treatment planning protocol for prostate EBRT is shown in Figure 1.4. The process begins with the acquisition of a CT image of the patient. The position must be comfortable so the patient is able to maintain it throughout the entire treatment course. The next step in the planning process is to contour the target and normal tissue on the planning CT. This is typically performed by a radiation oncologist employing the nomenclature outlined by the International Commission on Radiation Units and Measurements (ICRU) (ICRU Report 62, 1999). Supplementary images obtained using magnetic resonance imaging (MRI), positron emission tomography (PET), or ultrasound (US) may also be used to improve tissue contrast, thereby improving the accuracy of the target and normal tissue contours.

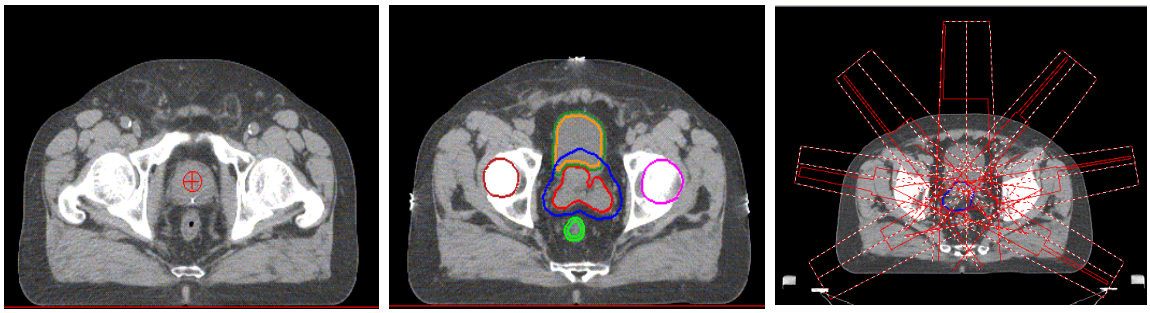


**Figure 1.4:** A typical treatment planning protocol for prostate cancer EBRT: 1. CT simulation (left), 2. Contouring of the target and surrounding healthy organs (center), and 3. Selection of beam parameters to deliver the prescribed dose (right).

The clinical target volume (CTV) comprises the gross tumor and areas at risk of subclinical involvement. In early stage prostate cancer the CTV includes the prostate gland and the base of the seminal vesicles. The planning target volume (PTV) consists of the prostate with

additional geometric margin. Targeting the PTV during treatment planning ensures that the prostate receives the prescribed dose despite uncertainties such as daily setup errors and patient motion due to organ filling. In addition to the prostate and PTV, all the healthy organs around the tumor volume are also outlined. For prostate cancer, the contoured OARs include the rectum (from the ischium to anterior flexion of sigmoid), entire bladder, and the right and left femoral heads (to the lesser trochanter inferiorly).

TPS offer various treatment planning alternatives for EBRT. For 3DCRT, a forward planning approach is used, while a computer based, inverse planning approach is necessary for IMRT/VMAT. Forward planning is a manual process that involves several steps. First, all beams parameters including the type of radiation and its energy along with geometrical parameters such as beam number, angles of incidence, and the MLC are defined. Second, the dose distribution is computed by the computer. Finally, the dose distribution is reviewed by the dosimetrist to confirm that the target volume is sufficiently covered and the OARs are sufficiently spared. If this is not achieved, various beam parameters may be modified until the goals of the treatment are achieved.

Manual dose optimization as used in 3DCRT planning is not possible in IMRT due to the complexity of the treatment. Therefore, a computer-based algorithm is used in which the dose to the OARs is limited though a number of user-defined objectives and constraints. For example, the dosimetrist may ask for “at least 95% of the PTV to be covered with dose greater than 78 Gy” while allowing “no more than 30% of the rectum to receive a dose of 40 Gy or higher”. The optimization algorithm then adjusts beam fluences to best meet the user-defined objectives. If

the result is not satisfactory, the planner will re-adjust the objectives appropriately and the optimization is repeated. At all stages, the treatment plan is reviewed using the dose volume histogram (DVH), which shows the fraction of targets or organs vs. the radiation dose. Once the plan is approved, the patient is carefully positioned on the linac treatment couch daily and the treatment is delivered as planned over several weeks.

### 1.3.3 Image-Guided Radiation Therapy (IGRT)

Radiotherapy is planned assuming that the prostate remains within the PTV despite organ motion or setup error. A large change from the planning geometry could increase the risk of recurrence due to reduced dose delivered to the prostate or the risk of complication due to over irradiation of OARs. One way to address daily uncertainty is to increase the margin between the prostate and PTV. However a larger margin increases the amount of dose to normal tissue. A small PTV margin is desired but requires the treatment to be corrected daily for anatomical variation and patient setup at each fraction. This requires IGRT, a process where patient images are acquired just before or during radiotherapy and the treatment is corrected online to ensure the planned dose is delivered.

Prostate IGRT typically involves implanting three or more gold fiducials into the prostate before treatment planning. These are clearly visible on the planning CT and act as surrogates for the position of the prostate at planning. Just before treatment, an orthogonal set of images of the patient is acquired employing MV photons. The treatment couch is then translated such that the fiducial positions at treatment align with their positions identified in planning. The treatment plan is delivered and the process is repeated for each fraction.

## 1.4 Image Registration

Medical image registration is the process of aligning homologous anatomy in two images by finding an optimal transformation. One image is defined as a fixed image $F$ (target) and the other is a moving image $M$ (source). An automated image registration algorithm comprises a transformation model that allows modifications of the source image to align it with the target, an interpolator to determine the intensity values at certain points when $M$ is moved, a similarity metric to quantify image alignment, and an optimization algorithm to maximize image similarity. When registering two images the objective is to find the spatial transformation that aligns the two images as measured by the similarity measure. The registration workflow is shown in Figure 1.5.

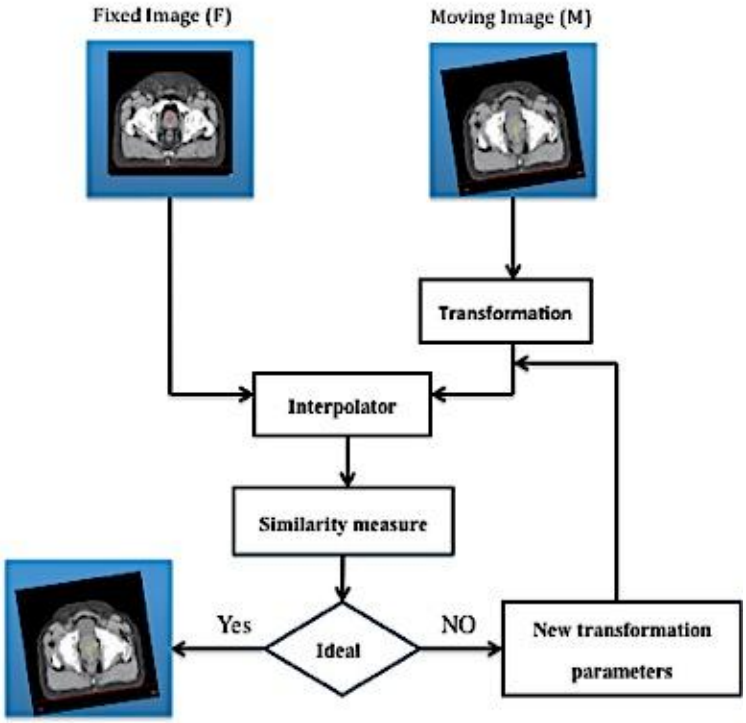



**Figure 1.5:** Image registration workflow.

## 1.4.1 Spatial Transformations

The transformation represents a spatial mapping from the source image to the target image space as shown in Figure 1.6.

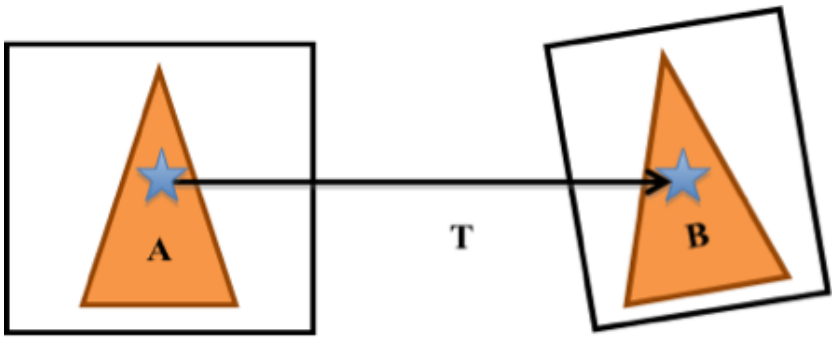



**Figure1.6**: Image registration determines the transformation that maps point A in one image to the homologous point B in the second image.

There are many ways to model transformations. Global methods affect the entire image simultaneously while local methods allow different areas of the image to undergo independent motion. One global transformation of interest is the 3D rigid transformation characterized by six degrees of freedom (DOF). In Cartesian coordinates there are three possible rotations about the x, y, and z axes ($\theta_x, \theta_y, \theta_z$) and three possible translations ($t_x, t_y, t_z$). Typically, this is written using a $4 \times 4$ matrix as follows:

$$TR = \begin{bmatrix} & R & & t_x \\ & & & t_y \\ & & & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad (1.1)$$

where $R = R_x R_y R_Z =$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & cos\theta_x & -sin\theta_x \\ 0 & sin\theta_x & cos\theta_x \end{bmatrix} \begin{bmatrix} cos\theta_y & 0 & sin\theta_y \\ 0 & 1 & 0 \\ -sin\theta_y & 0 & cos\theta_y \end{bmatrix} \begin{bmatrix} cos\theta_z & -sin\theta_z & 0 \\ sin\theta_z & cos\theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (1.2)$$

The process of finding a point in the target image $(x', y', z')$ that corresponds to a point in the source image $(x, y, z)$ using a rigid transformation may be written as follows:

$$(x', y', z') = TR(x, y, z) = TR \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}, \quad (1.3)$$

A slightly more complicated transformation model also involves scaling factors for the three axes. This gives a nine DOF transformation that may be written as follows:

$$TRS = \begin{bmatrix} & & & t_x \\ & RS & & t_y \\ & & & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (1.4)$$

$$RS = R \cdot \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & s_z \end{bmatrix}, \quad (1.5)$$

where R is the rotation matrix given in equation (1.3) while $s_x$ , $s_y$ , $s_z$ are the scaling factors for the three axes. When s = 1 there is no scaling, $|s| < 1$ is a compression, and $|s| > 1$ is an expansion. In general, an affine transformation comprises translation, rotation, scaling, and shearing for a total of 12 DOF in 3D. The general affine transform can be denoted as follows:

$$TA = \begin{bmatrix} & RSH & & \begin{matrix} t_x \\ t_y \\ t_z \end{matrix} \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (1.6)$$

$$RSH = R \cdot S \cdot \begin{bmatrix} 1 & h_x & h_y \\ 0 & 1 & h_z \\ 0 & 0 & 0 \end{bmatrix}, \quad (1.7)$$

Where R and S are the rotation and scaling matrices, respectively, and $h_x$, $h_{y,}$ $h_z$ are the shear factors along the $x, y$ and $z$ axes.

## 1.4.2 Image Similarity

Similarity metrics quantify the alignment of images during registration. Various types of metrics have been described but intensity-based computations are the most convenient from the perspective of automation. Mutual information (MI) is an example of an intensity-based similarity metric that is of special interest in this thesis. Mutual information measures how well one image explains the other; it is the amount of information in one image that allows one to describe the other image (Hill *et al*., 2001). It is maximized at the optimal registration and minimized for unrelated source and target images. MI is defined as follows:

$$MI\ (source, target) = H\ (source) + H\ (target) - H\ (source, target), \quad (1.8)$$

where H (source) and H (target) are the marginal entropies of the source and target images and H(*source*, *target*) is the joint entropy. A common method to compute these quantities is Shannon-Wiener entropy (Shannon, 1948):

$$H(source) = -\sum_a p(a) log p(a), \quad (1.9)$$

$$H(target) = -\sum_b p(b) log p(b), \quad (1.10)$$

$$H(source, target) = -\sum_a \sum_b p(a,b) \log p(a,b), \quad (1.11)$$

where $p\ (a)$ is the probability that intensity $a$ occurs in the source image, $p\ (b)$ is the probability that intensity $b$ occurs in the target image, and $p\ (a, b)$ is the probability that $a$ occurs in the source at the same location as $b$ occurs in the target. Typically, the MI metric has been used for inter-modality registration. For example, Studholme *et al*. (1996, 1997) used MI to register MR with CT and PET images. MI is also useful for intra-modality registration, for example, to align a pair of CT images (Wierzbicki *et al.*, 2004).

### 1.4.3 Optimization

In image registration, optimization is the selection of transformation parameters that maximizes the similarity between the transformed source and target images. Two common algorithms for function optimization are the downhill simplex approach (Nelder and Mead, 1965) and, of most interest in this thesis, simulated annealing (Metropolis *et al*., 1953).

#### 1.4.3.1 Downhill Simplex

The Nelder and Mead multidimensional unconstrained method for function optimization does not impose limits on parameter values. The algorithm begins by creating a simplex in the transformation parameter space. This is a polytope consisting of n + 1 vertices where n is the number of DOFs associated with the transformation model being solved. For example, a simplex is a triangle or a tetrahedron in two-dimensional or three-dimensional parameter space, respectively. The optimization commences by evaluating the similarity metric at each of the vertices and depending on these values, the simplex undergoes reflection, expansion, or contraction. The process continues until the change in similarity metric is below a certain tolerance.

#### 1.4.3.2 Simulated Annealing

The name and concept of simulated annealing (SA) originate from annealing in metallurgy where a material is heated and then cooled in a controlled manner to increase the size of crystals and reduce imperfections. Simulated annealing optimization is a random search method meaning that the parameter space is sampled randomly at each iteration. To control the convergence of the algorithm, the user specifies a cooling schedule described by an initial

temperature, a function for lowering temperature with increasing iterations, and the final temperature (stopping criterion).

Consider the example annealing run shown in Figure 1.7. An initial guess starts the process at point 1. The first iteration places the point of interest at point 2 which is selected purely randomly since nothing is known about the problem landscape. Furthermore, initial temperature is high meaning large moves in parameter space are allowed. Point 2 provides a higher function value than point 1 so it is automatically accepted as a potential solution. In the second iteration, the amount energy available for the entire system is less due to cooling. Thus, the move from point 2 to 3 is shorter than from 1 to 2. The function value at point 3 happens to be lower than 2 representing a less optimal set of parameters. At this point the algorithm allows selection of parameter sets that give lower function evaluations to ensure a large portion of parameter space is sampled and to prevent the algorithm from getting stuck near local maxima. As iterations continue the system is cooled down and smaller and smaller moves are allowed. Eventually, the algorithm may become greedy, accepting only parameter sets that increase the function value. Execution is terminated when temperature reaches the stopping criterion or if the change in function value between iterations is below tolerance. The algorithm may now be re-run several times to obtain new parameter solutions. If this is performed indefinitely all possible parameters will be explored so the global maximum of the function is guaranteed to be identified.

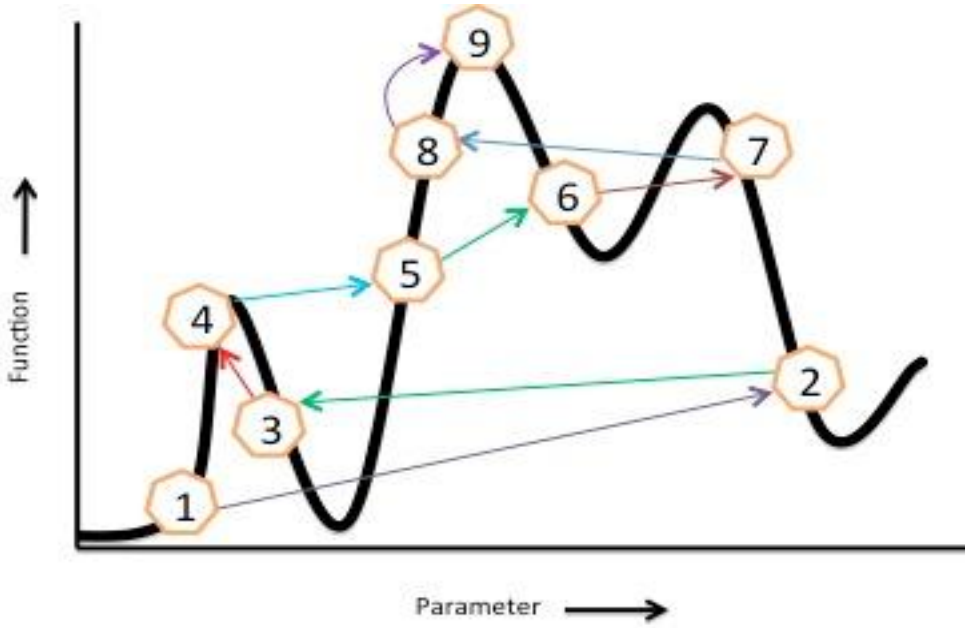



**Figure 1.7:** Example simulated annealing track.

Selecting the specifics of SA algorithms has a large impact on performance. Typically, the various parameters and subroutines are selected empirically although there is much research on determining appropriate approaches (Park *et al.*, 1996; Aarts *et al*., 1997). One method available in SciPy (The SciPy Community) employs the fast cooling schedule and is described in the following pseudo-code:

For k = 0 to $k_{max}$:

$T_{new} = T_0 \cdot \exp(-c \cdot k^{quench})$ with $c = n \cdot \exp(-n \cdot quench)$

For i = 0 to $i_{max}$:

$x_{new} = x_{old} + x_c$

$df = f(x_{old}) - f(x_{new})$

if $df < 0$:

accept $x_{new}$ as a solution

else: (

$p = \exp(-df \cdot 1/boltzmann/T)$

if ($p$ > random (0, 1)): accept $x_{new}$ as a solution

where k is the iteration number, $T_{new}$ is the temperature at iteration k, $T_0$ is the initial temperature, n and quench are user selectable parameters that modify the cooling schedule, i is the dwell sub-iteration number (T does not change during dwell iterations), $x_{new}$ is the new set of parameters, $x_c$ is the move in parameter space that is randomly selected with temperature dependence, $f(x_{old})$ is the evaluation of the function being optimized at the old set of parameters, boltzmann is a user selectable parameter that controls the probability of accepting parameters that are suboptimal, and random (0,1) is random number selected from a uniform distribution.

One advantage of SA over other methods is its ability to provide a good solution for complex problems. Moreover, this algorithm guarantees the optimal solution will be found given sufficient computational time (Kohonen, 1999). However, when computational time is limited, SA results depend greatly on characteristics such as the cooling schedule, and the algorithm becomes susceptible to being trapped at local maxima (Elmohamed, 1998; Elhaddad, 2012).

## 1.5 Atlas Based Segmentation (ABS)

Accurate image delineation in radiotherapy is important to ensure the prostate is sufficiently irradiated and that OARs are appropriately spared. Manual contouring is subject to significant inter and intra-observer variability (Mitchell *et al*., 2009; Livsey *et al.,* 2004) with most of the uncertainty attributed to inter-observer variability (Jameson, *et al.,* 2010). Recently, ABS has been applied in radiotherapy treatment planning to automate prostate and OAR delineation. In one ABS approach, an atlas is constructed by compiling many images of different subjects with corresponding segmentations generated manually by an expert. Thus, the atlas describes the location, shape, and spatial relationship between anatomical structures (Rohlfing *et al*., 2005). The first step for a new patient image is to select the most appropriate atlas subject. This is followed by image registration of the atlas subject and target subject images. Finally, the resulting transformation is used to propagate the atlas subject contours onto the new patient image. This process is represented in Figure 1.8. Clearly, ABS will produce more accurate results if the selected atlas subject is similar to the image to be segmented.

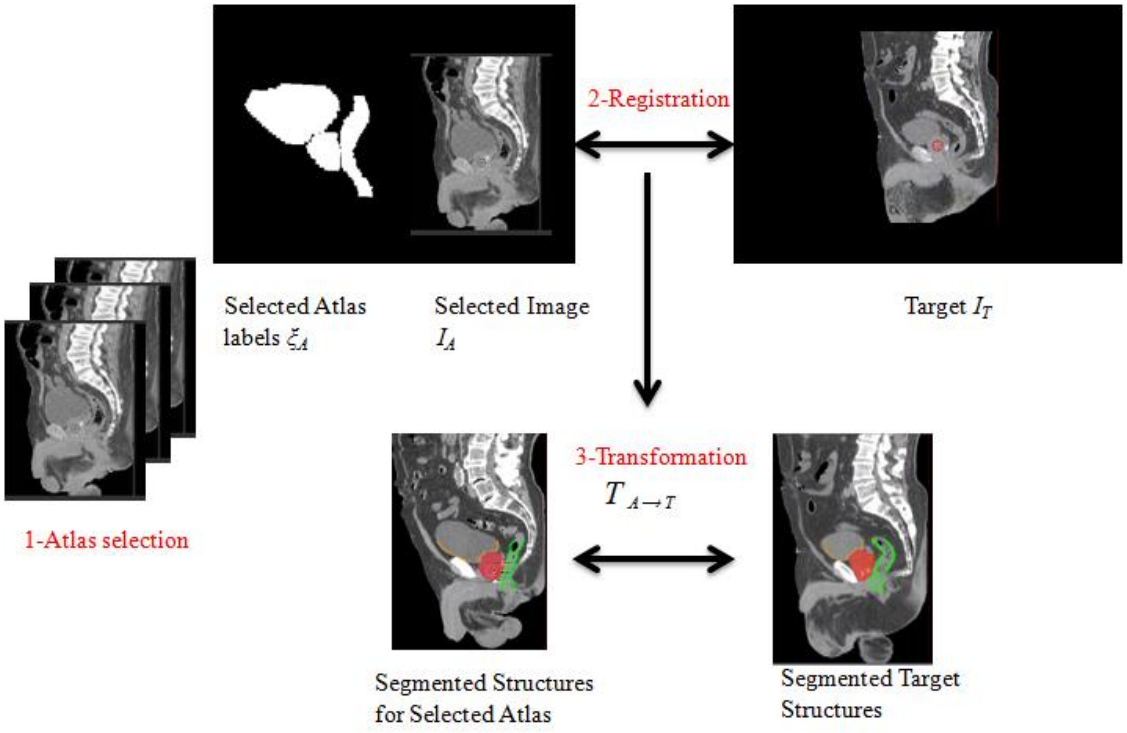



**Figure 1.8:** An example ABS process applied to prostate radiotherapy treatment planning. First, an appropriate atlas subject is selected from a multi-subject database. The selected atlas subject comprises a reference image $I_A$ and contours $\xi_A$ = {prostate, rectum, bladder, right femoral head, left femoral head}. The reference image $I_A$ is then registered to a new target image $I_T$ to obtain the transformation $T_{A\rightarrow T}$ that maps $\xi_A$ to $\xi_T$.

## 1.5.1 Atlas Selection Strategies

Various ABS approaches have been described in the literature. The focus of this thesis is on atlas selection since the decision about what atlas subject to use strongly impacts final segmentation accuracy, yet atlas selection has received little attention in the literature. Furthermore, strategies used for atlas selection are rarely explained in the literature, while simple selection techniques may not be effective (Rolfing *et al.*, 2005). Four main strategies for atlas subject selection have been identified by Rolfing *et al.*, (2005): single fixed atlas subject,

selecting the most similar atlas subject, generation of an average atlas subject, or using multiple atlas subjects.

### 1.5.1.1 Single Atlas Subject

The simplest approach involves careful selection of a single atlas subject for ABS. Clearly, the spatial relationship between various contours in the atlas must be sufficiently generic to represent the anatomical variation expected among patients. The atlas subject is usually selected based on visual inspection to ensure high image quality, contour accuracy, and anatomical variation that best captures the variation expected in the patient population. Single atlas based segmentation was employed by Kikinis *et al.* (1996) who selected one subject to form a brain MRI atlas; not much effort was made to select an appropriate atlas subject. Rohlfing *et al*. (2005) studied the effect of different atlas selection strategies on the accuracy of contours generated for confocal microscopy images of the bee brain. His results showed that constructing a single subject atlas based on visual assessment yielded lower final segmentation results compared to the other selection approaches described below.

### 1.5.1.2 Most Similar Atlas Subject

Selecting the optimal subject from an atlas assumes that for any given image there is one atlas subject that would produce the best segmentation accuracy (Rohlfing *et al*., 2005). The optimal atlas subject may be obtained by registering all atlas subject images to the target, computing image similarity metrics, and selecting the subject that is most similar with the target. This strategy simplifies the remainder of the ABS process since only a single atlas subject is used

but there is a strong dependence between the atlas subject selection process and the final segmentation accuracy (Hoang Duc *et al*., 2013). Leung *et al*., (2010) selected the most similar atlas subject for hippocampal volume in brain MRI according to similarity metric after rigid registration while Rohlfing *et al*. (2005) selected the most similar atlas subject for a bee brain confocal microscopy image based on image similarity after affine and non-rigid registration. Selection after affine registration was shown to be less computationally intensive. Regardless of the transformation model, these techniques rely heavily on the performance of the image registration method that aligns atlas subject and target images. On a practical level, registering all atlas subjects to the target image is extremely time consuming. Finally, similarity metrics do not always detect a best match in terms of final contouring accuracy in ABS (Sanroma *et al*., 2014).

### 1.5.1.3 Average Shape Atlas

An average atlas can be constructed by creating an average image over images of many individuals. One way to obtain an average atlas from a population is to select an arbitrary but representative individual as a reference to which all the original images are registered. An average image is then generated. There are different proposed methods to create an average shape atlas from the population; one way is to obtain an active shape model (ASM) as presented by Cootes *et al.*, (1994) and later used by Rohlfing *et al*., (2005). Creating an ASM requires identification of corresponding landmarks on the shapes (contours) of all subjects, affinely aligning all the landmarks so they correspond as closely as possible, and calculating the mean shape from the aligned shapes. The variation of contour points across the population can then be measured and mathematically described by applying principle component analysis (PCA) (Johnson and Wichern, 1988). Once the average image and average shape are generated, all the

original data may be re-registered with the current averages to generate a new average to use in subsequent iterations. This iterative process may be carried on until convergence (Rohlfing *et al*., 2005).

Acosta *et al*. (2010) used another method to build an atlas for prostate CT. This technique was first used by Downling *et al.* (2009) for prostate MRI segmentation. An average image was constructed as described above. The affine plus non-rigid transformations obtained during this process were then used to align binary images representing population contours onto the average image space. An averaging process was then performed on the binary data to obtain probabilistic maps for each contour.

In the methods above a separate procedure for atlas subject selection is unnecessary because shape variability across the population is contained within the atlas itself (using PCA or as a probability map). Another advantage of the image averaging process is that it reduces imaging artifacts and noise. However, the average atlas approach requires substantial computational time. Furthermore, adding subjects into the atlas is difficult since a full recomputation of the average is necessary.

### 1.5.1.4 Multiple Atlas Subjects

This process requires the collection of many subjects' images and their corresponding contours. To segment an unknown target image, atlas subjects within the database are registered to the target image and their transformed segmentations combined to estimate contours in the target image. One way to combine several segmentations is using a "majority voting" algorithm (Kittler *et al*., 2003). The outputs of the subject atlas contours are determined for each voxel in

the target image, their "votes" are counted and the label that gained the highest number of votes in a particular voxel is selected to represent that voxel. This process requires contours to be interpolated onto the 3D image grid.

Recently, multi-subject ABS was shown to be more accurate compared to a single subject or average atlas image approach (Rohlfing *et al.*, 2005). This is because the multi-subject atlas inherently contains large anatomical variability across the population. However, a multi-subject atlas requires additional computational time since all atlas subjects have to be aligned with a given target. Furthermore, the process of combining segmentations into one final contour set is not trivial. For example, Acosta *et al.*, (2013) used a technique where the final segmentation is a probability map – this describes the inherent uncertainty of contouring CT images. However, in the conventional radiotherapy process, a single, definite set of final contours is required.

Currently, there are commercially available ABS algorithms such as the one offered by MIMvista (MIMvista Corp, Cleveland, OH). From the quoted literature, it is not clear what atlas selection strategy is employed. The company compared two approaches for a prostate cancer CT atlas: most similar atlas subject and multiple atlas subjects. It appears that the most similar subject selection process involved computing an intensity-based similarity metric (e.g. MI) for all atlas subjects and the new target image, and selecting the atlas subject with the highest metric. In the multiple subject atlas method, the same process was repeated but the top three to five atlas subjects were selected and carried through the ABS pipeline. The final segmentations were then combined using a "vote rule" (Pirozzi *et al.*, 2012). The results were similar to what Rholing *et al.* (2005) concluded: multi subject atlas approaches yield higher accuracy than the most similar atlas subject. In a slightly different approach, a user provided bladder volume was used to

perform the most similar atlas subject selection (Lin *et al.*, 2008). This procedure requires bladder volume to be known or estimated and the images employed had contrast agent in the bladder to assist in the contouring. There were no further details on atlas selection for the other ROIs.

## 1.6 Evaluation of Image Segmentation

In this thesis, segmented anatomy is represented by a binary image with 1s and 0s representing inside and outside, respectively. Similarly, a geometrical surface may be used to represent anatomical boundaries. To quantify segmentation accuracy it is important to measure the difference between contours obtained using one method (e.g. ABS) vs. a gold standard (e.g. manual segmentation). Two metrics for measuring segmentation differences are the Dice Similarity Coefficient and Hausdorff distance. These evaluation methods usually take the manually drawn contour as the ground truth. However, manual segmentation results are subject to inter and intra-observer variability, complicating the process of quantifying the accuracy of an automated segmentation method such as ABS.

### 1.6.1 Dice Similarity Coefficient

The Dice Similarity Coefficient (DSC) quantifies the spatial overlap between two segmentation results on a 0 to 1 scale, where 0 indicates no overlap and 1 indicates perfect overlap (Lee, 1945).

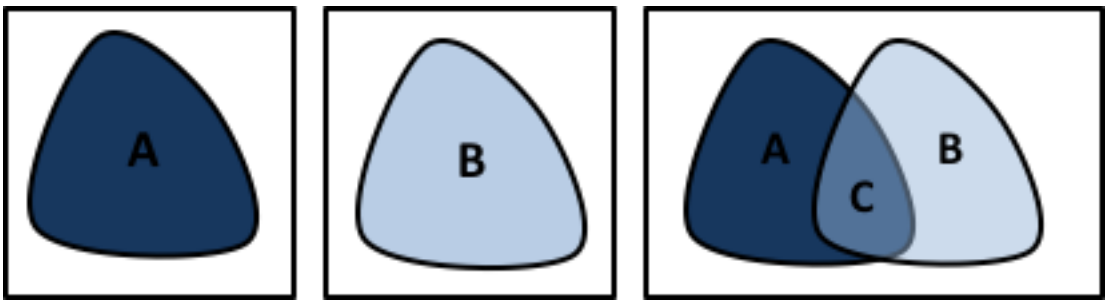



**Figure 1.9:** DSC is calculated based on the area or volume contained in regions A and B, and their overlap C.

For the example in Figure 1.9, the DSC is given by:

$$\mathrm{DSC} = \frac{2\,|A \cap B|}{|A|+|B|} = \frac{2|C|}{|A|+|B|}\ , \qquad (1.12)$$

where A and B are the areas of the two independent regions and C is their overlapping area. In prostate ABS, DSC was used to evaluate accuracy versus manually drawn contours (Velker *et al*., 2013).

## 1.6.2 Hausdorff Distance

Another common method for evaluating segmentation accuracy is to measure the distance between two segmentations (Huttenlocher *et al*., 1993). The Hausdorff distance is the maximum of all possible minimum distances between two sets of segmentations. Consider the example in Figure 1.10. The Hausdorff distance between A and B can be obtained by calculating the minimum of the distance between each point on A and every point on B and taking the maximum value labeled as $d_1$ in the given example. The same procedure is repeated to calculate

the minimum of the distance between each point on B and every point on A and taking the maximum value labeled as $d_2$. The final Hausdorff distance is the largest of $d_1$ and $d_2$.

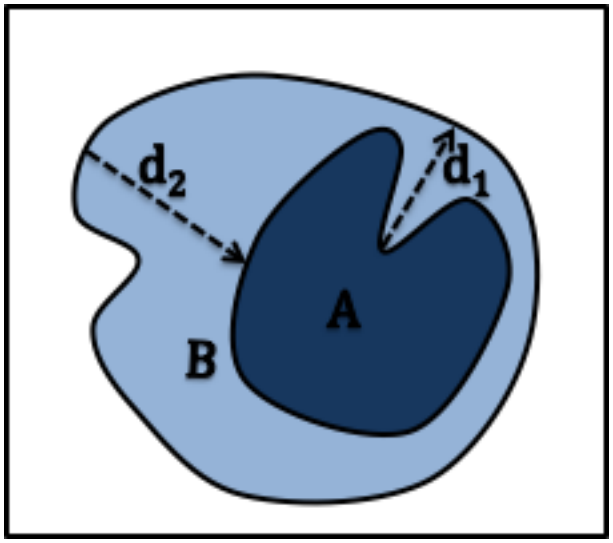



**Figure 1.10:** The Hausdorff distance for the contours A and B is given by $d_2$ since it is larger than $d_1$.

## 1.7 Purpose of Thesis

Successful IMRT of prostate cancer relies on accurate delineation of the prostate and OARs. Recently, ABS was employed in radiotherapy treatment planning to automate prostate and OAR delineation. In a typical approach, the essential step in ABS is the selection of an atlas subject from a database that best matches the target image. Pervious atlas selection methods employed various approaches with different advantages and disadvantages.

This thesis presents an alternative atlas selection strategy in the most similar atlas subject framework. Specifically, a single atlas subject will be selected for a given region in a new target image based on the similarity between anatomical characteristics. Therefore, different atlas subjects will be used to contour different anatomy in the target image. The anatomical characteristics used for selection include measurements such as prostate length and patient thickness. These were chosen because prostate length indicates the overall prostate shape while the overall body shape may be represented using the anterior-posterior and lateral separation.

The aim of this work is to propose an atlas selection strategy and evaluate its impact on the final segmentation accuracy. In order to do that, the first step was to create a training set of 20 patients. The data for each patient comprised a planning CT image and segmented prostate and OARs (rectum, bladder, and both femoral heads). A brute force procedure was then performed using affine image registration to measure the volume of overlap between segmentations. These values were used to identify subject pairs with the most compatible contours. The agreement in various anatomical measurements for best subject pairs was analyzed and an atlas-target subject similarity score was proposed. The similarity computation

was optimized to maximize the final DSC obtained with the training set. This new atlas selection approach was tested on 10 new patients to evaluate the selection performance.

ABS holds considerable promise for future application in clinical use to improve contouring efficiency and reduce variability. For such an application, the selection of the best atlas subject has a critical impact on segmentation accuracy. This work simultaneously proposes an atlas selection technique and evaluates its success. These methods may extended to process additional patient data, contour additional anatomy, and even extended to different areas of the body.

# Chapter 2

# Materials and Method

## 2.1. Computational Environment

The main part of this work was performed on a workstation running the CentOS 5.10 Linux operating system (The CentOS Project). Images were processed using the Python 2.4 scripting language (Python Software Foundation, Delaware, USA) and a custom version of the Visualization Tool Kit (Kitware Inc., Clifton Park, USA). A detailed list of scripts used in this work may be found in Appendix 1.

## 2.2. Patient Data

Anonymized planning CT images of 30 prostate cancer patients undergoing IMRT at the Juravinski Cancer Center between June 2012 and January 2013 were collected for this work. Patients were selected randomly although images with unusual features such as hip prostheses or uncommon size or shape of anatomy were excluded. It was assumed the patients had early stage prostate cancer given that the prescription was either 76 or 78 Gy over seven weeks. Ethics

approval for collecting this dataset was obtained from the McMaster University Research Ethics Board.

All images were acquired with the patient supine with a full bladder using a Philips Brilliance Big Bore CT scanner (Philips Medical Systems, Amsterdam, Netherlands) and the following parameters: helical mode, 120 kVp, 212-267 mAs, and 50-65 cm field of view. Images were reconstructed on 3 mm slices. The CTV, rectum, bladder, right femoral head, and left femoral head were contoured either by a radiation oncologist or a dosimetrist with later approval from the attending oncologist. The CTV contours were modified to remove the base of the seminal vesicles, obtaining a prostate ROI for use for the remainder of this study. This was deemed necessary due to the large variability of the seminal vesicles across the population and previous reports of difficulties in contouring this area of an image accurately (Fiorino *et a*l., 1998).

The following parameters were measured for each patient image for later analysis: anterior-posterior (AP) and lateral (LR) patient thicknesses at the prostate centroid, volume of each ROI, distance between the prostate centroid and centroids of all other ROIs, cranial-caudal prostate length (PL), and the maximum diameters of the right femoral head (RFHD) and the left femoral head (LFHD). Example measurements are shown in Figure 2.1. All measurements except PL, RFHD, and LFHD were performed automatically using the scripting environment in the Pinnacle TPS (Philips NV, Amsterdam, Netherlands). PL was measured in the sagittal plane using the prostate contour as a guide while femoral head diameters were measured in the transverse slice showing the largest bone cross section.

As explained below, these parameters will be used to select an atlas subject that is most compatible with a given target image. Thus, it is important that they can be measured precisely and efficiently. This can be achieved by having the user place points near the centroids of the prostate, rectum, bladder, and femoral heads. These points can be used to estimate most of the necessary measurements. The user will then have to measure PL and femoral head diameters manually. It is estimated that these measurements and the placing of the centroid points will require up to five minutes. Estimating the volume of ROIs is more complicated, but as discussed below, it may not be necessary.

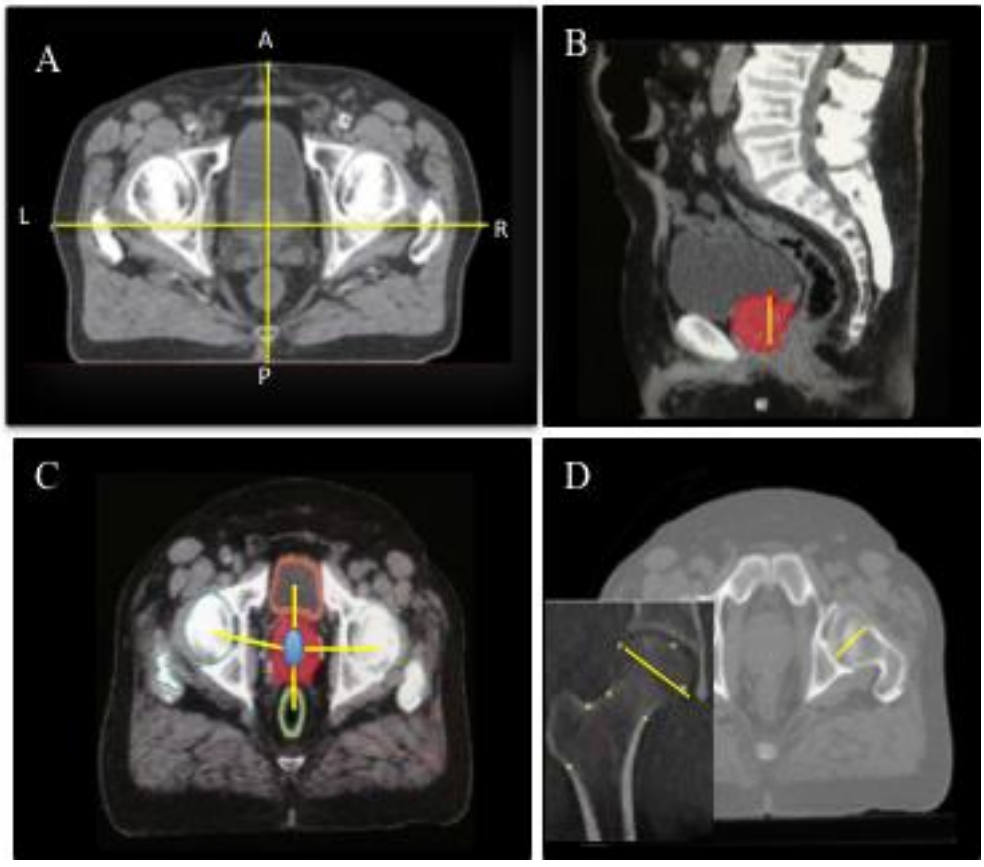



**Figure 2.1:** Parameters measured for each patient: panel A shows the anterior-posterior (AP) and lateral (LR) patient thickness, panel B shows the cranial-caudal prostate length (PL), panel C shows the distance between the prostate centroid and centroids of all other ROIs, and panel D shows the right and left femoral head diameters (RFHD and LFHD)

## 2.3. Image Preparation

The planning CT images and ROIs were obtained from the Pinnacle TPS. Several pre-processing steps were necessary for further analysis. First, all ROIs were converted into binary images with 1s indicating inside the ROI and 0s outside. The second step was to create binary mask images corresponding to the area of interest in each CT image. This was constructed by applying the OR logical operator to binary images representing the prostate, rectum, bladder, and both femoral heads. Finally, the result was morphologically dilated by 2 cm. The process is illustrated in Figure 2.2. The last preprocessing step involved resampling each CT image and its derived binary data to match the voxel size in all remaining images. This creates one-to-one correspondence between voxels in each image pair, simplifying various data analyses described later.

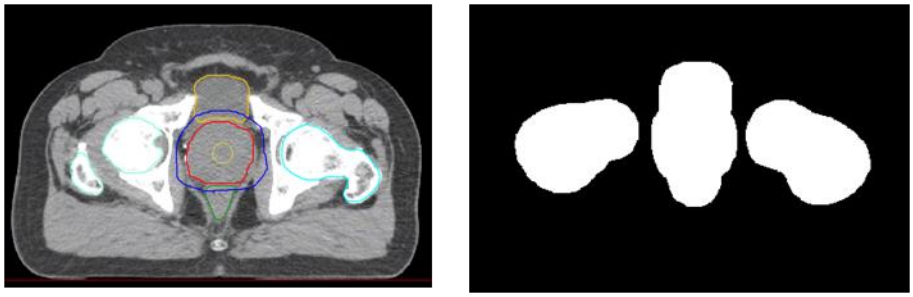


**Figure 2.2:** Procedure used to derive the mask representing the region of interest in each image for further analysis.

## 2.4. Image Registration

The pre-processed images were registered to enable the analysis described in Section 2.5.1. All registrations were performed using a modified version of a previously validated algorithm (Wierzbicki *et al.*, 2010). A nine DOF transformation model was employed for registration allowing 3D rotation, translation, and scaling. The downhill simplex optimizer was employed to

maximize MI in the volume described by the binary mask described above. Registration constraints were imposed to disallow any rotations greater than 10 degrees and scaling over 10 percent. This strategy reduced the optimization search space, improving the robustness of the registration algorithm. Every registration result was visually validated by comparing the transformed CT with the planning CT image.

## 2.5 Atlas Selection

An atlas was constructed comprising CT images of 20 out of the 30 patients analyzed in this work along with the associated contours of the prostate, rectum, bladder, and both femoral heads. This atlas was used to evaluate and develop the following atlas selection methods.

### 2.5.1 Brute Force

This method is not directly applicable in ABS because it requires that any new target image has already been contoured. However, it does allow the identification of atlas-target subject pairs that have the most compatible contours. This was achieved by performing 20-leave-one-out experiments where each subject image was selected as the target to which all remaining 19 subject images were registered as described in Section 2.4. There were a total of 380 registrations required (20 target subjects x 19 source subjects per target). The resulting transformations were applied to the source binary images representing the prostate, rectum, bladder, and both femoral heads. The agreement between any transformed contours and manually derived target contours was quantified using DSC. Finally, for each subject, the remaining 19 subjects were ranked according to the achieved DSC. The ranking was performed separately for each ROI to determine the best and poorest source subjects for any target. Thus, it

is possible that different subjects have the most compatible ROIs depending on the type of ROI. This method is illustrated in Figure 2.3.

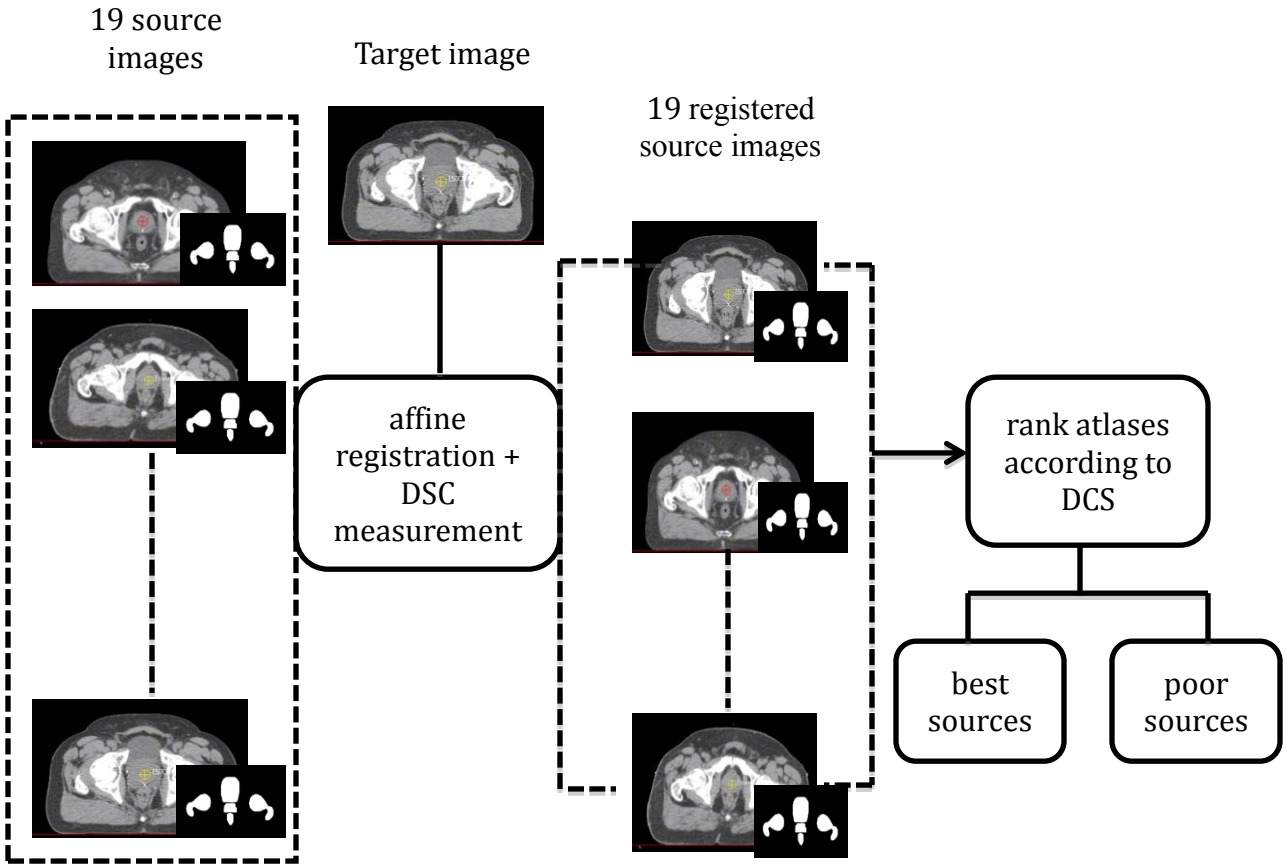



**Figure 2.3:** Brute force quantification of subject contour compatibility based on DSC.

## 2.5.2 Correlation of Single Anatomical Measurements

This atlas selection strategy compares anatomical measurements performed on images to find an atlas subject that is most similar to a target. Details of these measurements are described in Section 2.2. This method is feasible for use in ABS since it is not time consuming to make the measurements on a new target image and, since only a few parameters are involved, an exhaustive comparison with all atlas subjects is possible. The first goal was to measure correlation in particular anatomical measurements between each of the 20 targets and their best matched atlas subjects. In this case, "best matched" means the atlas subject that gave the highest DSC for the target as determined in the brute force method. For the best matched subject pairs plots of all anatomical measurements in target vs. best atlas subject were generated and the Pearson's product-moment correlation coefficient was computed. This process was used to identify the anatomical measurement that was most correlated for atlas selection. For a new target image, the measurement would be made and the atlas subject having the most similar measurement would be selected for ABS.

## 2.5.3 Correlation of Anatomical Measurement Combinations

It is possible that single parameters measured in targets and best-matched atlas subjects are poorly correlated. In this case, an optimal combination of several parameters may prove more useful. As described in Section 2.2, a total of 14 parameters were available and it is not trivial to determine how to combine them to improve correlation. To reduce complexity, ROI volumes were excluded because a full ROI contour is required to compute volume, eliminating the need for ABS. Moreover, volumes of most ROIs in the pelvic region have a large inter and intra patient variability (Livsey *et al*., 2004; Nishioka *et al*., 2013). It was proposed that the

anatomical measurements remaining in the analysis should be combined to compute a cost for each particular ROI and atlas-target subject combination as follows:

$$C_{ROI} = \sum_{M} W_{ROI,M} \left| 1 - \frac{M_A}{M_T} \right|, \qquad (2.1)$$

where M is an anatomical measurement (PL, AP, LR, RFHD, LFHD, prostate-BW, prostate-RW, prostate-RF, prostate-LF), $W_{ROI,M}$ is the weight assigned to M for the particular ROI, $M_A$ is the measurement in an atlas subject, $M_T$ is the measurement in the target subject, and $C_{ROI}$ is the cost for a particular ROI scoring the discrepancy between measurements in an atlas and a target image (C decreases as the atlas and target subjects become more similar). During atlas selection, the cost function may be computed between the target and all atlas subjects to find the atlas subject that is the closest match (minimal C). This is then repeated for each ROI to identify the optimal atlas subject for each contour. This form uses weighting factors to designate the importance of each anatomical measurement in selecting the atlas subject. An optimization problem was then formulated where the weights are the parameters and the goal is to maximize the DSC possible for each target and ROI. The problem was solved for the 20 images, with each subject acting as a target and the remaining 19 acting as an atlas. No additional image registration or DSC computation was necessary since all results were already established in the brute force analysis.

Optimization was performed using the SA method described in section 1.4.3.2**.** The algorithm randomly selects a set of weighting factors. For each target subject, the function C is computed for all 19 subjects comprising each individual atlas. The atlas subject giving the lowest C is chosen and the final DSC following affine registration is looked up from the results of the brute force experiment. This is repeated for all 20 targets to obtain 20 DSC values for any

ROI. Thus, the mean DSC represents the contouring accuracy possible for the ROI if atlas selection is performed using equation (2.1) with the current weight factors and if the remaining steps of ABS only involve affine registration. Overall, the SA method selects weight factors that maximize mean DSC. This process was automated using a Python script that employs the SA module included in the SciPy package. Table 2.1 outlines the parameters of the optimizer, description of SA optimization code may be found in Appendix 2. The optimization was repeated 10,000 times since the chance of finding the global maximum with SA improves with additional computational time (Kohonen, 1999).

| Parameter | Value |
|---|---|
| Initial solution $x_0$ (initial weights) | 1.0 |
| Initial temperature $T_0$ | None (automatically computed) |
| Final temperature $T_f$ | 1e-12 |
| maxiter ($k_{max}$) | 400 |
| dwell ($x_{max}$) | 250 |
| Boltzmann Constant | 1.0 |
| n, quench **(**parameters to change the cooling schedule) | 1.0 |

**Table 2.1:** Simulated annealing parameters used to optimize the atlas selection strategy. All parameters were set to their default values except maxiter, which was set to 400 using trial and error.

For comparison, an additional experiment were carried out using a leave-one-out scheme to assess segmentation accuracy based on random selection, where for each subject in the training set, an atlas subject was a randomly selected.

## 2.6. Atlas Selection Validation

Validation is considered one of the most important aspects when a new algorithm is proposed for clinical use. Thus, the proposed atlas selection algorithm was validated using 10 additional planning CT images (testing dataset). The goal was to validate not only the atlas selection algorithm but also to measure final segmentation accuracy. The process of validation is summarized as follows:

1. New data were prepared as described in Section 2.3. Manually drawn contours were used to obtain the gold standard segmentation result. To assess the possibility of maximizing accuracy by adding more patients to the atlas dataset, brute force was repeated with all 30 patients as described in section 2.5.1.
2. Anatomical parameters were measured for each new subject as outlined in Section 2.2.
3. Sets of weights resulting in high DSCs for the 20 subject training data set were identified from the total of 10,000 generated for each ROI. Details of the selection process are described later.
4. The optimized atlas selection method was applied for each new target image in the testing dataset. The atlas for this process comprised the 20 subject training set. For each pair the segmentation accuracy was measured using DSC and Hausdorff distance.
5. A random atlas selection process was also performed where, for each of the 10 subjects in the validation set, an atlas subject was randomly selected from the 20 subject training set.

Chapter 3

# Results and Discussion

## 3.1. Anatomical Characteristics

The measured parameters for the 20 study subjects used to design an atlas selection strategy are summarized in Table 3.1. Some parameters showed larger variability across the population, making them potentially more important among the other parameters when attempting to pair a target image with an atlas subject. Parameters such as AP and LR patient thickness are important because they indicate the overall body shape. Since the prostate is approximately bracketed by bladder, rectum and femoral heads, the distance between those organs indicates the position of the prostate. As already mentioned, ROI volumes were removed from further analysis despite providing valuable information since accurate computation of volume requires the ROI to be delineated.

The remaining parameters were relatively easy to measure in the TPS. For the 20 subjects, it took about 10 minutes to measure PL on the sagittal slice. Similarly, measuring RFHD and LFHD took a combined time of 15 minutes on a transverse slice. Distances between prostate and other organ centroids were obtained automatically from existing contours. However, it is reasonable to assume this may be performed efficiently since manually placing points near the centroid of each ROI and automatically measuring the distance between them is not expected to be time consuming. Furthermore, there is a potential to reduce the time necessary to perform the measurements. For example, it appears that only one femoral head diameter is needed since there was no statically significant difference between the right and left measurements ($p > 0.05$). All the anatomical characteristics can be found in Appendix 3.

| Parameter | Mean ± SD (cm) | Parameter | Mean ± SD (cm) |
|---|---|---|---|
| **PL** | 4.99 ± 0.93 | **prostate-BW distance** | 4.60 ± 0.65 |
| **RFHD** | 4.69 ± 0.40 | **prostate - RW distance** | 3.79 ± 0.44 |
| **LFHD** | 4.71± 0.38 | **prostate - RF distance** | 11.38 ± 0.57 |
| **AP thickness** | 23.56 ± 2.33 | **prostate - LF distance** | 11.26 ± 0.45 |
| **LR thickness** | 37.57 ± 1.96 | | |

**Table 3.1:** Means and standard deviation (SD) for the measured anatomical characteristics for the 20 subjects used to design the atlas selection method.

## 3.2. Atlas Selection

### 3.2.1. Brute Force

Image registration took approximately 20 minutes per image pair, summing to approximately 5 hours for the 20-leave-one-out experiments performed on the training dataset. Visual assessment of affine registration results indicated that differences in patient positioning and posture were corrected; however, differences in volume and shape of internal soft tissues were not. Figure 3.1 demonstrates a typical registration result showing that bones are well registered while some soft tissue needed further improvement. It is clear that the two coccyx bones are not aligned, lateral edges of the bladder are misaligned, and there is poor overlap for rectum and seminal vesicles. In general, the affine alignment of rectum and bladder was unacceptable due to large variations of these organs across the population. Thus, the remaining focus of this work was on ABS of the femoral heads and prostate only.

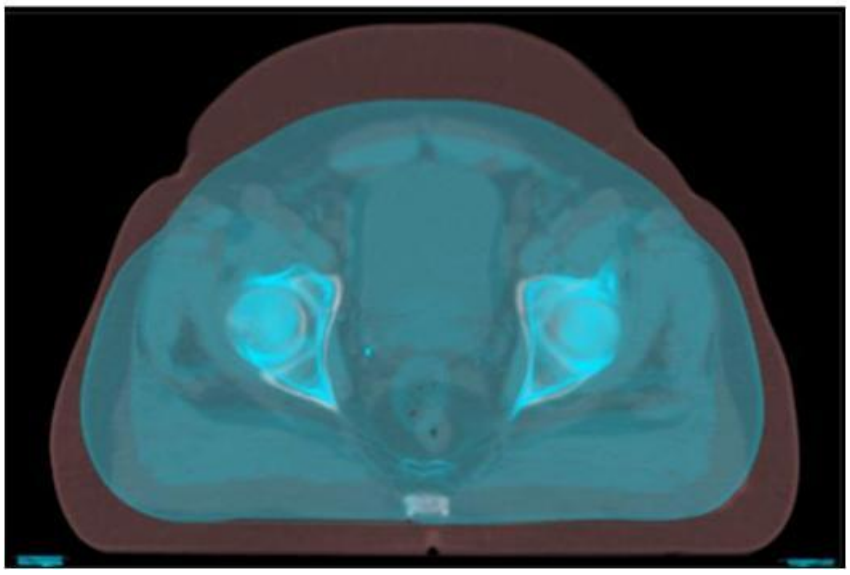


**Figure 3.1:** Typical affine registration result. The target image is shown in blue while the aligned source image is in red.

All DSCs obtained by brute force are shown in Tables 3.2, 3.3, and 3.4. Identifying the maximum DSC for each subject produced a mean ± standard deviation (SD) DSC across the population of 0.90 ± 0.02 and 0.78 ± 0.05 for both femoral heads and prostate, respectively. It is not surprising that excellent automated contours were possible for the femoral heads due to the strong bone contrast seen in CT; this assists the registration algorithm in aligning bone. The mean DSC for the prostate was nearly 0.8 which compares favorably to 0.67 achieved by Pate *et al.* (2014) for generating a prostate ABS. The brute force DSCs were above the 0.7 level previously achieved through manual re-contouring of the prostate by the same observer (Granberg *et al.,* 2011) and 0.65 achieved by Hwee *et al.* (2011) for manual re-contouring of the prostate bed by different observers.

Overall, it appears that excellent automated segmentation results are possible if the target image is paired with the most appropriate subject in the atlas. This is encouraging since these results were obtained using only affine registration. It is reasonable to assume that employing deformable image registration in the ABS technique would provide additional gains in contouring accuracy as would increasing the size of the atlas.

| Target Subjects \ Atlas Subjects | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | | 0.57 | 0.47 | 0.49 | 0.56 | 0.49 | 0.52 | 0.53 | 0.46 | 0.63 | 0.84 | 0.67 | 0.61 | 0.76 | 0.72 | 0.54 | 0.70 | 0.68 | 0.54 | 0.56 |
| **2** | 0.56 | | 0.20 | 0.17 | 0.35 | 0.21 | 0.56 | 0.52 | 0.20 | 0.46 | 0.55 | 0.54 | 0.42 | 0.45 | 0.46 | 0.31 | 0.48 | 0.56 | 0.34 | 0.31 |
| **3** | 0.46 | 0.23 | | 0.81 | 0.67 | 0.75 | 0.35 | 0.23 | 0.81 | 0.64 | 0.50 | 0.34 | 0.43 | 0.65 | 0.46 | 0.85 | 0.39 | 0.39 | 0.53 | 0.58 |
| **4** | 0.50 | 0.25 | 0.83 | | 0.75 | 0.22 | 0.17 | 0.84 | 0.75 | 0.58 | 0.40 | 0.49 | 0.67 | 0.55 | 0.84 | 0.47 | 0.46 | 0.59 | 0.65 | 0.50 |
| **5** | 0.54 | 0.30 | 0.66 | 0.76 | | 0.70 | 0.36 | 0.27 | 0.75 | 0.72 | 0.66 | 0.57 | 0.62 | 0.68 | 0.58 | 0.70 | 0.58 | 0.62 | 0.65 | 0.60 |
| **6** | 0.45 | 0.25 | 0.71 | 0.65 | 0.68 | | 0.25 | 0.25 | 0.57 | 0.63 | 0.55 | 0.45 | 0.52 | 0.51 | 0.40 | 0.65 | 0.50 | 0.44 | 0.48 | 0.42 |
| **7** | 0.50 | 0.60 | 0.14 | 0.18 | 0.32 | 0.29 | | 0.76 | 0.20 | 0.72 | 0.55 | 0.64 | 0.65 | 0.35 | 0.32 | 0.21 | 0.70 | 0.70 | 0.29 | 0.17 |
| **8** | 0.47 | 0.54 | 0.06 | 0.13 | 0.25 | 0.22 | 0.72 | | 0.11 | 0.69 | 0.48 | 0.64 | 0.50 | 0.25 | 0.32 | 0.20 | 0.65 | 0.70 | 0.56 | 0.15 |
| **9** | 0.46 | 0.20 | 0.82 | 0.85 | 0.76 | 0.61 | 0.21 | 0.10 | | 0.68 | 0.52 | 0.40 | 0.47 | 0.66 | 0.56 | 0.80 | 0.40 | 0.40 | 0.64 | 0.72 |
| **10** | 0.72 | 0.50 | 0.70 | 0.77 | 0.74 | 0.61 | 0.70 | 0.73 | 0.70 | | 0.71 | 0.66 | 0.66 | 0.68 | 0.59 | 0.66 | 0.84 | 0.82 | 0.61 | 0.58 |
| **11** | 0.84 | 0.55 | 0.47 | 0.52 | 0.65 | 0.56 | 0.56 | 0.51 | 0.50 | 0.80 | | 0.74 | 0.71 | 0.72 | 0.66 | 0.59 | 0.71 | 0.74 | 0.64 | 0.55 |
| **12** | 0.71 | 0.61 | 0.33 | 0.41 | 0.58 | 0.43 | 0.69 | 0.67 | 0.39 | 0.67 | 0.76 | | 0.78 | 0.62 | 0.61 | 0.41 | 0.70 | 0.71 | 0.65 | 0.38 |
| **13** | 0.62 | 0.41 | 0.39 | 0.48 | 0.60 | 0.56 | 0.62 | 0.56 | 0.47 | 0.78 | 0.69 | 0.79 | | 0.61 | 0.58 | 0.47 | 0.80 | 0.80 | 0.55 | 0.44 |
| **14** | 0.76 | 0.47 | 0.60 | 0.67 | 0.69 | 0.53 | 0.37 | 0.33 | 0.68 | 0.70 | 0.76 | 0.57 | 0.61 | | 0.61 | 0.74 | 0.70 | 0.57 | 0.74 | 0.70 |
| **15** | 0.71 | 0.49 | 0.44 | 0.58 | 0.59 | 0.42 | 0.36 | 0.57 | 0.52 | 0.61 | 0.74 | 0.61 | 0.64 | 0.74 | | 0.57 | 0.50 | 0.64 | 0.77 | 0.59 |
| **16** | 0.52 | 0.30 | 0.85 | 0.48 | 0.68 | 0.69 | 0.23 | 0.21 | 0.77 | 0.67 | 0.59 | 0.41 | 0.49 | 0.69 | 0.53 | | 0.42 | 0.46 | 0.55 | 0.64 |
| **17** | 0.66 | 0.51 | 0.34 | 0.42 | 0.56 | 0.50 | 0.66 | 0.63 | 0.39 | 0.70 | 0.66 | 0.67 | 0.77 | 0.55 | 0.52 | 0.46 | | 0.77 | 0.57 | 0.42 |
| **18** | 0.65 | 0.57 | 0.34 | 0.43 | 0.60 | 0.42 | 0.72 | 0.69 | 0.40 | 0.83 | 0.72 | 0.69 | 0.79 | 0.61 | 0.53 | 0.40 | 0.79 | | 0.59 | 0.42 |
| **19** | 0.74 | 0.43 | 0.53 | 0.61 | 0.68 | 0.50 | 0.66 | 0.27 | 0.63 | 0.61 | 0.67 | 0.61 | 0.51 | 0.79 | 0.81 | 0.56 | 0.60 | 0.65 | | 0.61 |
| **20** | 0.56 | 0.32 | 0.58 | 0.66 | 0.61 | 0.44 | 0.17 | 0.14 | 0.72 | 0.58 | 0.57 | 0.38 | 0.45 | 0.70 | 0.61 | 0.65 | 0.44 | 0.43 | 0.61 | |

**Table 3.2:** DSC values for prostate obtained during the brute force experiment. White cells indicate DSC values from 0.8 to unity, blue from 0.60 to 0.79, and red below 0.6.

| Target Subjects \ Atlas Subjects | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | | 0.85 | 0.80 | 0.88 | 0.82 | 0.81 | 0.83 | 0.84 | 0.88 | 0.81 | 0.91 | 0.91 | 0.88 | 0.86 | 0.84 | 0.84 | 0.81 | 0.87 | 0.85 | 0.89 |
| **2** | 0.83 | | 0.82 | 0.87 | 0.77 | 0.78 | 0.77 | 0.81 | 0.82 | 0.71 | 0.80 | 0.80 | 0.89 | 0.88 | 0.86 | 0.81 | 0.74 | 0.57 | 0.81 | 0.84 |
| **3** | 0.79 | 0.80 | | 0.91 | 0.87 | 0.82 | 0.59 | 0.77 | 0.85 | 0.84 | 0.88 | 0.91 | 0.86 | 0.82 | 0.86 | 0.86 | 0.77 | 0.82 | 0.88 | 0.91 |
| **4** | 0.86 | 0.82 | 0.88 | | 0.91 | 0.87 | 0.76 | 0.87 | 0.88 | 0.86 | 0.85 | 0.90 | 0.88 | 0.89 | 0.89 | 0.90 | 0.81 | 0.80 | 0.85 | 0.87 |
| **5** | 0.83 | 0.77 | 0.87 | 0.92 | | 0.85 | 0.75 | 0.82 | 0.86 | 0.88 | 0.82 | 0.85 | 0.86 | 0.89 | 0.89 | 0.91 | 0.78 | 0.75 | 0.81 | 0.88 |
| **6** | 0.81 | 0.78 | 0.81 | 0.88 | 0.85 | | 0.74 | 0.78 | 0.83 | 0.83 | 0.76 | 0.81 | 0.78 | 0.84 | 0.83 | 0.85 | 0.83 | 0.75 | 0.73 | 0.82 |
| **7** | 0.82 | 0.76 | 0.71 | 0.78 | 0.75 | 0.74 | | 0.77 | 0.79 | 0.48 | 0.82 | 0.82 | 0.83 | 0.79 | 0.81 | 0.74 | 0.74 | 0.74 | 0.77 | 0.83 |
| **8** | 0.86 | 0.80 | 0.82 | 0.87 | 0.83 | 0.79 | 0.77 | | 0.90 | 0.52 | 0.84 | 0.87 | 0.86 | 0.88 | 0.81 | 0.85 | 0.86 | 0.81 | 0.59 | 0.82 |
| **9** | 0.89 | 0.82 | 0.83 | 0.88 | 0.86 | 0.87 | 0.77 | 0.91 | | 0.78 | 0.83 | 0.85 | 0.85 | 0.85 | 0.82 | 0.85 | 0.89 | 0.80 | 0.76 | 0.84 |
| **10** | 0.51 | 0.59 | 0.84 | 0.86 | 0.89 | 0.78 | 0.38 | 0.47 | 0.78 | | 0.51 | 0.77 | 0.80 | 0.79 | 0.89 | 0.84 | 0.52 | 0.39 | 0.84 | 0.83 |
| **11** | 0.91 | 0.79 | 0.90 | 0.88 | 0.83 | 0.77 | 0.86 | 0.85 | 0.85 | 0.82 | | 0.90 | 0.91 | 0.84 | 0.87 | 0.83 | 0.79 | 0.83 | 0.81 | 0.90 |
| **12** | 0.89 | 0.80 | 0.90 | 0.90 | 0.89 | 0.83 | 0.83 | 0.87 | 0.86 | 0.41 | 0.91 | | 0.89 | 0.88 | 0.88 | 0.89 | 0.76 | 0.83 | 0.87 | 0.92 |
| **13** | 0.88 | 0.90 | 0.86 | 0.88 | 0.85 | 0.79 | 0.84 | 0.86 | 0.86 | 0.35 | 0.91 | 0.90 | | 0.83 | 0.79 | 0.86 | 0.79 | 0.76 | 0.79 | 0.88 |
| **14** | 0.85 | 0.88 | 0.86 | 0.89 | 0.89 | 0.85 | 0.77 | 0.87 | 0.86 | 0.81 | 0.83 | 0.87 | 0.83 | | 0.89 | 0.92 | 0.81 | 0.76 | 0.81 | 0.89 |
| **15** | 0.84 | 0.84 | 0.87 | 0.90 | 0.89 | 0.84 | 0.82 | 0.66 | 0.83 | 0.90 | 0.83 | 0.86 | 0.74 | 0.88 | | 0.88 | 0.75 | 0.57 | 0.85 | 0.89 |
| **16** | 0.84 | 0.80 | 0.85 | 0.90 | 0.90 | 0.84 | 0.75 | 0.84 | 0.84 | 0.87 | 0.82 | 0.88 | 0.85 | 0.92 | 0.86 | | 0.79 | 0.74 | 0.80 | 0.88 |
| **17** | 0.81 | 0.73 | 0.77 | 0.82 | 0.79 | 0.83 | 0.73 | 0.87 | 0.89 | 0.67 | 0.77 | 0.77 | 0.79 | 0.81 | 0.75 | 0.79 | | 0.85 | 0.63 | 0.77 |
| **18** | 0.87 | 0.54 | 0.83 | 0.83 | 0.76 | 0.74 | 0.73 | 0.82 | 0.81 | 0.44 | 0.85 | 0.82 | 0.78 | 0.77 | 0.71 | 0.76 | 0.85 | | 0.63 | 0.81 |
| **19** | 0.68 | 0.72 | 0.88 | 0.84 | 0.82 | 0.76 | 0.42 | 0.76 | 0.76 | 0.84 | 0.88 | 0.86 | 0.78 | 0.80 | 0.86 | 0.81 | 0.58 | 0.45 | | 0.87 |
| **20** | 0.87 | 0.84 | 0.90 | 0.89 | 0.88 | 0.83 | 0.84 | 0.83 | 0.85 | 0.79 | 0.90 | 0.92 | 0.88 | 0.88 | 0.88 | 0.88 | 0.77 | 0.81 | 0.87 | |

**Table 3.3:** DSC values for the right femoral head obtained during the brute force experiment. White cells indicate DSC values from 0.8 to unity, blue from 0.60 to 0.79, and red below 0.6.

| Target Subjects \ Atlas Subjects | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | | 0.86 | 0.84 | 0.91 | 0.93 | 0.87 | 0.86 | 0.80 | 0.81 | 0.80 | 0.86 | 0.90 | 0.89 | 0.92 | 0.87 | 0.86 | 0.78 | 0.82 | 0.86 | 0.91 |
| 2 | 0.87 | | 0.77 | 0.87 | 0.78 | 0.80 | 0.68 | 0.76 | 0.80 | 0.70 | 0.70 | 0.80 | 0.90 | 0.88 | 0.77 | 0.71 | 0.71 | 0.67 | 0.77 | 0.79 |
| 3 | 0.86 | 0.76 | | 0.86 | 0.83 | 0.79 | 0.77 | 0.72 | 0.82 | 0.76 | 0.88 | 0.84 | 0.82 | 0.79 | 0.83 | 0.78 | 0.74 | 0.83 | 0.89 | 0.89 |
| 4 | 0.91 | 0.85 | 0.84 | | 0.90 | 0.84 | 0.84 | 0.81 | 0.85 | 0.83 | 0.87 | 0.88 | 0.88 | 0.88 | 0.86 | 0.83 | 0.77 | 0.81 | 0.87 | 0.88 |
| 5 | 0.93 | 0.81 | 0.84 | 0.90 | | 0.85 | 0.87 | 0.81 | 0.85 | 0.82 | 0.84 | 0.87 | 0.90 | 0.91 | 0.87 | 0.86 | 0.79 | 0.81 | 0.85 | 0.90 |
| 6 | 0.87 | 0.79 | 0.76 | 0.84 | 0.86 | | 0.85 | 0.79 | 0.84 | 0.70 | 0.76 | 0.83 | 0.82 | 0.86 | 0.82 | 0.87 | 0.87 | 0.81 | 0.79 | 0.84 |
| 7 | 0.84 | 0.65 | 0.76 | 0.86 | 0.87 | 0.86 | | 0.88 | 0.83 | 0.40 | 0.85 | 0.83 | 0.87 | 0.87 | 0.81 | 0.87 | 0.80 | 0.88 | 0.85 | 0.87 |
| 8 | 0.82 | 0.76 | 0.76 | 0.82 | 0.81 | 0.81 | 0.88 | | 0.86 | 0.35 | 0.81 | 0.63 | 0.79 | 0.80 | 0.75 | 0.85 | 0.86 | 0.92 | 0.62 | 0.78 |
| 9 | 0.81 | 0.81 | 0.81 | 0.86 | 0.84 | 0.89 | 0.86 | 0.86 | | 0.75 | 0.85 | 0.87 | 0.81 | 0.81 | 0.82 | 0.89 | 0.84 | 0.86 | 0.85 | 0.88 |
| 10 | 0.70 | 0.51 | 0.74 | 0.81 | 0.82 | 0.65 | 0.39 | 0.36 | 0.74 | | 0.48 | 0.39 | 0.74 | 0.80 | 0.77 | 0.42 | 0.50 | 0.42 | 0.75 | 0.77 |
| 11 | 0.87 | 0.71 | 0.89 | 0.90 | 0.84 | 0.79 | 0.85 | 0.81 | 0.83 | 0.39 | | 0.86 | 0.88 | 0.82 | 0.82 | 0.79 | 0.72 | 0.85 | 0.83 | 0.90 |
| 12 | 0.92 | 0.81 | 0.84 | 0.88 | 0.87 | 0.84 | 0.74 | 0.64 | 0.87 | 0.85 | 0.86 | | 0.88 | 0.90 | 0.86 | 0.84 | 0.79 | 0.83 | 0.81 | 0.90 |
| 13 | 0.89 | 0.89 | 0.84 | 0.89 | 0.90 | 0.82 | 0.87 | 0.85 | 0.82 | 0.81 | 0.89 | 0.88 | | 0.85 | 0.75 | 0.83 | 0.78 | 0.82 | 0.86 | 0.87 |
| 14 | 0.91 | 0.88 | 0.79 | 0.87 | 0.91 | 0.86 | 0.84 | 0.80 | 0.80 | 0.84 | 0.82 | 0.90 | 0.86 | | 0.85 | 0.84 | 0.79 | 0.77 | 0.82 | 0.87 |
| 15 | 0.88 | 0.75 | 0.81 | 0.86 | 0.87 | 0.83 | 0.83 | 0.52 | 0.81 | 0.80 | 0.75 | 0.84 | 0.64 | 0.86 | | 0.81 | 0.80 | 0.56 | 0.82 | 0.87 |
| 16 | 0.87 | 0.71 | 0.78 | 0.83 | 0.85 | 0.90 | 0.85 | 0.85 | 0.87 | 0.50 | 0.77 | 0.84 | 0.80 | 0.85 | 0.79 | | 0.87 | 0.81 | 0.79 | 0.84 |
| 17 | 0.80 | 0.70 | 0.73 | 0.78 | 0.79 | 0.87 | 0.82 | 0.85 | 0.85 | 0.64 | 0.76 | 0.80 | 0.80 | 0.80 | 0.81 | 0.84 | | 0.84 | 0.79 | 0.78 |
| 18 | 0.83 | 0.68 | 0.81 | 0.80 | 0.80 | 0.82 | 0.88 | 0.91 | 0.85 | 0.45 | 0.83 | 0.81 | 0.81 | 0.79 | 0.75 | 0.86 | 0.81 | | 0.83 | 0.83 |
| 19 | 0.88 | 0.72 | 0.88 | 0.89 | 0.87 | 0.82 | 0.81 | 0.80 | 0.84 | 0.78 | 0.83 | 0.84 | 0.86 | 0.80 | 0.80 | 0.81 | 0.69 | 0.78 | | 0.87 |
| 20 | 0.91 | 0.79 | 0.87 | 0.89 | 0.90 | 0.85 | 0.84 | 0.79 | 0.88 | 0.68 | 0.90 | 0.90 | 0.86 | 0.87 | 0.89 | 0.83 | 0.81 | 0.83 | 0.89 | |

**Table 3.4:** DSC values for the left femoral head obtained during the brute force experiment. White cells indicate DSC values from 0.8 to unity, blue from 0.60 to 0.79, and red below 0.6.

### 3.2.2 Correlations of Single Anatomical Measurements

The brute force experiment allowed the identification of best matched atlas subject and target subjects in terms of final contour DSC. The relationship in any single anatomical measurement between best subject pairs was assessed by computing the Pearson's product-moment correlation coefficient. For instance, the correlation coefficient for PL was found to be 0.42, as shown in Figure 3.2. Generally, the correlation coefficient varied among parameters with a maximum value of 0.448 achieved for the left femoral head diameter and a minimum value of 0.088 found for AP thickness. Correlation plots for all other anatomical measurements can be found in Appendix 4.

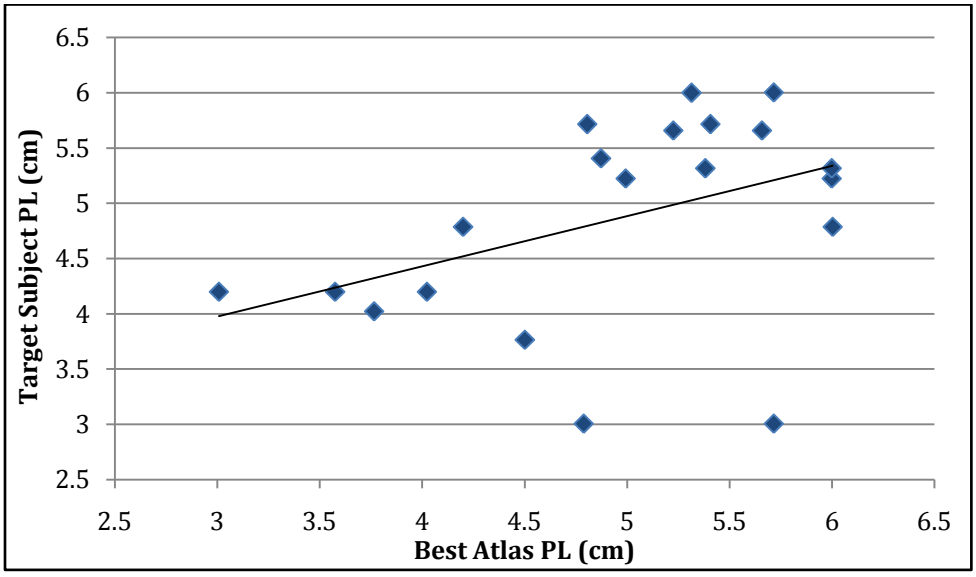



**Figure 3.2:** Correlation of target subject and best atlas subject PL for the 20 subject training dataset (r = 0.42).

It was not possible to identify a single anatomical parameter that was well correlated between a target and the best atlas subject. A low correlation coefficient does not mean parameters are completely independent but does indicate there is no linear relationship. Despite considering

other relationships, trends in anatomical measurements between the best image pairs were not identified, and this atlas selection approach was abandoned.

### 3.2.3. Correlation of Anatomical Measurement Combinations

Previously, a single anatomical parameter approach to atlas selection was shown to be ineffective. Accurate atlas selection may be achieved by using combinations of anatomical measurements. Therefore, an optimization procedure was performed to identify weighting factors in Equation 2.1 such that the final mean DSC for the training database was maximized. As described in Section 2.5.3, a total of 10,000 optimizations were performed to identify 10,000 sets of weights. Two final sets of weights were obtained for prostate: one set was chosen from the 52 sets that achieved a maximum mean DSC of 0.73 for the training dataset while giving the highest DSC for the validation dataset. The other set was chosen from the 300 sets of weights that achieved a DSC between 0.72 and 0.73 for the training set while giving the highest DSC for the validation dataset. For the femoral heads, the 10,000 weight sets were ranked according to DSC and the highest ranking set was chosen. The sets of selected final weights are shown in Table 3.5.

Higher weights for particular anatomical measurement indicate that that parameter was more important in identifying a good match in the atlas database. For the second set of prostate weights, the distances from the prostate to both femoral heads were the most important for selecting an atlas subject for prostate contouring. This may be because the position of the prostate in axial CT varies most in the AP direction (Kyenzeh *et al*., 2010). Lateral distance from the prostate to the femoral heads is less variable than the distance to the bladder or rectal

wall. The next most important parameter for prostate selection was PL. This is not a surprise because prostate size is variable among patients and PL describes this in one dimension. Furthermore, there is a direct relationship between PL and age (Zhang, 2012) so it is expected PL should be an important parameter in selecting a well matched atlas subject for prostate. In this work, $W_{CTV-RF}$ was higher than $W_{PL,}$ however, reducing the uncertainty in measuring PL might increase the dependency of atlas selection on $W_{PL}$.

| | ROI | | | |
|---|---|---|---|---|
| **Weight** | **Prostate (W1)** | **Prostate (W2)** | **RF** | **LF** |
| $\boldsymbol{W_{PL}}$ | 0.15 | 0.15 | 0.07 | 0.06 |
| $\boldsymbol{W_{RFHD}}$ | 0.06 | 0.13 | 0.10 | 0.25 |
| $\boldsymbol{W_{LFHD}}$ | 0.32 | 0.07 | 0.003 | 0.17 |
| $\boldsymbol{W_{AP}}$ | 0.13 | 0.07 | 0.16 | 0.12 |
| $\boldsymbol{W_{LR}}$ | 0.0016 | 0.10 | 0.15 | 0.01 |
| $\boldsymbol{W_{CTV-BW}}$ | 0.0031 | 0.01 | 0.14 | 0.06 |
| $\boldsymbol{W_{CTV-RW}}$ | 0.09 | 0.03 | 0.04 | 0.07 |
| $\boldsymbol{W_{CTV-RF}}$ | 0.13 | 0.27 | 0.02 | 0.14 |
| $\boldsymbol{W_{CTV-LF}}$ | 0.11 | 0.18 | 0.31 | 0.12 |

**Table 3.5**: Optimized weighting values for prostate and femoral heads obtained for each anatomical parameter. These weights were normalized to unity.

The distribution of importance was different in the two prostate weight sets demonstrating that the method was sensitive to the way in which the weight sets were selected. However, there was some agreement between the two sets. For example, $W_{PL}$ were identical while $W_{CTV-RW}$ and $W_{CTV-BW}$ values were low. It is not surprising that the distance from the prostate to the bladder or rectum does not help to position the prostate since those organs are quite variable across the population and not aligned following affine registration. It is believed that $W_{CTV-RW}$ and $W_{CTV-BW}$ would be more important for selecting atlas subjects for rectum and

bladder contouring. For clarity, the remaining discussion will refer to results generated with the second set of prostate weights.

Selection of an atlas using the proposed method was nearly instantaneous, excluding the time necessary for performing the anatomical measurements. The atlas selection method was applied to the training dataset in a leave-one-out scheme, where one subject was the target and the remaining 19 formed the atlas. This experiment tested the ability of the proposed selection method to reproduce the brute force result. For comparison, a random selection process was also performed where the atlas subject was selected randomly for each target.

The proposed algorithm had excellent ability to predict the most similar atlas subject for femoral heads, achieving a mean DSC of 0.87 ± 0.02 following affine registration. For prostate, a mean DSC of 0.72 ± 0.11 was obtained. The achieved DSCs scores using the proposed atlas selection for each target subject in the training set are reported in Table 3.6. The obtained results are encouraging since a DSC of > 0.7 was identified as acceptable following consideration of intra and inter-observer variability in the manual contouring process (Zijdenbos *et al*., 1994). The minimum DSC value obtained with the atlas selection method did decrease below the 0.7 threshold, especially for the prostate. This highlights the need for human surveillance during the automatic segmentation procedure.

| | DSC for | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Prostate | | | RF | | | LF | | |
| Subject ID | Brute force | Proposed | Random | Brute force | Proposed | Random | Brute force | Proposed | Random |
| 1 | 0.84 | 0.76 | 0.61 | 0.91 | 0.86 | 0.85 | 0.93 | 0.92 | 0.86 |
| 2 | 0.56 | 0.34 | 0.42 | 0.89 | 0.89 | 0.81 | 0.90 | 0.90 | 0.88 |
| 3 | 0.85 | 0.81 | 0.50 | 0.91 | 0.91 | 0.82 | 0.89 | 0.86 | 0.82 |
| 4 | 0.84 | 0.83 | 0.55 | 0.91 | 0.88 | 0.86 | 0.91 | 0.91 | 0.85 |
| 5 | 0.76 | 0.72 | 0.57 | 0.92 | 0.86 | 0.83 | 0.93 | 0.93 | 0.85 |
| 6 | 0.71 | 0.57 | 0.45 | 0.88 | 0.83 | 0.81 | 0.87 | 0.84 | 0.86 |
| 7 | 0.76 | 0.76 | 0.50 | 0.83 | 0.83 | 0.83 | 0.88 | 0.88 | 0.81 |
| 8 | 0.72 | 0.72 | 0.47 | 0.90 | 0.87 | 0.88 | 0.92 | 0.88 | 0.76 |
| 9 | 0.85 | 0.85 | 0.56 | 0.91 | 0.88 | 0.77 | 0.89 | 0.86 | 0.84 |
| 10 | 0.84 | 0.74 | 0.66 | 0.89 | 0.89 | 0.83 | 0.82 | 0.82 | 0.82 |
| 11 | 0.84 | 0.74 | 0.64 | 0.91 | 0.90 | 0.85 | 0.90 | 0.90 | 0.82 |
| 12 | 0.78 | 0.76 | 0.58 | 0.92 | 0.91 | 0.83 | 0.92 | 0.87 | 0.85 |
| 13 | 0.80 | 0.55 | 0.55 | 0.91 | 0.90 | 0.83 | 0.90 | 0.86 | 0.82 |
| 14 | 0.76 | 0.76 | 0.61 | 0.92 | 0.85 | 0.85 | 0.91 | 0.91 | 0.82 |
| 15 | 0.77 | 0.77 | 0.58 | 0.90 | 0.85 | 0.82 | 0.88 | 0.82 | 0.86 |
| 16 | 0.85 | 0.68 | 0.53 | 0.92 | 0.90 | 0.84 | 0.90 | 0.85 | 0.87 |
| 17 | 0.77 | 0.70 | 0.55 | 0.89 | 0.82 | 0.77 | 0.87 | 0.85 | 0.82 |
| 18 | 0.83 | 0.72 | 0.60 | 0.87 | 0.85 | 0.81 | 0.91 | 0.83 | 0.83 |
| 19 | 0.81 | 0.81 | 0.61 | 0.88 | 0.86 | 0.85 | 0.89 | 0.86 | 0.86 |
| 20 | 0.72 | 0.57 | 0.43 | 0.92 | 0.90 | 0.81 | 0.91 | 0.90 | 0.88 |
| Mean | 0.78 | 0.72 | 0.55 | 0.90 | 0.87 | 0.83 | 0.90 | 0.87 | 0.83 |
| ± SD | ±0.07 | ±0.11 | ±0.07 | ±0.02 | ±0.03 | ±0.02 | ±0.03 | ±0.03 | ±0.02 |

**Table 3.6**: DSC between contours obtained using the proposed ABS, brute force, and random selection compared to the manual contour for all 20 subjects in the training data set.

The mean DSCs for the proposed atlas selection were slightly lower than the values established using brute force. Small decreases in performance may be acceptable given that the current evaluation does not include the potential benefit of deformable image registration. This analysis shows that the proposed atlas selection method is able to nearly reproduce the results obtained with brute force for the same subjects. Random atlas selection achieved DSCs of 0.55±0.07 and 0.83±0.02 for the prostate and femoral heads respectively. Thus, the mean DSC for the proposed method demonstrated significantly higher segmentation accuracy compared

with random atlas selection. The mean accuracy values across all subjects for three methods are shown in Figure 3.3. Analysis of variance (ANOVA) on DSCs between the selection methods was performed to test for statistical significance. The result showed there is statically significantly difference between the mean DSCs for all contours. A Tukey's honestly significant difference (HSD) test correction indicates that the three methods are significantly different from each other. Among the three strategies, brute force was the most accurate, followed by the proposed method, then random selection.

In summary, the proposed method demonstrated a modest reduction in prostate accuracy compared to the brute force method and a minimal loss for both femoral heads. The optimized method nearly reproduced the brute force result without requiring the target images to be contoured earlier.

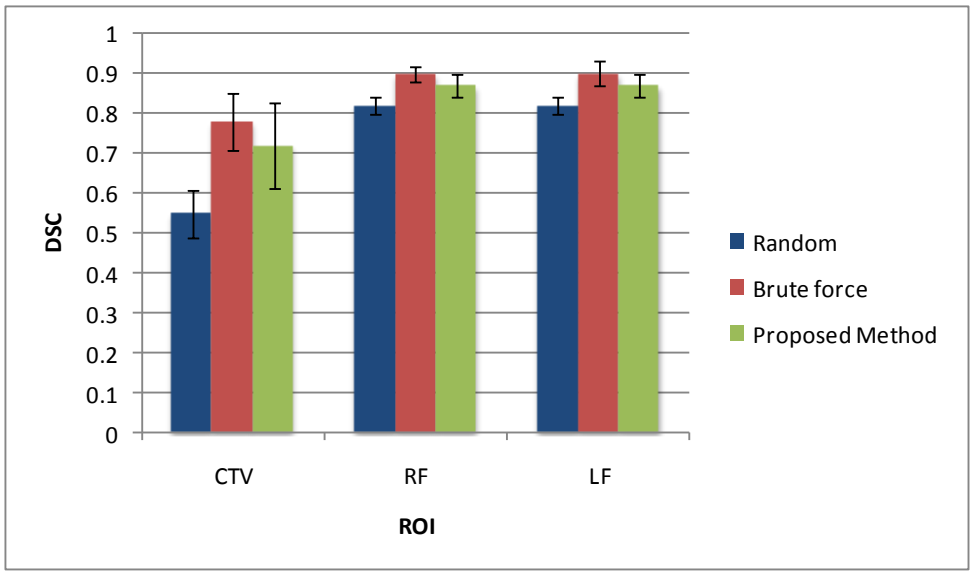



**Figure 3.3:** Mean DSC± standard deviation obtained for the proposed method, brute force, and random atlas selection. CTV includes the prostate gland only

## 3.3. Validation

To test the proposed atlas selection method, 10 new subjects were randomly selected for inclusion into a validation dataset. For these subjects, the measured anatomical characteristics are summarized in Table 3.7.

| Parameter | Mean ± SD (cm) | Parameter | Mean ± SD (cm) |
|---|---|---|---|
| **PL** | 4.65 ± 0.92 | **prostate - BW distance** | 4.43 ± 0.66 |
| **RFHD** | 4.57 ± 0.38 | **prostate - RW distance** | 3.88 ± 0.42 |
| **LFHD** | 4.62 ±0.43 | **prostate - RF distance** | 11.45 ± 0.84 |
| **AP thickness** | 22.70 ±1.98 | **prostate - LF distance** | 11.14 ± 0.75 |
| **LR thickness** | 36.36 ±1.74 | | |

**Table 3.7:** Means and standard deviation (SD) for several anatomical measured characteristics for 10 subjects used to validate the atlas selection method.

It appeared that the there was no difference in the measured anatomical parameters between the training and validation datasets, suggesting that the training dataset may already include much of the variability associated with patients' geometry. The test subjects were added to the brute force procedure to first assess the possibility of increasing accuracy by increasing the number of atlas subjects. Encompassing more anatomical variability with a larger atlas database increases computational time. A comparison of mean DSCs across all subjects obtained from brute force on the 20 and 30 subject sets is shown in Table 3.8. These results suggest that increasing the number of atlas subjects will not have a large impact on the final DSC. This may indicate that

the current values are the maximum achievable considering the fact that the gold standard segmentations were obtained manually.

| Number of Subjects | DSC for (Mean ± SD) Prostate | RF | LF |
|---|---|---|---|
| **20** | 0.78±0.05 | 0.90±0.02 | 0.90±0.03 |
| **30** | 0.80±0.05 | 0.91± 0.03 | 0.90±0.02 |

**Table 3.8:** The mean and standard deviation of DSCs achieved using the brute force procedure.

As discussed previously, two sets of weights for the prostate were obtained for the proposed atlas selection algorithm. The first set was the one that achieved the highest DSC for the validation set (0.60) from 52 sets that achieved the highest DSC on the training data (0.73). The second set was the one that achieved the highest DSC for the validation set (0.64) from 300 sets that achieved the highest DSC on the training data (0.72 to 0.73). The remainder of the discussion will focus on atlas subject selection using the second set of weights. However, it must be noted that selecting the optimal set of weights was not trivial. This is because the DSCs obtained on the training set are not exact due to reliance on manual segmentation as the gold standard. Furthermore, many (300) optimizations resulted in high DSCs, complicating the process of identifying only one final set of parameters.

Comparison of DSCs between manual and ABS for the test subjects using the anatomical selection method is summarized in Figure 3.4. High DSC values were achieved for femoral heads, which are clearly visible structures in CT images. The mean DSCs ± SD were 0.86 ± 0.03 for either femoral head. High bone contrast in CT imaging improves affine image registration.

Furthermore, bony anatomy tends to be similar across patients so atlas selection is not as important. The mean DSC for the prostate was 0.64 ± 0.09. This is a respectable result considering that no deformable image registration was employed to compensate for variability in position and shape due to patient features and variability in bladder and rectum filling. Random atlas selection achieved DSCs of 0.55±0.03 and 0.86±0.03 for the prostate and either femoral head, respectively. ANOVA was performed and identified the differences between the mean DSCs for all contours for the three-atlas selection methods were significant. The proposed selection results compare well with previous ABS work in prostate radiotherapy. For example, using a multi-subject atlas selection approach, Pirozzi *et al*. (2012) obtained mean DSCs of 0.64, 0.82, and 0.84 for the prostate, right femoral head, and left femoral head respectively. The proposed approach provides the same or better DSCs without the advantage of deformable image registration in the ABS pipeline

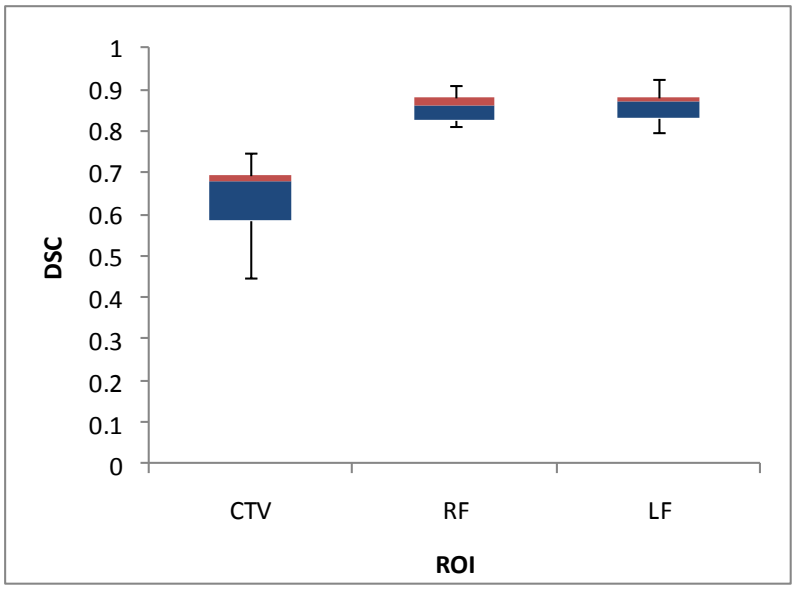



**Figure 3.4:** Box and whisker plots for prostate and femoral head auto contouring. The figure shows the quartiles and the median. The maximum and minimum are the ends of the whisker. CTV includes the prostate gland only.

Table 3.9 illustrates the individual auto contouring results compared to the manual gold standard. For the prostate contours the Hausdorff distance was also computed to provide an indication of the absolute contouring error as seen in Table 3.10. As discussed before, a DSC of 0.7 or higher may be considered acceptable. This criterion is met for all patients for femoral head contours. Only 40 % of the patients had prostate contours with sufficient DSC

| | DSC for | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | RF | | | LF | | | Prostate | | |
| Subject ID | Brute force | Proposed | Random | Brute force | Proposed | Random | Brute force | Proposed | Random |
| **1** | 0.90 | 0.88 | 0.79 | 0.88 | 0.88 | 0.83 | 0.74 | 0.45 | 0.52 |
| **2** | 0.90 | 0.83 | 0.89 | 0.91 | 0.88 | 0.9 | 0.76 | 0.64 | 0.61 |
| **3** | 0.91 | 0.88 | 0.84 | 0.88 | 0.87 | 0.88 | 0.72 | 0.70 | 0.51 |
| **4** | 0.92 | 0.85 | 0.85 | 0.90 | 0.86 | 0.81 | 0.81 | 0.75 | 0.57 |
| **5** | 0.91 | 0.81 | 0.83 | 0.91 | 0.88 | 0.83 | 0.74 | 0.57 | 0.58 |
| **6** | 0.91 | 0.91 | 0.85 | 0.93 | 0.93 | 0.85 | 0.78 | 0.68 | 0.59 |
| **7** | 0.88 | 0.82 | 0.8 | 0.83 | 0.80 | 0.83 | 0.69 | 0.68 | 0.55 |
| **8** | 0.89 | 0.87 | 0.83 | 0.85 | 0.80 | 0.84 | 0.76 | 0.70 | 0.53 |
| **9** | 0.91 | 0.82 | 0.78 | 0.85 | 0.84 | 0.77 | 0.80 | 0.53 | 0.50 |
| **10** | 0.88 | 0.88 | 0.81 | 0.90 | 0.83 | 0.81 | 0.78 | 0.72 | 0.57 |
| **Mean** | 0.90 | 0.86 | 0.83 | 0.88 | 0.86 | 0.83 | 0.76 | 0.64 | 0.55 |
| **± SD** | ±0.01 | ±0.03 | ±0.03 | ± 0.03 | ±0.04 | ±0.04 | ±0.03 | ±0.09 | ±0.03 |

**Table 3.9:** DSC between contours obtained using the proposed ABS pipeline or brute force and random selection compared to the manual contour for the 10-subject validation dataset.

| Subject ID | Hausdrorff distance - prostate (mm) |
|---|---|
| **1** | 25.7 |
| **2** | 14.2 |
| **3** | 10.7 |
| **4** | 11.1 |
| **5** | 15.7 |
| **6** | 18.7 |
| **7** | 10.4 |
| **8** | 9.6 |
| **9** | 13.2 |
| **10** | 10.6 |
| **Mean** | 14.00 |
| **± SD** | ±5.0 |

**Table 3.10:** The Hausdorff distance obtained for the prostate using the proposed ABS compared to the manual contour for the 10-subject validation dataset.

The best and worst prostrate segmentation results were obtained for subject 4 and 1 with DSCs of 0.75 and 0.45, respectively. The subject with the lowest DSC had a Hausdorff distance of 25.7 mm to the manual contour, while the distance for the subject with the highest DSC was 11.1 mm. This distance may be acceptable giving that deformable registration may be able to correct this difference as the capture range of deformable registration in pelvic CT is on the order of 30 mm (Wierzbicki *et al*., 2010). Moreover, manual segmentation for the prostate is associated with a relatively large inter observer variability so smaller difference between two contours would be acceptable. The manual and auto segmentation for best and the worst result as measured by DSC are shown in Figure 3.5. For the worst case prostate, the auto segmentation performance was poor due to the anterior position of the prostate with respect to the bony anatomy for this patient. For femoral heads, both best and worst cases demonstrate minimal difference from the manual contour.

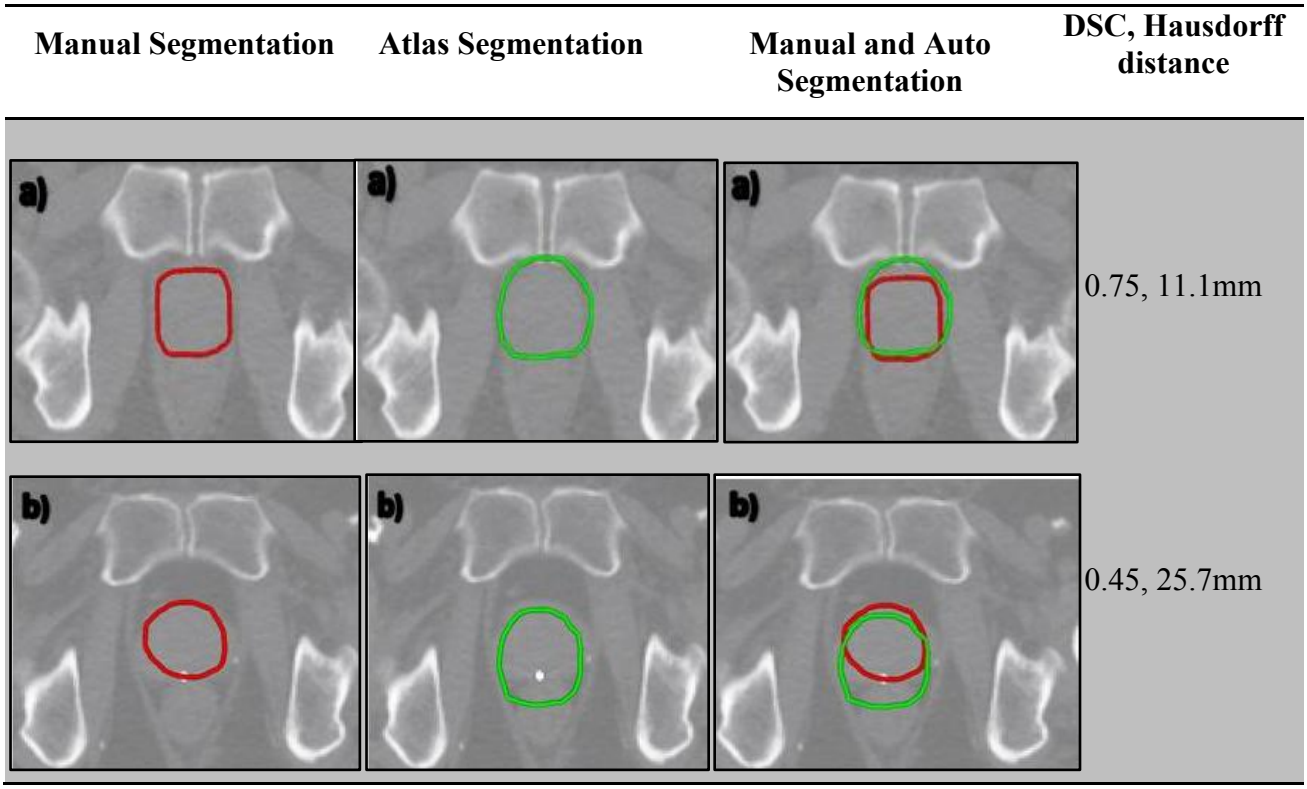



**Figure 3.5:** Axial displays for the best (a) and worst (b) cases of prostate segmentation evaluated using DSC and Hausdorff distance. Green curves represent the automated segmentation while red curves represent the manual segmentation.

Different approaches have been reported in literature for atlas selection in ABS as described in Section 1.5.1. In contrast with the previous methods, selection in this work was based on anatomical characteristics. This made it possible to use a large atlas data set, which was previously unfeasible due to the large computation time required. In general this approach is able to provide accurate selection for femoral head contours with reasonable accuracy for the prostate.

Chapter 4

# Conclusion

This study tested an atlas selection strategy that combined image registration with anatomical measurement matching to segment CT images for radiotherapy treatment planning. The technique provided contours of the prostate and both femoral heads that are similar to those drawn manually, but with significantly less user interaction.

Selection of an atlas subject for a new subject using the proposed method begins with the measurement of several anatomical characteristics in the CT image. These values are entered into the optimized atlas selection algorithm and a target-atlas subject cost value is computed for each subject in the atlas. The cost value is a weighted sum of differences between the anatomical measurements performed in the target and atlas subject images. Finally, the atlas subject that gives the lowest cost value is selected for ABS. This process is efficient, taking less than 5 minutes for any new CT image, including the time it takes to make the anatomical measurements.

The proposed most similar atlas selection approach based on anatomical characteristics was tested for the prostate and both femoral heads. Validation was initially performed using a 20-subject atlas training dataset with manual delineation taken as the gold standard. For this data, a brute force approach provided mean DSCs of 0.78 ± 0.11 and 0.90 ± 0.02 for the prostate and either femoral head, respectively. These results indicate that selecting the best matching atlas subject for a given target achieves a similar level of segmentation accuracy as manual re-contouring (Hwee *et al.*, 2011), indicating that if the atlas selection problem is solved, ABS may replace manual contouring in radiotherapy. After optimizing the proposed atlas selection process on the same training dataset, DSCs of 0.72 ± 0.11 and 0.87 ± 0.03 for the prostate and both femoral heads were achieved. This demonstrated that the proposed method is capable of nearly reproducing the brute force results with the same set of data. The mean DSC for the proposed method indicated significantly higher segmentation accuracy compared with random atlas selection, which achieved DSCs of 0.55±0.07 and 0.83±0.02 for the prostate and femoral heads, respectively. ANOVA showed significant difference on the mean DSCs the three atlas selection methods. Finally, testing the optimized atlas selection method on a new, 10 subject validation set yielded DSCs of 0.64 ± 0.09 and 0.86 ± 0.03 for the prostate and either femoral head, respectively. This represented a significant loss of accuracy from 0.76 ± 0.03, 0.90 ± 0.01, and 0.88 ± 0.03 for prostate, right femoral head and left femoral head, respectively, achieved for the validation set using brute force. However, significantly better results were obtained compared to randomly selecting with a DSC of 0.55±0.03 and 0.86±0.03 for the prostate and femoral heads respectively. Statistically significant difference in ANOVA was found between the mean DSCs for all contours obtained from the three-atlas selection method. Furthermore, similar results of 0.65 for prostate, 0.812 for right femoral head, and 0.834 for left femoral head were achieved in

previous ABS work (Pirozzi *et al.*, 2012). It is important to note that the impact of deformable registration was not included in this work whereas Pirozzi *et al*. did have this benefit. Thus, the current method may be considered successful.

Despite that the work has reached its aim, there were unavoidable limitations. First, inter and intra observer variability is the most significant contributor to error in ABS results, particularly for prostate cancer radiotherapy due to large variation in pelvic region causing uncertainty in the manual segmentation (gold standard). A poor level of segmentation accuracy was also achieved for the prostate due to organ variation by Acosta *et al.* (2010) with a DSC of 0.583. With this finding, it is important to continue address the variability challenge. In this work, manual segmentation was used as the gold standard during training, resulting in a variety of weights sets that were suitable candidates for the proposed method and complicating the process of selecting the best weights to use. Similarly, validation of the final segmentation relied on manual segmentation as the gold standard complicating the process of qualifying the accuracy of the automated segmentation result. Furthermore, due to variability in weights the optimization algorithm was repeated several times to insure the optimal set of weights has been selected.

Due to time limitations, this study did not gain the potential advantage of deformable registration. Affine registration was able to align bones, but this was not sufficient for aligning the motion accruing in pelvic region due to rectum and bladder filling. This also forced the removal of rectum and bladder from further analysis due to the large variability in size, shape, and location of these organs. However deformable registration is able to correctly align the prostate by masking bladder and rectum, so the registration algorithm will concentrate more on

registering their large discrepancies, which lead to correctly registering the prostate (Godly *et al.*, 2009). Finally, this work did not analyze the impact of errors in the manual measurements performed on CT images. These measurements would be made by placing points near the centroids of the prostate, rectum, bladder, and femoral heads. These points can be used to estimate most of the measurements. This must be obtained accurately since selection of the most similar atlas subject is based on these measurements. Given that PL and femoral heads diameter were obtained manually, there is an error is assisted with these measurement.

There are several potential strategies for improving the proposed method. For example, including additional subjects in the analysis would alleviate the effect of uncertainties in the manually draw contours used as the gold standards. This would improve the outcome of the weight optimization, consequently reducing the final ABS error quantification since the optimized atlas selection method depends on the anatomical parameters weights. It should be noted that additional subjects would probably not improve final accuracy since brute force results obtained with 20 and 30 subjects showed similar results.

Including deformable registration in the process will also improve contouring accuracy (Godly *et al.*, 2009). Moreover, it would allow comparison with different, non-rigid registration-based, atlas selection techniques. Deformable registration is necessary for contouring the rectum and bladder. This would require the investigation of other anatomical characteristics that are correlated to bladder and rectum such as their volumes along with testing the ability of distance between the prostate and rectum and bladder to predict rectum and bladder contouring. Although deformable registration would improve final contouring accuracy, accurate delineation

of the rectum and bladder will probably require a significant amount of user input. The impact of errors in the anatomical measurements on the final contouring accuracy can be assessed by repeating the proposed atlas selection method with the manual anatomical parameters measured different times. The difference in the final contouring accuracy can be easily evaluated. Furthermore, the optimized method is focused on single atlas per contour. The method may be extended by testing multiple atlas subject segmentation techniques where several estimates of the same organ are obtained for a target image. This has been shown to improve contouring accuracy previously but does increase computational time (Aljabar *et al*., 2009).

Another future direction will involve treatment planning of a dose distribution with the ABS contours and a comparison with the dose distribution obtained clinically using manually drawn contours. If the dose distributions are clinically comparable, the ABS errors might be acceptable from the radiotherapy treatment planning perspective. It is reasonable to assume this is the case because the prostate (CTV) is expanded to the PTV for treatment planning; an averaging process that may alleviate minor contouring errors. Furthermore, techniques like VMAT may not have the ability to control the dose distribution at the same spatial resolution as the potential contouring errors expected in ABS.

ABS in the pelvic region has previously demonstrated the potential to improve efficiency and reduce variability associated with manual segmentation (Young *et al*., 2011). This work presented an atlas selection technique based on matching various anatomical measurements. Simultaneously, the method was validated for a set of 10 patient images. The approach demonstrated great potential for selecting the best atlas for CT segmentation in radiotherapy

treatment planning. The method achieved high accuracy for femoral heads with reasonable accuracy obtained for the prostate in comparison to other results found in the literature. The proposed method may extended to process further patient data, contour additional anatomy, and even be extended to different areas of the body. In conclusion, this work demonstrates a promising approach for delineating the prostate and femoral heads based only on limited knowledge of the subject anatomy. The proposed atlas selection method could help increase consistency between different centers as well as increase the efficiency of the contorting process.

# Bibliography

Aarts, E., & Lenstra, J. K. (1997). Local Search in Combinatorial Optimization. John Wiley & Sons, Inc.

Abdel-Wahab, M., Mahmoud, O., Merrick, G., Hsu, I. C. J., Arterbery, V. E., Ciezki, J. P., Yamada, Y. (2012). ACR Appropriateness Criteria external-beam radiation therapy treatment planning for clinically localized prostate cancer. Journal of the American College of Radiology : JACR, 9(4), 233–8.

Acosta, O., Jason, D., Deran, G., Antoine, S., De Crevoisier, R., *et al*. (2010). Atlas Based Segmentation and Mapping of Organs at Risk from Planning CT for the Development of Voxel-Wise Predictive Models of Toxicity in Prostate Radiotherapy. Prostate Cancer Imaging, 42–51.

Acosta, O., Jason, D., Deran, G., Antoine, S., De Crevoisier, R., *et al* (2014). Multi-atlas-based segmentation of pelvic structures from CT scans for planning in prostate cancer. Abdomen and thoracic imaging, pp 623–656.

Albert, K., Hoang, D., Modat, M., Kelvin, K., M. Jorge, C, Josephine, B, Timor, K, Sébastien, O. (2013). Using manifold learning for atlas selection in multi-atlas segmentation. PLOS One, 8(8), e70059.

Aljabar, P., Heckemann, R. A., Hammers, A., Hajnal, J. V., & Rueckert, D. (2009). Multi-atlas based segmentation of brain images: atlas selection and its effect on accuracy. Neuroimage, 46(3), 726-738.

Bethesda. (1999) ICRU Report 62 International Commission on Radiation Units and Measurements.

Cootes, T. F., Hill, A., Taylor, C. J., & Haslam, J. (1994). The Use of Active Shape Models For Locating Structures in Medical Images. Image and Vision Computing, 12(6), 355–366.

Canadian Cancer Society's Steering Committee. (2014). Canadian cancer statistic 2014. Canadian Cancer Society, ISSN 0835-2976.

Crawford, E. D, Flaig, T. W. (2012). Optimizing outcomes of advanced prostate cancer: drug sequencing and novel therapeutic approaches. Oncology (Williston Park, N. Y.) 6(1), 70 -7.

De Meerleer, G., Vakaet, L., Meersschout, S., Villeirs, G., Verbaeys, A., Oosterlinck, W., & De Neve, W. (2004). Intensity-modulated radiotherapy as primary treatment for prostate cancer: acute toxicity in 114 patients. International Journal of Radiation Oncology, Biology, Physics, 60(3), 777-787.

Dowling, J., Fripp, J., Freer, P., Ourselin, S., & Salvado, O. (2009). Automatic atlas-based segmentation of the prostate: A MICCAI 2009 Prostate Segmentation Challenge entry. Workshop in Medical Image Computer Assist, 17-24.

Elhaddad, y. (2012). Combined Simulated Annealing and Genetic Algorithm to Solve Optimization Problems. World Academy of Science, Engineering and Technology, 6(8), 1270–1272.

Elmohamed, M. S., Coddington, P., & Fox, G. (1998). A comparison of annealing techniques for academic course scheduling. In Practice and Theory of Automated Timetabling II (pp. 92-112). Springer Berlin Heidelberg.

Fiorino, C., Reni, M., Bolognesi, A., Cattaneo, G. M., & Calandrino, R. (1998). Intra-and inter-observer variability in contouring prostate and seminal vesicles: implications for conformal treatment planning. Radiotherapy and oncology, 47(3), 285-292.

Godley, A., Ahunbay, E., Peng, C., & Li, X. A. (2009). Automated registration of large deformations for adaptive radiation therapy of prostate cancer. Medical physics, 36(4), 1433-1441.

Hill, D. L., Batchelor, P. G., Holden, M., & Hawkes, D. J. (2001). Medical image registration. Physics in medicine and biology, 46(3), R1.

Huttenlocher, D. P., Klanderman, G. A., & Rucklidge, W. J. (1993). Comparing images using the Hausdorff distance. Pattern Analysis and Machine Intelligence, IEEE Transactions on Communications, 15(9), 850-863.

Hwee, J., Louie, A. V., Gaede, S., Bauman, G., D'Souza, D., Sexton, *et al* (2011). Technology assessment of automated atlas based segmentation in prostate bed contouring. Radiation Oncology, doi: 10.1186/1748-717X-6-110.

Jameson, M. G., Holloway, L. C., Vial, P. J., Vinod, S. K., & Metcalfe, P. E. (2010). A review of methods of analysis in contouring studies for radiation oncology. Journal of medical imaging and radiation oncology, 54(5), 401-410.

Jukka, K. (1999). A brief comparison of simulated annealing and genetic algorithm approaches. Department of Computer Science, University of Helsinki.

Johnson, R. A., & Wichern, D. W. (1992). Applied multivariate statistical analysis (Vol. 4). Englewood Cliffs, NJ: Prentice hall.

Kittler, J., & Alkoot, F. M. (2003). Sum versus vote fusion in multiple classifier systems. Pattern Analysis and Machine Intelligence, IEEE Transactions on Communications, 25(1), 110-115.

Lee, D. (1945). Measures of the Amount of Ecologic Association Between Species. Ecology, 26(3), 297–302.

Lin, A, Kubicek, G, Piper, J.W, Nelson, A.S, Dicker, *et al.* (2008). Atlas-Based Segmentation in Prostate IMRT : time - savings in the clinical workflow, International Journal of Radiation Oncology, Biology, Physics, 72(1), S328–S329.

Livsey, JE., Wylie, JP., Swindell, R., Khoo, VS., Cowan, RA., Logue JP. (2004). Do differences in target volume definition in prostate cancer lead to clinically relevant differences in normal tissue toxicity? . International Journal of Radiation Oncology, Biology, Physics. 60(4), 1076–81.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. The journal of chemical physics. 21(6), 1087-1092.

Mitchell, DM., Perry, L., Smith, S., Elliott T., Wylie, JP., *et al* (2009). Assessing the effect of a contouring protocol on post prostatectomy radiotherapy clinical target volumes and inter physician variation. International Journal of Radiation Oncology, Biology, Physics, 75(4), 990–3.

Nelder, J. A., & Mead, R. (1965). A Simplex Method for Function Minimization. The Computer Journal, 7(4), 308–313.

Nishioka, K., Shimizu, S., Kinoshita, R., Inoue, T., Onodera, S., *et al* (2013) Evaluation of inter-observer variability of bladder boundary delineation on cone-beam CT. Radiation Oncology, 8(1), 185

Park, M., & Kimto, Y. (1998) A Systematic Procedure for Setting Parameters in Simulated Annealing Algorithm. Pattern Recognition, 25(3) pp.207-217.

Patel, R. B., Bryan, T., Kaminsky, D., Pirozzi, S., Nelson, A., Piper, J., ... & Ellis, R. J. (2014). Evaluation of an Atlas-Based Segmentation Method for High-Risk Prostate Cancer With RTOG-Defined Pelvic Lymph Node Levels. International Journal of Radiation Oncology, Biology, Physics, 90(1), S74-S75.

Pirozzi, SD., Horvat, M., Nelson, AS., Piper, JW. (2012). Atlas-based segmentation : evaluation of a multi-atlas approach for prostate cancer. ASTRO annual meeting, 2012, 81, 3677.

Rohlfing, T., Brandt, R., Menzel, R., Russakoff, D. B., & Maurer Jr., C. R. (2005). Quo vadis, atlas-based segmentation?. In Handbook of Biomedical Image Analysis (pp. 435-486). Springer US.

Sanroma, G., Guorong, W., Yaozong, G., & Dinggang, S. (2014). Learning to Rank Atlases for Multiple Segmentation, 33(10), 1939–1953.

Schuff, N., Woerner, N., Boreta, L., Kornfield, T., Shaw, L. M., *et al*. (2009). MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. Brain, 132(4), 1067-1077.

Sheets, NC, Goldin, GH, Meyer, A, Wu, Y, Chang Y, *et al*. (2012). Intensity-Modulated Radiation Therapy, Proton Therapy, or Conformal Radiation Therapy and Morbidity and Disease Control in Localized Prostate Cancer. The Journal of the American Medical Association. 18:307(15): 1611-20. doi:10.1001/jama.2012.460.

Snell, R. S. (2011). Clinical anatomy by regions. Lippincott Williams & Wilkins.

Strassmann, G., Abdellaoui, S., Richter, D., Bekkaoui, F., Haderlein, M., Fokas, E., Engenhart-Cabillic, R. (2010). Atlas-based semiautomatic target volume definition (PROSTATE) for head-and-neck tumors. International Journal of Radiation Oncology, Biology, Physics, 78(4), 1270 – 6.

Studholme, C., Hill, DL, Hawkes, DJ. (1996). Automated 3-D registration of MR and CT images of the head. Medical Image Analysis, 1(2), 163–175.

Studholme, C., Hill, DL, Hawkes, DJ. (1999). An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition, 32, 71–86.

Sylvester, J. E., Grimm, P. D., Wong, J., Galbreath, R. W., Merrick, G., Blasko, J. C. (2011). Fifteen-year biochemical relapse-free survival, cause-specific survival, and overall survival following I (125) prostate brachytherapy in clinically localized prostate cancer: Seattle experience. International Journal of Radiation Oncology, Biology, Physics, 81(2), 376–81.

Velker, V. M., Rodrigues, G. B., Dinniwell, R., Hwee, J., & Louie, A. V. (2013). Creation of RTOG compliant patient CT-atlases for automated atlas based contouring of local regional breast and high-risk prostate cancers. Radiation Oncology, 8(1), 188.

Vicini, F. A., Abner, A., Baglan, K. L., Kestin, L. L., & Martinez, A. A. (2001). Defining a dose–response relationship with radiotherapy for prostate cancer: is more really better? International Journal of Radiation Oncology, Biology, Physics, 51(5), 1200–1208.

Wierzbicki, M., Schaly, B., Peters, T., & Barnett, R. (2010). Automatic image guidance for prostate IMRT using low dose CBCT. Medical physics, 37(7), 3677-3686.

Wierzbicki, M. Drangova, G. Guiraudon, and T. M. P. (2004). Mapping template heart models to patient data using image registration. Medical Image Computing and Computer-Assisted Intervention, 87, 671–678.

Young, A. V, Wortham, A., Wernick, I., Evans, A., & Ennis, R. D. (2011). Atlas-based segmentation improves consistency and decreases time required for contouring postoperative endometrial cancer nodal volumes. International Journal of Radiation Oncology, Biology, Physics, 79(3), 943 –7.

Yu, C. X., Tang, G. (2011). Intensity-modulated arc therapy: principles, technologies and clinical implementation. Physics in Medicine and Biology, 56 (5), R31–54.

Zambrano, V., Furtado, H., Fabri, D., LÜtgendorf-Caucig, C., GÓra, J., Stock, M., Georg, D. (2013). Performance validation of deformable image registration in the pelvic region. Journal of Radiation Research, 54 (Suppl 1), i120–i128. doi:10.1093/jrr/rrt045.

Zelefsky, M., Leibel, S., Gaudin, P., Kutcher, G., Fleshner, *et al*. (1998) Dose escalation with three-dimensional conformal radiation therapy affects the outcome in prostate cancer. International Journal of Radiation Oncology, Biology, Physics, 41(3), 491-500.

Zhang, S-J., Qian, H-N., Zhao, Y., Sun, K., Wang, H-Q.*, et al.* (2013). Relationship between age and prostate size. Asian Journal of Andrology, 15(1), 116–120. doi:10.1038/aja.2012.127

Zijdenbos, A. P., Dawant, B. M., Margolin, R. A., & Palmer, A. C. (1994). Morphometric analysis of white matter lesions in MR images: method and validation. Medical Imaging IEEE Transactions on Communications 13(4) 716-724.

# Appendix 1

This section contains a detailed list of the scripts used for the image registration and image preparation process.

| Scripts | Action |
|---|---|
| **make_prostate_Mask.py** | Generates the mask that limits the area of the images considered during image registration. |
| **resample_Images.py** | Resamples two images to a common voxel size. |
| **ImageGlobalRegistration.py** | Registers two images using a combination of rotations, translation, and scaling. |
| **apply_XFM_to_Image.py** | Applies a transformation to an image. |
| **fill_Edge_Images.py** | Fills region of interest boundaries automatically with certain intensity. |
| **add_Image.py** | Sum the intensities for two images into single image. |
| **binarize_Image.py** | Assigns fixed intensity values to an image corresponding to region of interest. |
| **convert_ROI_to_Edge_Image.py** | Converts a Pinnacle ROI structure into the boundary of the contour. |
| **calaculate_Dice_coefficient.py** | Calculate the image similarity for the intensity value in an image. |
| **convert_ROI_to_poly_Image.py** | Convert Pinnacle ROI into visualization toolkit (VTK) mesh files. |
| **atlas_selector.py** | Select the atlas for each new patient |

**Table A1.1.** Description of the Scripts.

# Appendix 2

Description of simulating annealing optimization code

**Optimize the weights**

```
import scipy.optimize


Cost function = ((w[PL]              * abs(1 - [j][PL]              /  [i][PL]))              +
                 (w[RFHD]            * abs(1 - [j][RFHD]            /  [i][RFHD]))            +
                 (w[LFHD]            * abs(1 - [j][LFHD]            /  [i][LFHD]))            +
                 (w[AP]              * abs(1 - [j][AP]              /  [i][AP]))              +
                 (w[LR]              * abs(1 - [j][LR]              /  [i][LR]))              +
                 (w[prostate-BW]     * abs(1 - [j][prostate-BW]     /  [i][prostate-BW]))     +
                 (w[prostate-RW]     * abs(1 - [j][prostate-RW]     /  [i][prostate-RW]))     +
                 (w[prostate-RF ]    * abs(1 - [j][prostate-RF]     /  [i][prostate-RF]))     +
                 (w[prostate-LF]     * abs(1 - [j][prostate-LF]     /  [i][prostate-LF)])      )

 j = Atlas subject,  i = Target Subject

x0 = [1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0,1.0]

res = scipy.optimize.anneal (Cost Function,
                        x0,
                        args=(),
                        schedule='fast',
                        T0=None,
                        Tf=9.9999999999999998e-13,
                        maxeval=None,
                        maxaccept=None,
                        maxiter=400,
                        boltzmann=1.0,
                        learn_rate=0.5,
                        feps=9.9999999999999995e-07,
                        quench=1.0,
                        n=1.0,
                        dwell=250 )
```

# Appendix 3

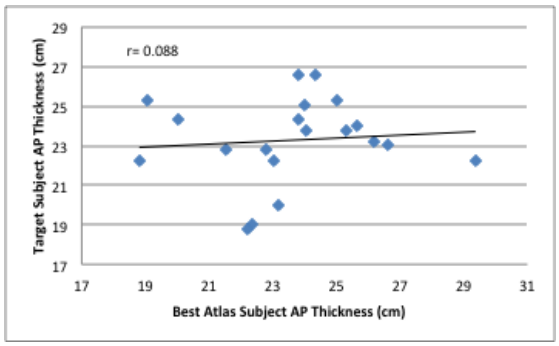
This section provides a patient-specific data collected during this work, the table contains patients' anatomical parameters.

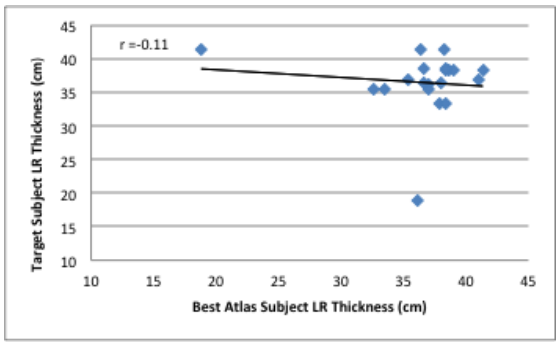
| Patient | PL (cm) | RFD (cm) | LFD (cm) | AP (cm) | LR (cm) | CTV-BW (cm) | CTV-RW (cm) | CTV-RF (cm) | CTV-LF (cm) | CTV (cc) | BW (cc) | RW (cc) | RF (cc) | LF (cc) |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 4.20 | 4.31 | 4.59 | 26.20 | 38.63 | 5.01 | 3.72 | 11.29 | 11.22 | 52.86 | 18.46 | 12.75 | 165.95 | 167.96 |
| **2** | 3.58 | 4.55 | 4.18 | 21.52 | 38.58 | 4.09 | 3.60 | 9.99 | 10.13 | 41.35 | 16.32 | 14.16 | 168.07 | 168.16 |
| **3** | 5.32 | 4.04 | 4.04 | 23.02 | 36.99 | 5.15 | 3.88 | 11.46 | 11.35 | 73.33 | 29.95 | 16.96 | 165.60 | 148.38 |
| **4** | 5.66 | 4.19 | 4.59 | 23.80 | 36.44 | 4.51 | 3.23 | 11.20 | 10.95 | 69.10 | 17.07 | 15.03 | 170.28 | 167.75 |
| **5** | 5.38 | 4.33 | 4.71 | 25.65 | 39.01 | 4.23 | 4.25 | 11.83 | 11.49 | 75.45 | 25.45 | 18.39 | 211.29 | 202.27 |
| **6** | 6.00 | 5.46 | 5.38 | 24.01 | 37.95 | 4.93 | 4.61 | 10.94 | 10.88 | 126.62 | 34.16 | 18.26 | 197.13 | 201.20 |
| **7** | 4.99 | 4.88 | 4.93 | 25.32 | 36.97 | 5.02 | 3.44 | 11.45 | 12.10 | 59.57 | 24.76 | 19.20 | 174.22 | 175.98 |
| **8** | 5.23 | 5.11 | 4.98 | 22.77 | 35.37 | 4.88 | 3.47 | 11.34 | 11.50 | 78.12 | 22.41 | 13.22 | 189.94 | 184.89 |
| **9** | 5.41 | 4.81 | 4.80 | 25.04 | 38.20 | 5.00 | 3.24 | 10.88 | 10.82 | 46.75 | 28.31 | 15.91 | 170.47 | 165.67 |
| **10** | 6.00 | 4.50 | 4.24 | 24.36 | 38.33 | 4.16 | 3.89 | 12.08 | 11.37 | 113.96 | 25.24 | 14.27 | 232.07 | 225.01 |
| **11** | 4.79 | 4.95 | 5.23 | 23.19 | 38.39 | 4.52 | 4.15 | 11.88 | 11.63 | 84.21 | 23.81 | 15.07 | 191.26 | 192.51 |
| **12** | 4.81 | 4.78 | 4.52 | 22.34 | 37.87 | 3.67 | 3.82 | 11.75 | 11.53 | 75.16 | 25.31 | 17.25 | 207.17 | 210.96 |
| **13** | 3.01 | 4.36 | 4.35 | 21.62 | 34.75 | 4.71 | 3.63 | 11.06 | 10.68 | 67.14 | 27.58 | 15.30 | 169.65 | 165.35 |
| **14** | 4.50 | 4.23 | 4.50 | 27.70 | 41.09 | 5.69 | 3.96 | 10.98 | 10.91 | 45.63 | 24.58 | 14.04 | 161.45 | 167.22 |
| **15** | 4.02 | 5.03 | 4.61 | 18.81 | 34.14 | 3.55 | 3.07 | 10.84 | 11.11 | 37.05 | 16.66 | 11.14 | 165.55 | 154.81 |
| **16** | 6.00 | 4.63 | 4.96 | 26.62 | 41.35 | 5.64 | 4.62 | 11.98 | 11.90 | 77.36 | 24.97 | 20.34 | 205.55 | 203.75 |
| **17** | 6.60 | 4.62 | 4.78 | 23.80 | 36.62 | 4.78 | 3.55 | 11.54 | 11.57 | 107.72 | 24.86 | 14.45 | 185.61 | 189.69 |
| **18** | 5.72 | 5.23 | 5.26 | 20.02 | 38.36 | 3.58 | 4.02 | 12.57 | 11.62 | 102.94 | 27.06 | 21.50 | 209.14 | 206.60 |
| **19** | 3.77 | 4.47 | 4.45 | 20.27 | 34.12 | 3.75 | 3.33 | 10.85 | 11.13 | 39.17 | 19.75 | 13.78 | 158.79 | 163.62 |
| **20** | 4.81 | 5.27 | 5.18 | 25.23 | 38.31 | 5.12 | 4.29 | 11.65 | 11.38 | 39.06 | 19.46 | 18.56 | 180.34 | 177.55 |
| **21** | 4.20 | 4.43 | 4.54 | 24.3 | 35.58 | 4.11 | 3.55 | 9.99 | 9.87 | 36.95 | 18.44 | 12.83 | 172.66 | 174.57 |
| **22** | 4.25 | 4.15 | 4.45 | 24.01 | 37.7 | 5.01 | 4.06 | 11.44 | 11.46 | 89.96 | 28.28 | 16.33 | 181.4 | 183.09 |
| **23** | 3.89 | 4.58 | 4.48 | 22.75 | 35.27 | 4.38 | 3.84 | 11.54 | 10.94 | 10.6 | 54.66 | 18.08 | 165.13 | 152.7 |
| **24** | 6.33 | 5.16 | 5.11 | 25.71 | 37.94 | 4.75 | 3.88 | 11.96 | 11.82 | 87.27 | 21.81 | 16.85 | 211.12 | 214.56 |
| **25** | 6.03 | 4.55 | 4.7 | 21.41 | 34.35 | 3.82 | 3.96 | 11.7 | 10.45 | 100.51 | 20.85 | 13.65 | 165.2 | 161.91 |
| **26** | 4.8 | 5.11 | 4.99 | 25.04 | 38.55 | 5.23 | 4.09 | 12.19 | 12.07 | 54.42 | 20.73 | 14.62 | 222.49 | 225.4 |
| **27** | 4.78 | 4.40 | 3.70 | 20.58 | 34.75 | 5.54 | 2.83 | 11.62 | 11.14 | 58.95 | 52.81 | 16.55 | 190.02 | 186.56 |
| **28** | 4.63 | 3.95 | 4.54 | 21.83 | 34.73 | 3.82 | 4.22 | 10.35 | 10.64 | 39.69 | 18.69 | 12.36 | 157.5 | 150.55 |
| **29** | 3.32 | 4.7 | 5.2 | 19.87 | 38.95 | 3.83 | 4.14 | 12.79 | 12.2 | 64.04 | 37.1 | 17.53 | 225.58 | 192.44 |
| **30** | 4.24 | 4.65 | 4.53 | 21.48 | 35.75 | 3.8 | 4.25 | 10.94 | 10.81 | 38.3 | 20.09 | 16.34 | 153.89 | 158.54 |

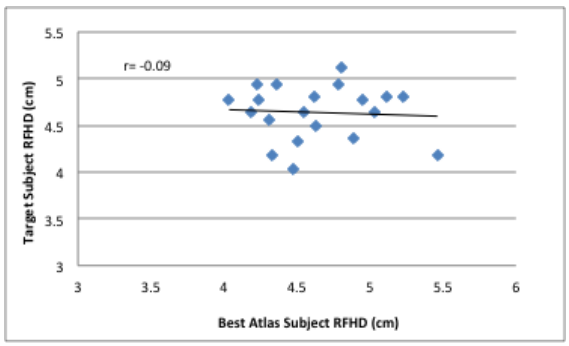
**Table A3.1.** The clinical parameters for each patient including, cranial-caudal prostate length (PL), right femoral head diameter (RFHD), left femoral head diameter (LFHD), anterior-posterior thickness (AP), lateral thickness (LR), distances between the CTV (prostate) centroid and the centroids of all other ROIs, and all contour volumes.

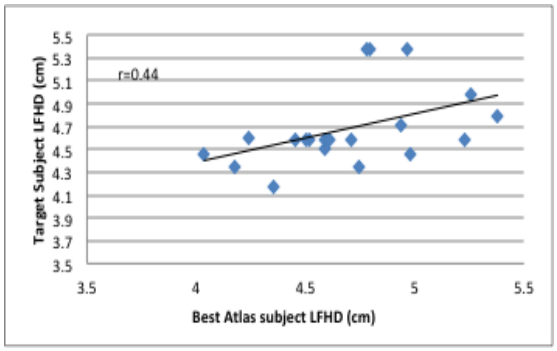
# Appendix 4

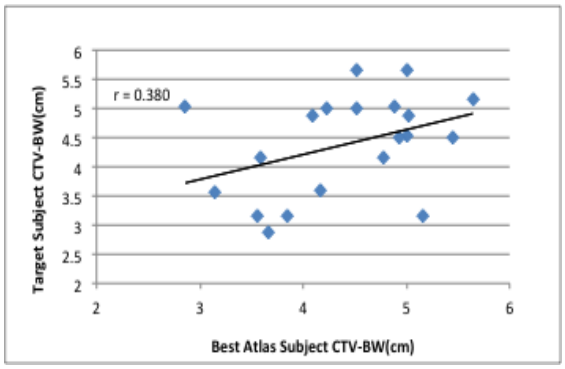
This appendix provides a result of the correlation obtained for the anatomic characteristics between the target and the best atlas. It was noticed that LR thickness for one patient was just less than 20 cm. This is obviously a mistake that should be corrected for future work.

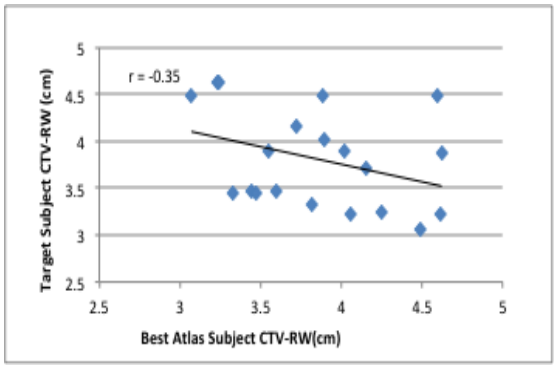


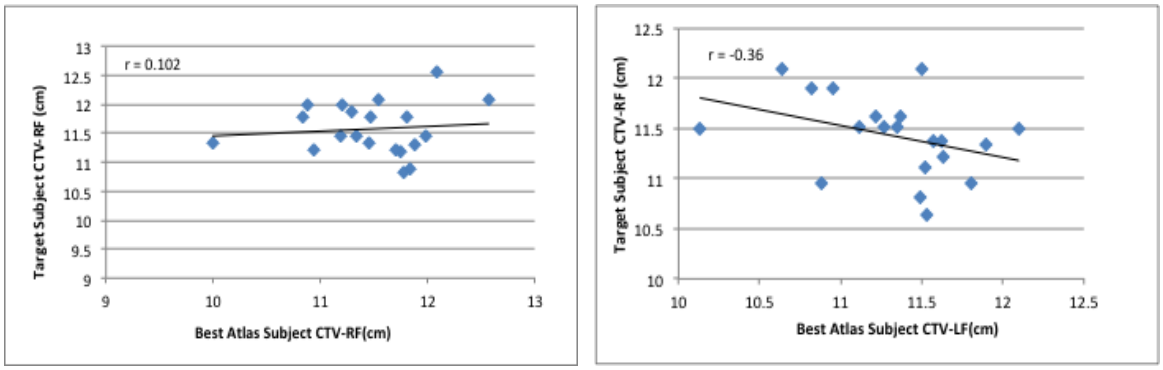



**Table A4.1**: Correlation between the target subject and the best atlas for all the anatomical parameters. The CTV in this case includes the prostate gland only.