

Knowledge-based IMRT Treatment Planning for Prostate Cancer

by

Vorakarn Chanyavanich

Graduate Program in Medical Physics  
Duke University

Date: \_\_\_\_\_

Approved:

\_\_\_\_\_  
Joseph Y. Lo, Co-Supervisor

\_\_\_\_\_  
Shiva K. Das, Co-Supervisor

\_\_\_\_\_  
Georgia Tourassi

\_\_\_\_\_  
Timothy Turkington

Dissertation submitted in partial fulfillment of  
the requirements for the degree of Doctor of Philosophy in the  
Graduate Program in Medical Physics in the Graduate School  
of Duke University

2011

ABSTRACT

Knowledge-based IMRT Treatment Planning for Prostate Cancer

by

Vorakarn Chanyavanich

Graduate Program in Medical Physics  
Duke University

Date: \_\_\_\_\_  
Approved: \_\_\_\_\_

\_\_\_\_\_  
Joseph Y. Lo, Co-Supervisor

\_\_\_\_\_  
Shiva K. Das, Co-Supervisor

\_\_\_\_\_  
Georgia Tourassi

\_\_\_\_\_  
Timothy Turkington

An abstract of a dissertation submitted in partial  
fulfillment of the requirements for the degree  
of Doctor of Philosophy in the  
Graduate Program in Medical Physics in the Graduate School  
of Duke University

2011

Copyright by  
Vorakarn Chanyavanich  
2011

## **Abstract**

The goal of intensity-modulated radiation therapy (IMRT) treatment plan optimization is to produce a cumulative dose distribution that satisfies both the dose prescription and the normal tissue dose constraints. The typical manual treatment planning process is iterative, time consuming, and highly dependent on the skill and experience of the planner. We have addressed this problem by developing a knowledge-based approach that utilizes a database of prior plans to leverage the planning expertise of physicians and physicists at our institution. We developed a case-similarity algorithm that uses mutual information to identify a similar matched case for a given query case, and various treatment parameters from the matched case are then adapted to derive new treatment plans that are patient specific.

We used 10 randomly selected cases matched against a knowledge base of 100 cases to demonstrate that new, clinically acceptable IMRT treatment plans can be developed. This approach substantially reduced planning time by skipping all but the last few iterations of the optimization process. Additionally, we established a simple metric based on the areas under the curve (AUC) of the dose volume histogram (DVH), specifically for the planning target volume (PTV), rectum, and bladder. This plan quality metric was used to successfully rank order the plan quality of a collection of knowledge-based plans. Further, we used 100 pre-optimized plans (20 query x 5 matches) to show that the average normalized MI score can be used as a surrogate of overall plan quality. Plans of lower pre-optimized plan quality tended to improve substantially after optimization, though its final plan quality did not improve to the same level as a plan that has a higher pre-optimized plan quality to begin with. Optimization usually improved PTV coverage slightly while providing substantial dose sparing for both

bladder and rectum of 12.4% and 9.1% respectively. Lastly, we developed new treatment plans for cases selected from an outside institution matched against our site-specific database. The knowledge-based plans are very comparable to the original manual plan, providing adequate PTV coverage as well as substantial improvement in dose sparing to the rectum and bladder.

In conclusion, we found that a site-specific database of prior plans can be effectively used to design new treatment plans for our own institution as well as outside cases. Specifically, knowledge-based plans can provide clinically acceptable planning target volume coverage and clinically acceptable dose sparing to the rectum and bladder. This approach has been demonstrated to improve the efficiency of the treatment planning process, and may potentially improve the quality of patient care by enabling more consistent treatment planning across institutions.

# Contents

Abstract .....	iv
List of Tables.....	viii
List of Figures .....	ix
List of Abbreviations .....	xi
Acknowledgements.....	xiii
1. Introduction.....	1
1.1 Preview of Chapters.....	1
2. Background.....	3
2.1 Clinical motivation.....	3
2.2 IMRT Treatment Planning .....	4
2.3 Knowledge-based applications .....	9
2.4 Features of the IMRT knowledge-base.....	11
2.4.1 Beam’s eye view (BEV) contour projections.....	11
2.4.2 Fluence maps .....	15
3. Knowledge-based Treatment Planning .....	17
3.1 Introduction .....	17
3.2 Methods .....	18
3.2.1 Knowledge-base reference library .....	18
3.2.2 Case similarity algorithm.....	18
3.2.3 Generating new treatment plans.....	24
3.2.4 Dose-volume constraints.....	25
3.3 Results .....	27
3.4 Discussion.....	33

3.4.1 Size of knowledge base .....	34
3.4.2 Improving case similarity retrieval .....	34
3.4.3 Assessing DVH plan quality .....	35
3.4.4 Reduction in Treatment Planning time .....	36
4. Evaluation of Treatment Plan Quality .....	37
4.1 Introduction .....	37
4.2 Material and Methods .....	38
4.2.1 Knowledge-base reference library .....	38
4.2.2 Generating new treatment plans.....	38
4.2.3 Defining a treatment plan quality metric .....	39
4.2.4 Evaluating pre-optimized plans using the plan quality index.....	40
4.3 Results .....	40
4.4 Discussions .....	52
5. Inter-institutional comparison of plan quality .....	55
5.1 Introduction .....	55
5.2 Material and methods.....	56
5.3 Results .....	59
5.4 Discussion.....	71
6. Summary, Conclusion and Future Work .....	75
6.1 Summary of Findings .....	75
6.2 Conclusion.....	77
6.3 Future Work .....	78
References .....	81
Biography .....	86

## List of Tables

Table I Duke University's prostate IMRT protocol with dose-volume constraints .....	26
Table II Percentage difference in PTV coverage between new and original plan.....	31
Table III Percentage difference in rectum dose between new and original plan .....	32
Table IV Percentage difference in bladder dose between new and original plan.....	33
Table V Effect of $AUC_{PTV}$ rank order for different weighting schemes .....	43
Table VI Percentage difference in AUC between knowledge-based plan vs. manual plan .....	71



## List of Figures

Figure 1 Seven co-planar beams and a 2D fluence map.....	5
Figure 2 Five convergent radiation beams .....	6
Figure 3 Volume rendered structures .....	13
Figure 4 Orientation of seven beam angles.....	14
Figure 5 Seven projection images .....	15
Figure 6 Seven fluence maps for each of the treatment beam orientations.....	16
Figure 7 Overlapping contour regions .....	19
Figure 8 Plot of MI values for 100 cases .....	21
Figure 9 Example of a query image and various matched images.....	22
Figure 10 Graphical user interface used to compare query and matched cases.....	23
Figure 11 Fluence maps, before and after registration with the PTV.....	25
Figure 12 Comparison of DVH for the new plan vs. original plan (case #8).....	28
Figure 13 Comparison of DVH for the new plan vs. original plan (case #7).....	29
Figure 14 Comparison of DVH for the new plan vs. original plan (case #5).....	30
Figure 15 Range of DVH plans using matched cases of high to low similarity .....	41
Figure 16 Example of un-weighted plan quality scoring method, rank-ordered.....	42
Figure 17 Rank order of 20 cases from #1 to #96 for PTV weighted scheme.....	44
Figure 18 Plan quality versus MI for each angle .....	45
Figure 19 Plan quality versus average normalized MI score .....	46
Figure 20 Percent change in PTV AUC versus the initial pre-optimized PTV AUC .....	48
Figure 21 Percent change in rectum AUC versus the initial pre-optimized AUC .....	49
Figure 22 Percent change in bladder AUC versus the initial pre-optimized AUC .....	50
Figure 23 Percent change in overall plan quality versus the initial PQ.....	51

Figure 24 Diagram of AUC shoulder and AUC tail regions of the PTV.....	59
Figure 25 Four examples of the set of seven fluence maps from Duke. ....	60
Figure 26 Four examples of the set of seven fluence maps from Institution #1. ....	61
Figure 27 Four examples of the set of seven fluence maps from Institution #2. ....	62
Figure 28 Twenty Duke cases sorted from high PQ (#1) to low PQ (#96) .....	63
Figure 29 Twenty cases from Institution #1, sorted from high PQ to low PQ.....	65
Figure 30 Twenty cases from Institution #2, sorted from high PQ to low PQ.....	67
Figure 31 Comparison of knowledge-based plan (Case #1) vs. manual plan .....	68
Figure 32 Comparison of knowledge-based plan (Case #7) vs. manual plan .....	69
Figure 33 Comparison of knowledge-based plan (Case #6) vs. manual plan .....	70

## List of Abbreviations

AAPM – American Association of Physicists in Medicine

ACS – American Cancer Society

ASTRO – American Society for Therapeutic Radiology and Oncology

AUC – area under the curve

BEV – Beam’s Eye View

CAD – Computer Aided Detection

cGy – centigray

CBIR – Content-Based Image Retrieval

CC – Cross-correlation

CERR – Computational Environment for Radiotherapy Research

CT – Computed Tomography

$D_x$  – dose to X-percent (*e.g.* 30%) of the volume

DVH – Dose-Volume Histogram

EUD – Equivalent Uniform Dose

Gy - Gray

IMRT – Intensity Modulated Radiation Therapy

IRB – Institutional Review Board

KB – knowledge-based

MI – Mutual Information

NCI – National Cancer Institute

OAR – organs at risk

PCRP- Prostate Cancer Research Program

PQ – Plan quality

PTV – Planning Target Volume

ROI – Region of Interest

RTOG – Radiation Therapy Oncology Group

$V_x$  – Volume receiving X Gray

## Acknowledgements

I would like to acknowledge my committee members for the mentoring and active involvement throughout this research project. In particular, I would like to thank my co-advisors, Dr. Joseph Lo and Dr. Shiva Das, for their generous guidance and continued support throughout this project. Thank you for providing the independence, trust, flexibility and the unwavering kind support throughout my time here at Duke. I am very fortunate to have had the rare opportunity to be treated as an equal partner with two outstanding gentlemen scholars. Dr. Georgia Tourassi has been generous with her support and enthusiasm for this project. She has shared the much needed wisdom, wit, and humor to make it all fun. Brian Harrawood, has engaged me daily with insightful and memorable discussions, of all things scientific, political as well as the ludicrous. Of course, I also want to thank both Gina and Brian for their permission to allow me to adapt their C++ mutual information code. I am very grateful for Dr. Tim Turkington's involvement with the project, particularly for his flexibility and generosity with his time. Additionally, Dr. William Robert Lee has provided valuable clinical feedback on multiple occasions to help advance this project towards clinical relevancy. I also want to thank Danthai, Matt, Carl, and Chris for their technical assistance throughout the various phases of the project.

Finally, to Christy, Ana, Justin, Jackie, Arnaud, and Xiang. I have developed many wonderful friendships here in North Carolina, and *y'all* have helped me to maintain balance and motivation throughout graduate school. Christine and Stephen, *muito obrigado* for being my family away from home. Lastly, I want to thank my wife Becky and my entire Chanyavanich and Powell family for their unwavering support -

from Denver, Bangkok and Jackson - know that you've each been by my side throughout the whole process, even if from afar.

This research was partially supported by a grant from the Wallace H. Coulter Translational Research Partnership Award. Additionally, Dr. Samei, Dr. Dobbins and Olga Baranova in the Duke medical physics program have been very supportive and have helped to provide bridge funding to complete this project.

# 1. Introduction

Knowledge-based IMRT treatment planning is a semi-automated approach of retrieving and adapting prior information from a similarly matched case in order to generate a new, clinically acceptable treatment plan. By leveraging the expertise and manually optimized efforts expended on prior cases, we can efficiently generate new treatment plans and skip ahead to the last few iterations of the optimization process. This approach has the potential to improve quality of treatment planning across institutions, as well as to provide a computer-assisted tool to ensure that treatment plans of acceptable plan quality are generated.

## 1.1 Preview of Chapters

*Chapter 2 Background* briefly presents the clinical motivation and an overview of intensity modulated radiation therapy (IMRT) treatment planning for prostate cancer, while highlighting the challenges of the treatment planning process. An overview of the knowledge-based approach is presented as a potential solution to some of these challenges.

*Chapter 3 Knowledge-based Treatment Planning* introduces the framework of using prior cases that have already been manually optimized and clinically approved, for developing new treatment plans. Mutual information is a promising method for developing a case-similarity algorithm to identify similar cases based on matching 2D beam's eye view (BEV) contour projections. Fluence map registration and the importing of optimization weights are important successive steps to developing new treatment

plans. Results of using the single best-matched case demonstrate the feasibility of the knowledge-based approach for developing new, clinically acceptable, treatment plans.

*Chapter 4 Evaluating Plan Quality* explores the range in case-similarity of a matched case that can be used to produce a clinically acceptable plan. In other words, the study examines whether the resulting plan quality is correlated with the degree of case-similarity. The knowledge base is expanded to include 250 cases from the Duke clinic, and a larger study is conducted to generate 100 new treatment plans. Whereas the initial study in *Chapter 3* used the single best-matched case, this larger study explores the effect of using matched cases that span the range in case-similarity. A plan quality metric is introduced and used as a figure of merit to quantify and compare the dose-to-volume coverage of knowledge-based treatment plans.

*Chapter 5 Inter-institutional Comparison of Plan Quality* presents the final validation study that examines the feasibility of applying this treatment planning scheme to independent cases from another institution. This retrospective study investigates whether it is feasible to use site-specific Duke knowledge base to develop clinically acceptable plans for new query cases drawn from outside institutions. Additionally, the study provides an opportunity to conduct an inter-institutional comparison of clinical plan quality, as well as a comparison of manual planning versus semi-automated knowledge-based planning.

*Chapter 6 Summary, Conclusions and Future Work* summarizes the knowledge-based IMRT treatment planning process, discusses additional approaches of using mutual information for determining case-similarity, and concludes with recommendations for the direction for future work.



## **2. Background**

We first introduce the clinical motivation and some of the clinical challenges in prostate cancer radiotherapy. Next, we present a brief overview of the IMRT treatment planning process and define the problem of IMRT fluence optimization. Finally, we review prior knowledge-based applications in the field, and examine the necessary features of a knowledge base of IMRT treatment plans.

### ***2.1 Clinical motivation***

Prostate cancer is the most common cancer among men of all races in the United States, accounting for 217,730 new cases in 2010[1]. The American Cancer Society (ACS) reported that prostate cancer is the second leading cause of cancer death in men over age 45[2]. Approximately 1.1 million cancer patients each year undergo radiation therapy at 2,169 hospitals and freestanding radiation therapy centers [3, 4]. Nearly 90% of those patients are treated with external-beam radiation therapy from a linear accelerator. Most prostate cancer patients undergo Intensity Modulated Radiation Therapy (IMRT), which result in 90% disease-free rates after five years for early stage cases [5].

The goal of prostate IMRT treatments is to deliver therapeutic dose while maximizing patient survival and quality of life. IMRT uses one external radiation beam, at different subsequent orientations, that converge to deliver a therapeutic dose to the target volume while sparing adjacent normal tissue, i.e. bladder, rectum, and femoral heads [6]. The most frequent use of IMRT in the United States is to treat prostate cancer, and prostate cases at our institution account for approximately one-third of all procedures [7]. Given the prevalence of prostate cancer and the reliance on IMRT for

treating it, any improvements to prostate IMRT can have an immediate and substantial clinical impact.

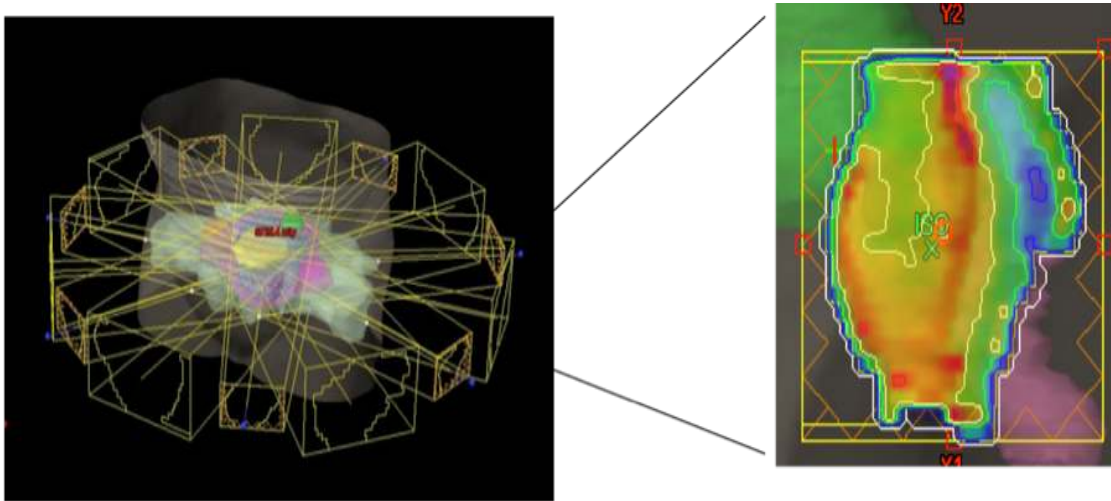
Currently, each IMRT plan is developed *de novo* in an inefficient manner that is both time consuming and subjective, and may not even be clinically optimal. The entire process of IMRT treatment planning can take several hours per case to achieve a clinically acceptable plan [8, 9]. The iterative process involves striking a compromise between the conflicting constraints of providing homogenous coverage of the prostate target volume while simultaneously sparing dose from the adjacent normal critical structures. While a plan can be deemed clinically acceptable, it may not be fully optimized to minimize normal tissue doses to the highest extent possible. The quality of a treatment plan is highly dependent on the skill and experience level of the physicist/dosimetrist planner. This implies that plans at centers with limited expertise may not be clinically optimal, in that they may not minimize the bladder and rectal doses to the maximum extent possible. It has recently been reported that there is wide variability across medical institutions between the prescribed radiation dose and the dose delivered, with a total of 46% of the patients receiving a maximum dose that was more than 10% higher than the prescribed dose and 63% of the patients receiving a dose that was more than 10% lower than the prescribed dose [10].

## **2.2 IMRT Treatment Planning**

Most institutions utilize the same fixed beam orientations to deliver IMRT for prostate cancer. In other words, the number and angular position of the beams are usually the same for all patients, typically five to seven co-planar beams. *Figure 1* depicts

the typical arrangement of seven co-planar beams for a prostate IMRT treatment plan, as displayed on the Eclipse™ commercial treatment planning system.

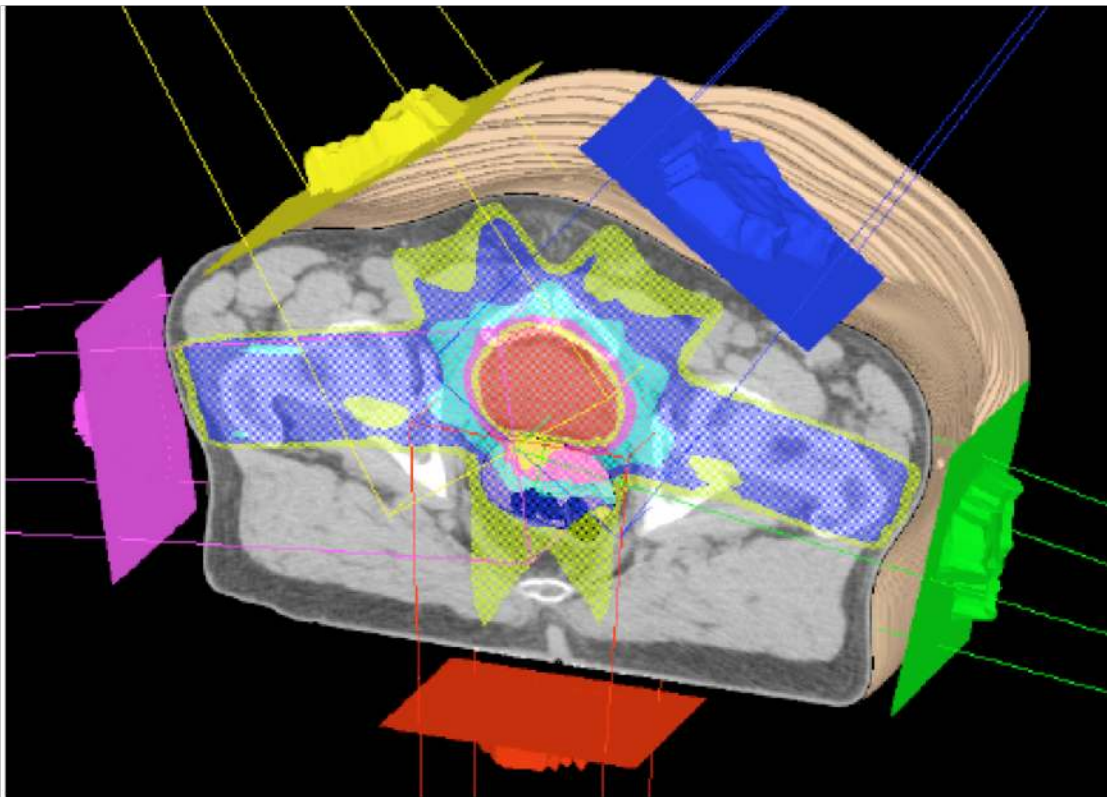
At our institution, the patient is immobilized and treated in the supine position with seven co-planar beams. Each beam in turn is composed of many “beamlets”, akin to pixels in a digital image. The radiation fluence of each beamlet can be individually modulated, resulting in a desired fluence pattern that can deliver greater dose to one area while avoiding another. Each such fluence pattern from the “beam’s eye view” (BEV) forms a 2D map of radiation intensity. Each patient’s unique anatomy requires deriving a new set of fluence patterns.



**Figure 1** Seven co-planar beams and a 2D fluence map

A typical arrangement of seven co-planar beams for a prostate IMRT treatment plan. Each of the beam angles delivers a unique 2D map of modulated radiation intensity, also known as the fluence map [11].

When multiple beams are focused together toward the target volume, the result is a 3D confluence that delivers therapeutic dose to the prostate while minimizing dose to surrounding critical structures (i.e., the rectum, bladder, and femoral heads). Refer to *Figure 2*. Undue dose to those critical structures can result in poor survival and reduced quality of life to the patient [12, 13]. Potential radiotoxicities to these critical structures include bladder, erectile and bowel dysfunction, urinary and rectal incontinence, bleeding and fistula [14-19].



**Figure 2** Five convergent radiation beams

Each of the five impinging beams has a different aperture and fluence, resulting in a convergence of the desired dose distributions [11].

Using a commercial treatment planning system, the planner specifies as inputs the prescription dose to the target volume, as well as dose constraints for organs at risk. Establishing clinically useful dose constraints is more subjective for several reasons: 1) dose constraints are based on clinical experience, 2) dose constraints are population-based, 3) dose constraints are not deterministic, and 4) the dose-complication relationship is not well understood.

The typical clinical practice strives to deliver a prescription dose (or higher) to at least 98% of the tumor volume while not exceeding normal structure dose constraints. In doing so, while this practice may meet minimal standards, there may be often no additional effort to reduce normal structure dose any further, thereby potentially resulting in radiation toxicity. We define high quality plans in term of the following characteristics: 1) prescription dose to at least 98% of the tumor volume, 2) normal structure dose constraints are not exceeded, 3) minimize normal structure doses as much as possible. Many community centers utilize dose constraints published in the literature as the limit below which there are no complications. However, since these limits are population-based, the best treatment planning strategy for any individual patient would be to reduce the dose as much as possible by pushing the limit of dose sparing. Currently, this strategy requires a manual, trial-and-error approach, which is both very time-consuming and highly dependent on the experience and skill of the physicist/dosimetrist planner.

Mathematically, IMRT is considered an inverse problem whereby the final desired dose distribution are the initial input, and the incident radiation intensity patterns necessary to obtain such a dose distribution must be solved for. This

assessment is typically expressed as an objective function,  $F_{PTV}$ , the squared-difference between the computed dose and the prescription dose [6].

$$F_{PTV} = \frac{1}{N_i} \sum_{i=1}^N (D_i - D_{presc})^2 \quad (1)$$

where  $N_i$  is the number of points in the target volume,  $D_i$  is the dose to point  $i$  and  $D_{presc}$  is the prescription dose. Similarly, the objective function for organs at risk (OAR),  $F_{OAR}$ , is defined in terms of the maximum dose constraints and dose-volume constraints, including penalty weights for each specific organ,

$$F_{OAR} = \frac{1}{N_{OAR}} \left[ w_{OAR,max} \sum_{i=1}^{N_{OAR}} (D_i - D_{max})^2 + w_{OAR,dv} \sum_{i=1}^{N_{dv}} (D_i - D_{dv})^2 \right] \quad (2)$$

where  $N_{OAR}$  is the number of points in the OAR,  $D_{max}$  is the maximum tolerance dose, and  $N_{OAR,dv}$  is the number of points whose dose must be below the dose-volume constraint dose  $D_{dv}$ . The penalty weights are  $w_{OAR,max}$  and  $w_{OAR,dv}$  [6].

In manual treatment planning, the iterative process involves adjusting dose constraints, performing the dose calculation, and reviewing the clinical plan. This trial and error process is repeated until the user is satisfied that a clinically acceptable plan has been reached. The dose distributions are reviewed as a dose-volume histogram plot and/or as color-wash isodose profiles overlaid on the CT. An ideal treatment plan would have 100% of the planning target volume receiving the prescription dose and none of the critical structures would receive any radiation dose. In reality, due to the relative proximity and/or overlap of the PTV and adjacent critical structures, there is

always a trade-off between maximizing therapeutic dose to the PTV and minimizing dose to normal critical structures.

### **2.3 Treatment planning optimization**

Many different issues concerning the optimization of IMRT treatment plans have been extensively published, and some of the pioneering works include those of Brahme [20-23], Bortfeld [24-26], Webb [6, 27, 28], and Mohan [29, 30]. The optimization of the treatment planning process can be considered in terms of the various components and processes of the “IMRT chain”, including: *variables, cost function, search algorithm, and treatment delivery* [31]. Since the scope of our optimization project concerns only the treatment planning process, which is accomplished prior to the treatment delivery, we are not concerned with optimization issues related to the actual treatment delivery (i.e. leaf sequencing). However, we fully recognize that a plan cannot be considered optimal if it is not feasible for delivery.

The many treatment planning variables that can be optimized include: radiation type (proton, photon, electrons), energy, number of beams, beam angle, beam weights, and number of fluence levels [9, 32, 33]. In clinical practice though, photon energies are determined by the delivery system, and typical energies for prostate IMRT are 6 MeV, 9 MeV, and 15 MeV. The cost function is typically defined in physical, dose-based terms, such as maximum dose limits or dose-volume thresholds. Most commercial treatment planning systems utilize dose-based or dose-volume constraints. There are also clinical and biologically-based criteria, including tumor control probability (TCP), normal tissue complication probabilities (NTCP), and equivalent-uniform-dose (EUD) [34]. However, these biological cost functions are not yet widely adopted in typical clinical practice.

Computer search algorithms can be divided into two categories: gradient methods and stochastic methods. Two gradient methods are steepest descent and conjugate gradient, with the conjugate gradient being faster for regions of narrow valley [35]. The widely used commercial treatment planning system, Eclipse (Varian Medical Systems, Palo Alto, CA), uses a conjugate gradient optimization algorithm that allows real-time modification of the optimization parameters, thereby minimizing the overall optimization time [36]. Stochastic algorithms are typically slower than gradient methods and are not common. An example of a stochastic algorithm is simulated annealing, an iterative method that permits the escape from local minima traps [37]. The CORVUS treatment planning system implements both simulated annealing and gradient descent algorithms [38]. The most physically realistic type of treatment planning algorithms is the Monte Carlo approach, but these are computationally extensive, and have yet to be widely adoption into clinical practice [39-41].

## **2.4 Knowledge-based applications**

Knowledge-based approaches in radiation oncology have previously been used to develop imaging informatics platforms for proton therapy and radiosurgery [42, 43]. Wu *et al.* have used a database of plans for IMRT treatment plan quality control as well as for generating dose-volume histogram objectives [44, 45]. Additionally, the challenges of addressing issues of variability in manual contouring and segmentation, particularly for head and neck cancers, are being explored with atlas-based solutions that utilize patient databases [46-48].

We developed a novel knowledge-based approach of using similarity metrics to search a reference library of previously generated and clinically approved treatment



plans. These existing plans represent a valuable expertise knowledge base that “understands” how best to treat the cancer while minimizing dose to normal tissues. If a new patient can be matched to a similar, existing patient, then the existing data from that match can be used to derive new treatment plans that meet prescription objectives and are patient specific. In other words, by utilizing a knowledge base of clinically approved plans that have previously been manually optimized, we hypothesize that a semi-automated system can be used to develop new plans that are also clinically acceptable. If successful, this knowledge base of treatment plans can potentially be made available for use at other institutions, and thus, may lead to more consistent treatment planning quality across institutions.

## ***2.5 Features of the IMRT knowledge-base***

Each IMRT treatment plan includes the CT dataset, tumor/target volume and normal structure contours, beam geometry specifications, beam intensity (fluence) maps, dose prescription and dose-distributions. For our knowledge base, we use a subset of the data from the treatment plan to build a case, including the following: 1) the tumor/target volume and normal structure contours delineated by the dosimetrist/physician, 2) the beam geometry specifications, 3) the optimal fluence maps, and 4) the final optimization weights for each structure.

### **2.5.1 Beam’s eye view (BEV) contour projections**

Since each treatment beam “sees” a 2D projection of the 3D treatment volume along the angle of the radiation beam, the projections of the anatomy become the inputs that determine the aperture and fluence maps in IMRT treatment planning. One of the

most important inputs to the treatment planning dose calculation is the structure contour. In prostate IMRT, the contours for five of the most important anatomic structures include the following: planning target volume (PTV), rectum, bladder, left femoral head and right femoral head. The PTV is defined as the expansion of the clinical target volume (CTV) to provide a treatment margin that accounts for possible setup errors, patient and organ motion, and alignment errors. In this study, the contours for the small bowel, penile bulb, and seminal vesicles were excluded, since these structures are not always contoured for every case, and more importantly, the dosimetry for these structures are often less important than the major structures. A 3D volumetric model of each structure is generated from the stack of 2D slices, each of which consists of the manually contoured points for the structures visualized on the initial patient CT images taken prior to treatment. The 3D model of the five primary rendered structures was generated using the Computational Environment for Radiotherapy Research (CERR) tool, an open-source MATLAB-based radiotherapy software tool [49]. This volume rendering is illustrated in *Figure 3*, (i.e. yellow = bladder, green = PTV, olive = rectum, cyan = Left femoral head, blue-gray = Right femoral head). For many cases, the dosimetrist does not contour the shafts of the femur; therefore these volumes consist only of the femoral heads.

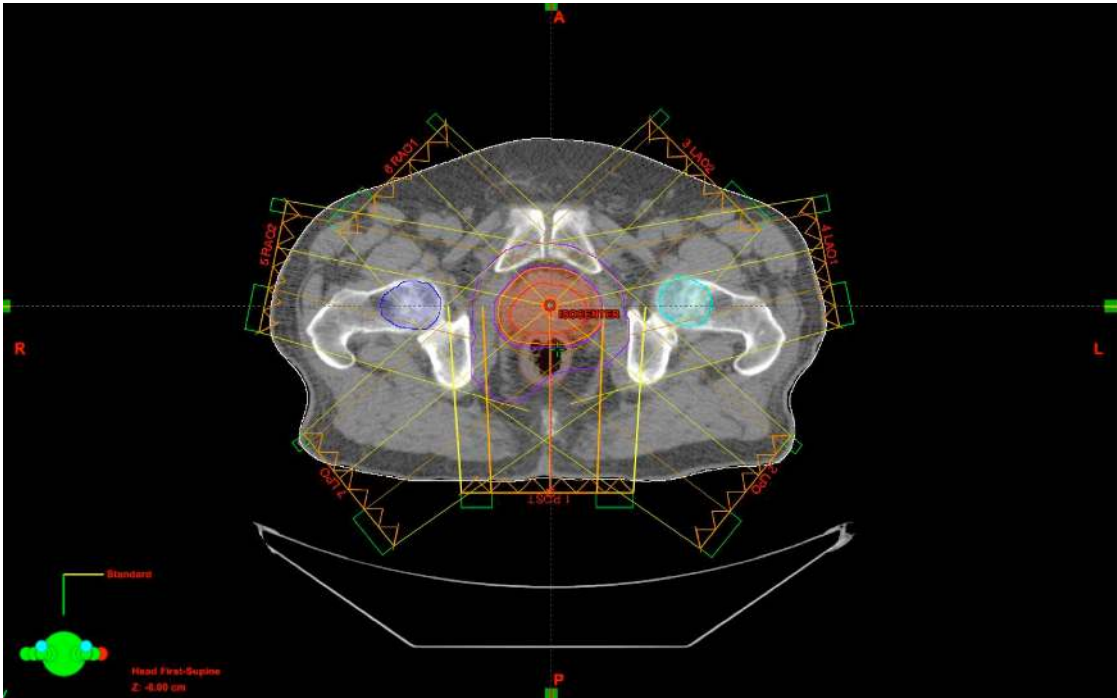


**Figure 3** Volume rendered structures

CERR was used to generate the volume rendered structures (green = bladder, red = PTV, brown = rectum, cyan = left femoral head, blue-gray = right femoral head, tan = body). The red bar and green bar respectively denotes the body orientation in the superior-inferior and left-right directions.

Next, 2D beam's eye view (BEV) projection images of the contour volumes are generated at each of the seven standard gantry angles. A single-point perspective projection is generated at each treatment angle, with the center of the field of view positioned at the isocenter, which roughly corresponds to the centroid of the PTV.

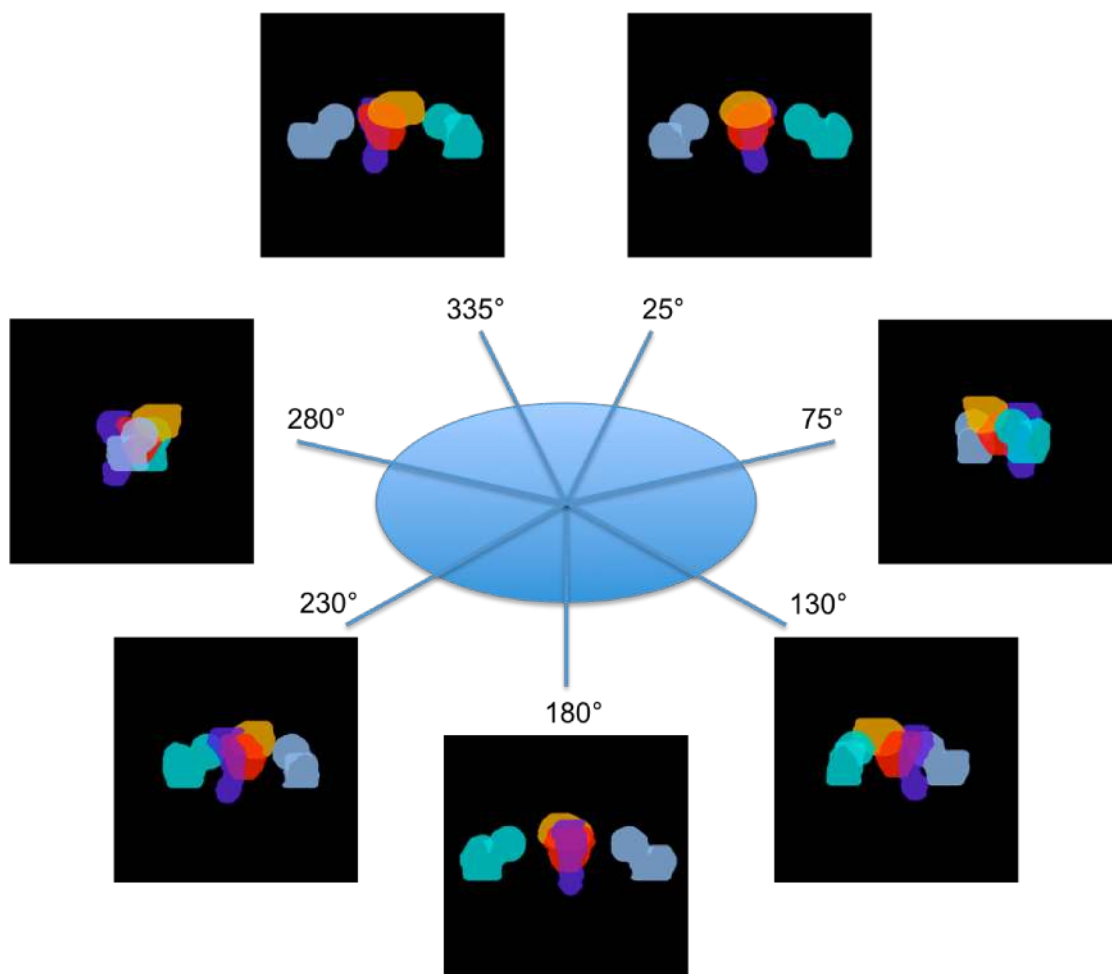
*Figure 4* depicts the orientation of these seven beam angles.



**Figure 4** Orientation of seven beam angles

The standard seven beam angles: 25°, 75°, 130°, 180°, 230°, 280°, and 335° (clockwise from top in this axis slice).

Thus, each case consists of seven BEV contour images. *Figure 5* illustrates the seven BEV projection images, showing the PTV (red), bladder (yellow), rectum (dark blue), left femoral head (cyan) and right femoral head (gray). At Duke, the majority of cases are designed using the seven-beam geometry; however, there are other institutions that may use a five-beam geometry.



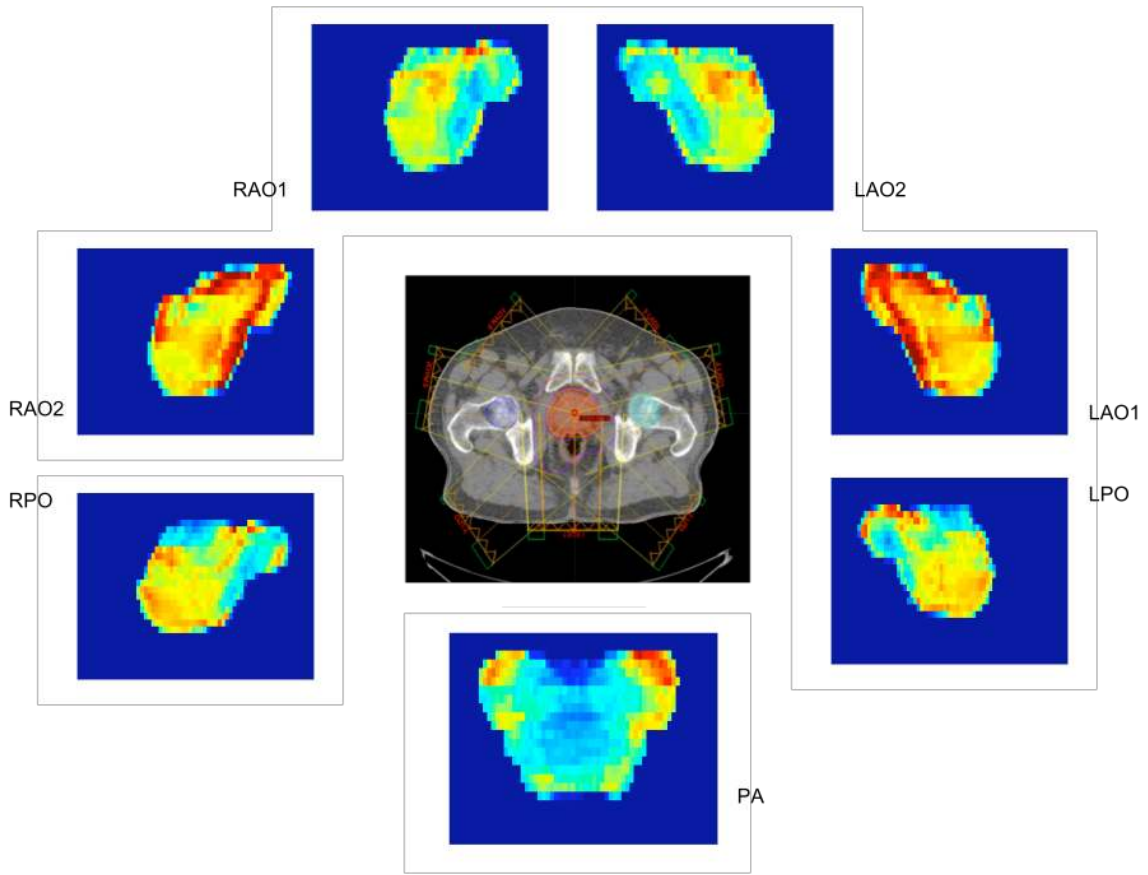
**Figure 5** Seven beam's eye view projection images

Each case consists of seven projection images, showing the PTV (red), bladder (yellow), rectum (dark blue), left femoral head (cyan) and right femoral head (gray).

### 2.5.2 Fluence maps

Each fluence map consists of a unique pattern of radiation intensity levels. For example, a typical PA fluence map may be a matrix of 72 x 48 pixels, where each pixel dimension in the x- and y- direction corresponds to 2.5 mm. Thus, each fluence map

corresponds to a field size of approximately 18 cm x 12 cm. The number of unique intensity levels can vary dramatically, depending on such factors as the beam angle, the regions where steep dose gradients are required, as well as institutional practice.



**Figure 6** Seven fluence maps for each of the treatment beam orientations

The seven fluence maps each correspond to the seven treatment beam orientations: PA, LPO, LAO1, LAO2, RAO1, RAO2, and RPO.

In summary, a knowledge-base case consists of the CT dataset, the dose prescription, the beam geometry parameters, the BEV contour projections, the fluence intensity maps and the optimization weights.

### **3. Knowledge-based Treatment Planning**

In the previous chapter, we proposed a knowledge-based approach to IMRT treatment planning, and described the components of the knowledge base containing all the relevant clinical data. In this chapter, we will demonstrate initial feasibility of our knowledge-based treatment planning system. Using existing human subject data from our institution, we generated new treatment plans using the clinical data from the best matching case in the knowledge base. These new treatment plans were assessed qualitatively and quantitatively.

#### **3.1 Introduction**

The clinical challenge in IMRT treatment planning requires solving the inverse problem in order to derive a set of fluence patterns that would produce a dose distribution that provides an effective coverage of the planning target volume while sparing the normal structures (i.e. rectum and bladder) to the greatest extent possible. We address this problem by proposing a knowledge-based solution, whereby the manually optimized prior cases that reflect the expertise and time spent by a human planner is leveraged as a semi-automated algorithm. First, a knowledge base consisting of CT data, and structure contours as well as various treatment parameters is assembled from prior optimized and clinically approved IMRT cases from the Duke clinic. Second, a case-matching algorithm is developed using mutual information as the basis for identifying similar patient cases. A query case is selected from the knowledge base, and considered as if it were a new patient case for which a treatment plan was needed. The case-similarity algorithm identifies a range of matched cases similar to the query case. Thirdly, the best-matched case having the highest overall MI score is selected and its treatment parameters are adapted to design a new treatment plan for the query case.

Lastly, we evaluate the resulting new semi-automated plans and compare these to the original plans developed by a human planner. We evaluate the plans based on dose-to-volume criteria that reflect the initial prescribed protocol.

## **3.2 Methods**

### **3.2.1 Knowledge-base reference library**

In this initial study, we assembled a database of 100 prostate cancer IMRT cases, all of which were previously planned, reviewed and delivered within the Duke clinic. IRB permission was obtained for the use of these retrospective cases in this HIPAA-compliant study, and the cases were anonymized prior to use. All clinical treatment plans were developed using the Eclipse treatment planning system (Varian Medical Systems, Palo Alto, CA). Each treatment plan includes the CT dataset, tumor/target volume and normal structure contours, beam geometry specifications, beam intensity (fluence) maps, and dose-distributions. At our institution, seven co-planar beam angles are typically used: 25°, 75°, 130°, 180°, 230°, 280°, and 330°.

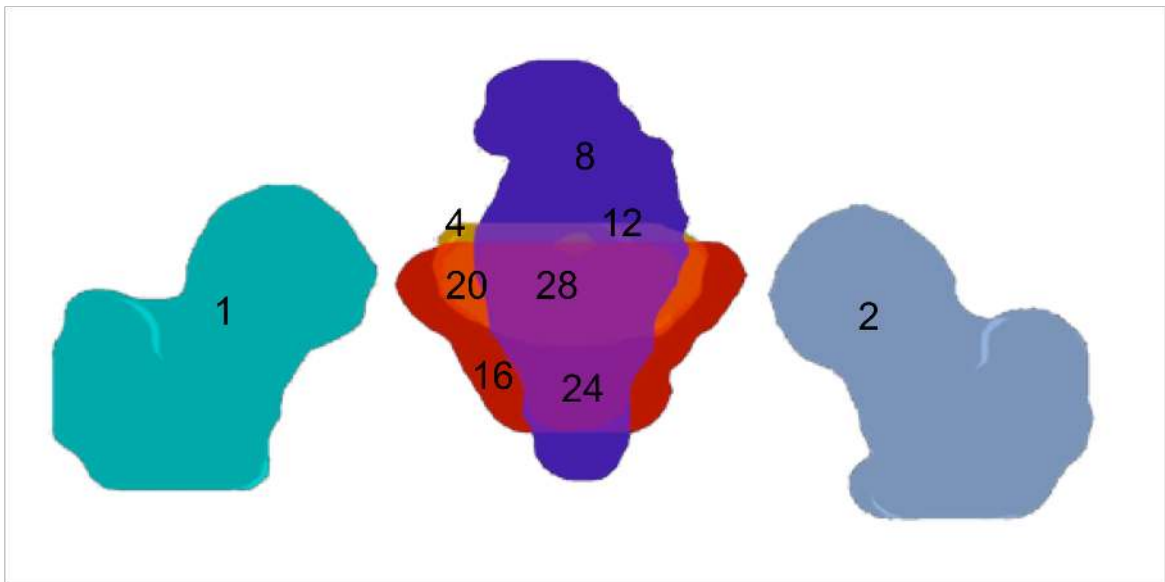
### **3.2.2 Case similarity algorithm**

Our hypothesis is that if the relative spatial locations of target volume and normal structures for a new “query” case are geometrically similar to a prior “match” case from the knowledge base, then it may be possible to use the treatment plan parameters from the matched case as an improved starting point in the planning/optimization process. The algorithm for retrieving similar patient cases is based on matching of 2D images, specifically the beam’s eye view (BEV) projections of the structure contours. Since each treatment beam “sees” a 2D projection of the 3D treatment volume along the angle of the radiation beam, the projections of the contours



become the inputs that determine the aperture and fluence maps in IMRT treatment planning.

Within each BEV projection image, a unique index number is assigned to correspond to the projection of each structure, or the overlapping regions of two or more structures. The seven 2D BEV projection images show the PTV, bladder, rectum, left and right femoral heads and the various possible overlapping regions are shown in *Figure 7*. This scheme is referred to as bit masking, where a bit represents each tissue. In each image, the five structures were assigned a unique index value (Lt. fem head = 1, Rt. fem head = 2, bladder = 4, rectum = 8, and PTV = 16), and thus, each overlapping region would also have a unique index sum (i.e. bladder + rectum = 12, PTV + rectum = 24, PTV + bladder = 20, PTV + rectum + bladder = 28).



**Figure 7 Overlapping contour regions**

Each of the five structures is assigned a unique index value (Lt. fem head = 1, Rt. fem head = 2, bladder = 4, rectum = 8, and PTV = 16). Thus, each overlapping region would also have a unique index sum (i.e. bladder + rectum = 12, PTV + rectum = 24, PTV + bladder = 20, PTV + rectum + bladder = 28).

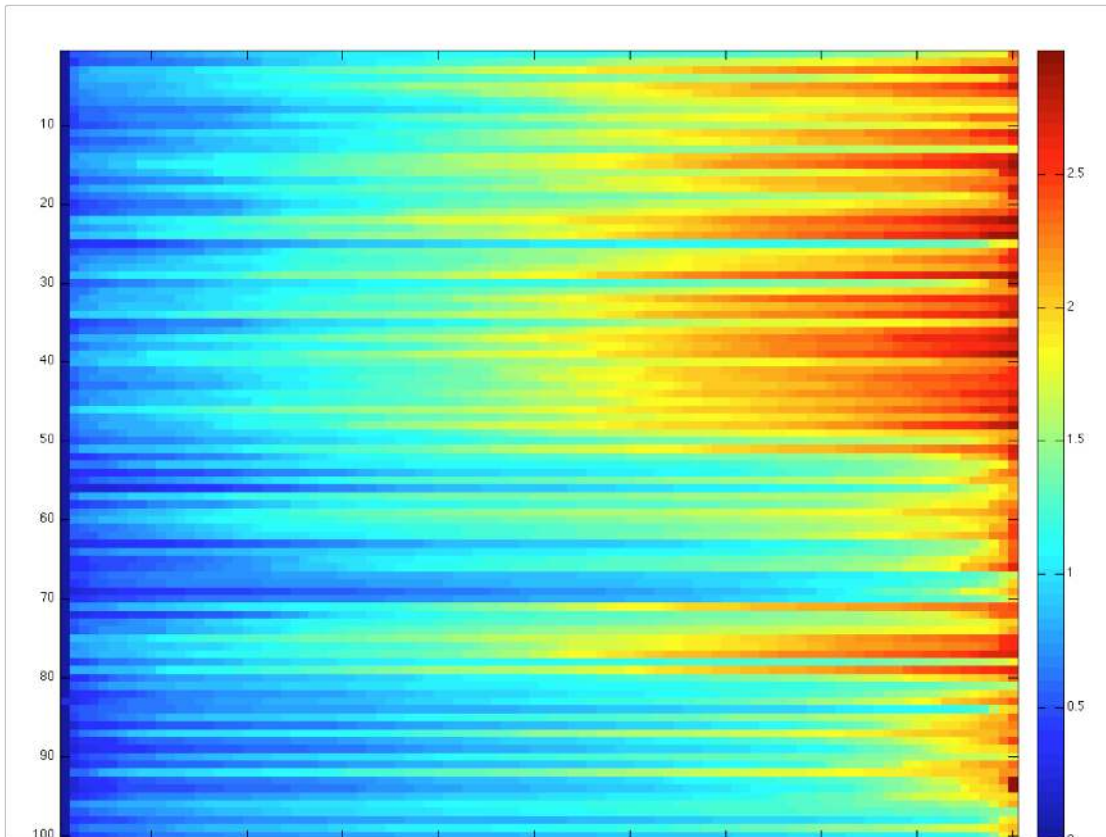
The goal of the case similarity algorithm is to match, to the best extent possible, the projection images for each of the seven beam angles of the query case to the corresponding seven beam angles of prior cases from the knowledge base. We use mutual information as the similarity metric for image matching:

$$MI(X; Y) = \sum_x \sum_y P(X, Y) \log_2 \frac{P(X, Y)}{P(X)P(Y)} \quad (3)$$

where  $P(X)$  and  $P(Y)$  are the marginal distributions and  $P(X, Y)$  is the joint probability distribution of two images  $X$  and  $Y$  [50]. MI is a general measure of statistical dependence between two images. The comparison of the histograms of the two images considers the marginal probability  $P(X)$ , corresponding to the probability of an image pixel will have an intensity value  $X$ . The more similar two images are, the higher the MI value. Because mutual information relies only on the statistical properties of the image histograms, it is more robust at tasks such as image registration and matching compared to traditional linear approaches such as cross correlation [51]. Mutual information (MI) has been widely and effectively used for image registration and fusion of multi-modality images [52]. Tourassi et al. have used MI extensively in computer aided detection applications to match 2D ROIs of mammography masses. [53-56].

The mutual information for the query case and a reference case was calculated for each of the seven treatment angles. A single composite MI value was calculated from the simple average of the seven MI values corresponding to each of the treatment angle image pairs. Using leave-one-out sampling, each of the 100 cases was used as a query and matched against the other 99 cases, resulting in a total of 9,900 matches. Since MI is not a linear metric and cannot be used as a measure of distance, the MI value has meaning only in so far as they represent relative ranking of case similarity. *Figure 8* is a plot of the MI values for 9,900 matches, sorted in ascending order from left to right,

where each row element corresponds to the MI value of a single query case matched against the other 99 cases. The plot reveals that while some cases are highly matched with a large number of other cases, there are other cases that have relatively few highly similar matches.

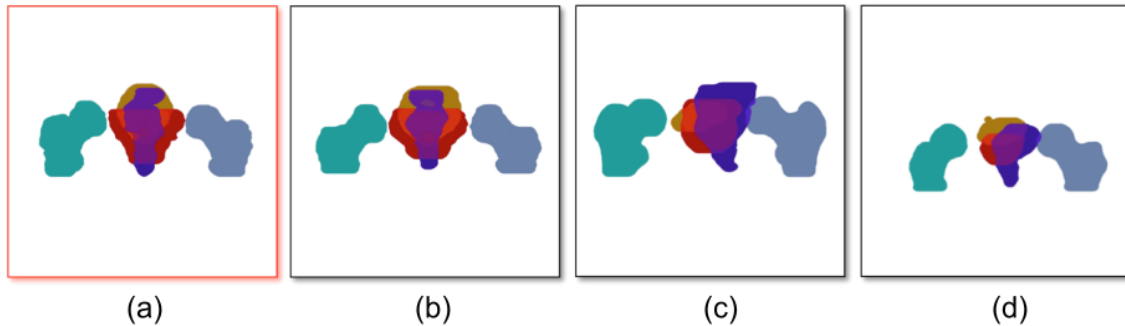


**Figure 8 Plot of MI values for 100 cases**

A plot of the MI values for the 100 cases, sorted in ascending order from high MI (left) to low MI (right). Each row element corresponds to the MI value of a single query case matched against the other 99 cases.

The overall MI scores of all the query-reference pairs in the reference library were rank-ordered. A graphical user interface (GUI) was developed to review the query case images compared with the best matches, and visual inspection of each beam angle match confirmed that the similarity algorithm successfully retrieved cases with similar

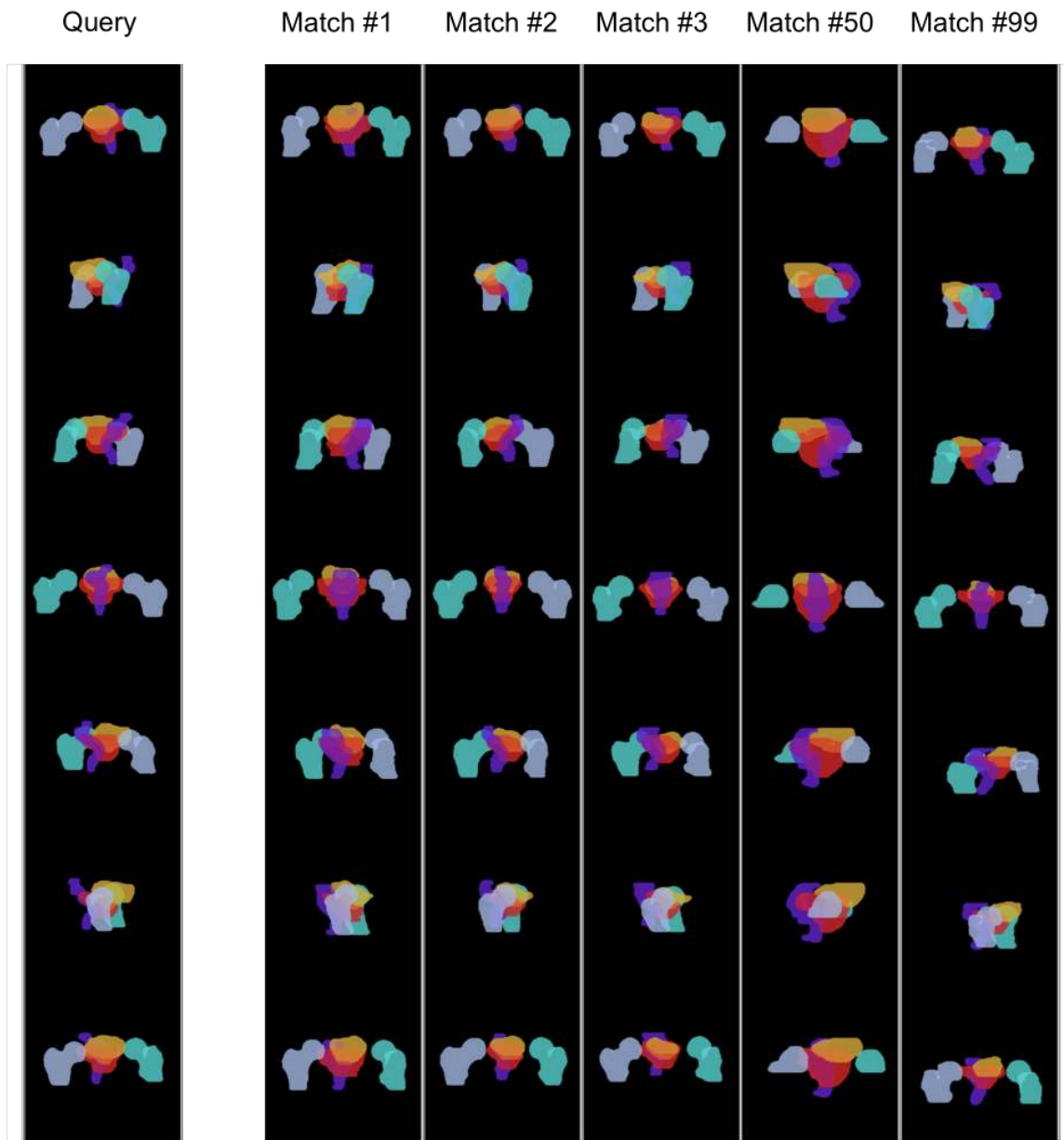
contours. *Figure 9* illustrates an example for one beam angle, a query case and the varying matched images in descending order of mutual information score.



**Figure 9** Example of a query image and various matched images

Projection image matching example: (a) Query image, (b) Best match (MI = 2.8), (c) Poor match (MI = 1.6), (d) Worst match (MI = 0.8). Planning target volume (red), bladder (yellow), rectum (dark blue), femoral heads (cyan and blue-gray).

*Figure 10* illustrates a visual comparison of the seven beam angles for a given query case matched against several other cases. A qualitative visual inspection of relative positions of PTV, bladder, rectum, and the femoral heads indicate that this simple average of the seven individual MI values do suggest the feasibility of using mutual information as a case-similarity approach to retrieve and rank matched cases.



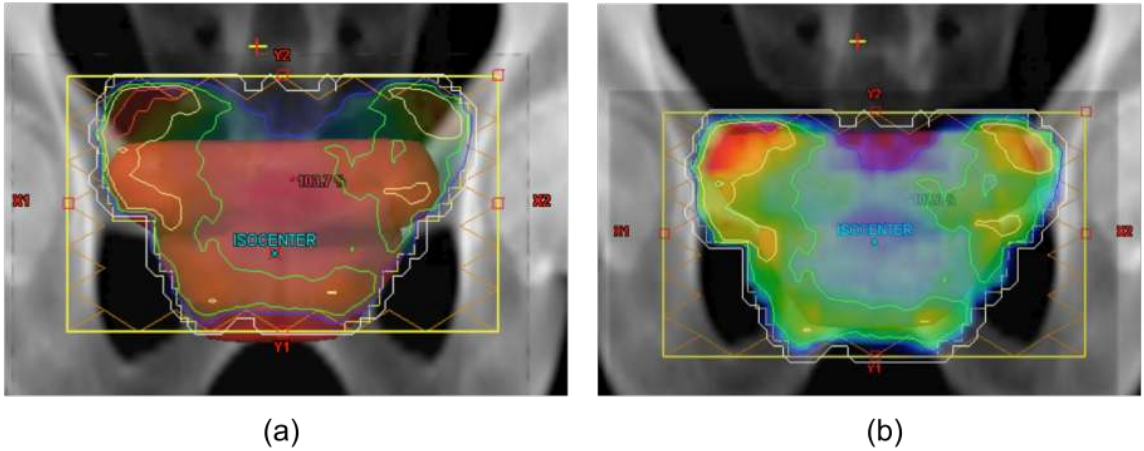
**Figure 10** Graphical user interface used to compare query and matched cases

A visual comparison of a query case matched against several other cases. The first, left column is the projection images of the query case. The top three best-matched cases (highest overall MI scores) are shown in columns two to four. A poor match is shown in column five, and the worst matched case is shown in column six.

### 3.2.3 Generating new treatment plans

Once the best-matching library reference case was found, the treatment parameters from that case were imported and utilized to develop a new treatment plan for the query case. Those treatment parameters included the beam geometry specifications, the fluence maps, and the final structure constraints and weights used to optimize the matched case during IMRT planning.

After the best-matched case is identified and selected, the fluence intensity maps from the matched case are imported into the query case. Since the fluence maps are the prior solutions to the optimization from an existing matched case, we can bypass many of the early optimization iterations for the new case by importing fluence maps from a best-matched prior case. In general, the fluence maps are geometrically aligned with the centroid of the planning target volume, since the majority of the radiation intensity is delivered to the tumor volume. However, we found that these imported fluence maps from the matched case often do not exactly align with the PTV contour, and it was thus necessary to apply a deformable transformation to each fluence map in order to improve the alignment of the radiation fluence with the target anatomy. *elastix*, a deformable registration toolkit using b-splines, was utilized to perform minor scaling and translations to improve alignment of the imported fluence to the planning target volume (PTV) of the query case (See *Figure 11*) [57].



**Figure 11** Fluence maps, before and after registration with the PTV

Imported fluence map overlay on PTV: (a) Before registration, fluence map is shown as contour lines, and (b) After registration, fluence map is shown as color wash. After deformable fluence registration, the resulting dose distribution will likely still need further minor optimization.

Under the assumption that the reference and query case are similar, the final constraints and priorities used to previously manually optimize the reference case would also likely almost optimally maximize the query case. Thus, the final step consisted of further optimizing the query case, from the starting point of the elastically deformed fluence maps, using the constraints and priorities from the reference case. The optimization was initiated and allowed to run for a total of 50 iterations (approximately 2-3 minutes on the Eclipse treatment planning system) without any manual intervention. Note that the prescription doses and constraints from the reference case were appropriately scaled to match the query case.

### 3.2.4 Dose-volume constraints

Table I depicts the prostate IMRT protocol used at our institution to specify the dose-volume constraints for critical structures. The protocol is similar to those

established by the Radiation Oncology and Therapy Group (RTOG) for prostate IMRT clinical trials [58].

**Table I Duke University’s prostate IMRT protocol with dose-volume constraints**

Critical Structure	Dose (Gy)	% Volume
Bladder	70	20
Bladder	62.4	30
Bladder	39	50
Rectum	98% of Rx dose	0
Rectum	70	20
Rectum	56	30
Rectum	39	50
Lt Femoral Head	50	0
Rt Femoral Head	50	0
Penile bulb	30	15
Small bowel	45	0

\* All constraints are based on target dose of 78 Gy.

The typical clinical practice strives to deliver a prescription dose (or higher) to at least 95% of the tumor volume while not exceeding normal structure dose constraints.

In this study, the quality of each of treatment plan was evaluated by comparing the dose volume histograms of the new semi-automated plan to that of the original plan developed manually by an expert (human) planner for the query case. We considered a plan to be ‘high quality’ if it both i) met the 95% dose to PTV prescription, and ii) reduces normal structure dose as much as possible even below guidelines set forth for bladder, rectum and femoral heads. That is, any further reduction would compromise the PTV coverage. A comparison of the various DVH cut-points considered the dose to



percent volume coverage ( $D_x$ , which is the dose to the highest  $x\%$  of volume) for the PTV and two normal structures (i.e. bladder and rectum). The cut-points for the PTV were  $D_{98}$ ,  $D_{95}$ , and  $D_1$ . Specifically,  $D_1$  was used to quantify the maximum target dose. For both the bladder and the rectum, the respective cut-points at  $D_{50}$ ,  $D_{30}$ , and  $D_{20}$  were evaluated.

In this initial study, ten query cases were randomly selected from the knowledge base of 100 cases. Only the constraints for five major structures were selected and imported, including: PTV, bladder, rectum, and the two femoral heads. Each plan was normalized to deliver the prescription dose to approximately 97 - 98% of the planning target volume. Using the most similar matched case for ten query cases randomly selected from the knowledge base of 100 cases, a total of ten new treatment plans were developed in this manner.

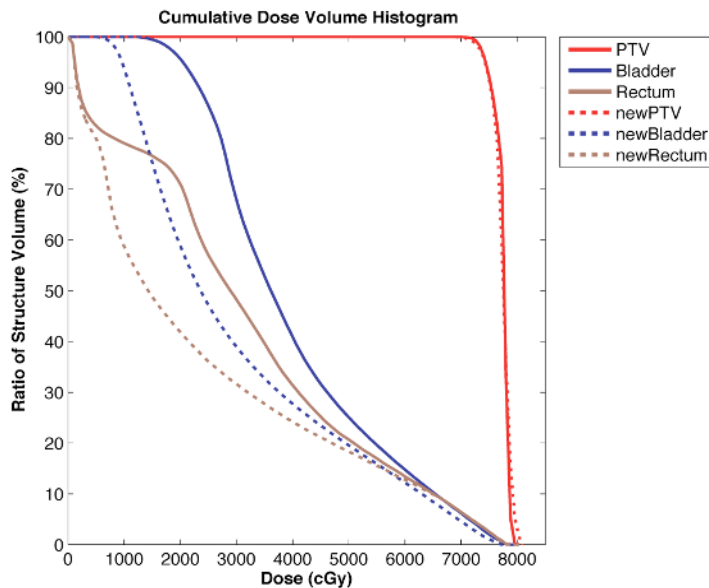
### 3.3 Results

While the results of all ten query cases are summarized, three of the ten cases have been selected to illustrate, in greater detail, the comparison between the new semi-automatically generated plan and original manual plan DVHs. Treatment plans were evaluated based on the relative percentage differences at specific DVH points:

$$\% \text{ diff} = 100 * (D_{vol - new \text{ plan}} - D_{vol - original \text{ plan}}) / D_{vol - original \text{ plan}} \quad (4)$$

A negative percentage difference indicates that the new plan has reduced dose as compared to the original plan. For a critical structure, such a negative percentage difference indicates an improvement, that is greater dose sparing. The DVH of the femoral heads are not plotted to make it easier to visualize the PTV, bladder and rectum. *Figure 12* is the DVH for case #8, which is one of the best of ten cases, in that it shows a very comparable PTV dose distribution between the new plan (dashed lines) and the

original plan (solid lines) and considerably greater dose sparing to both bladder and rectum.

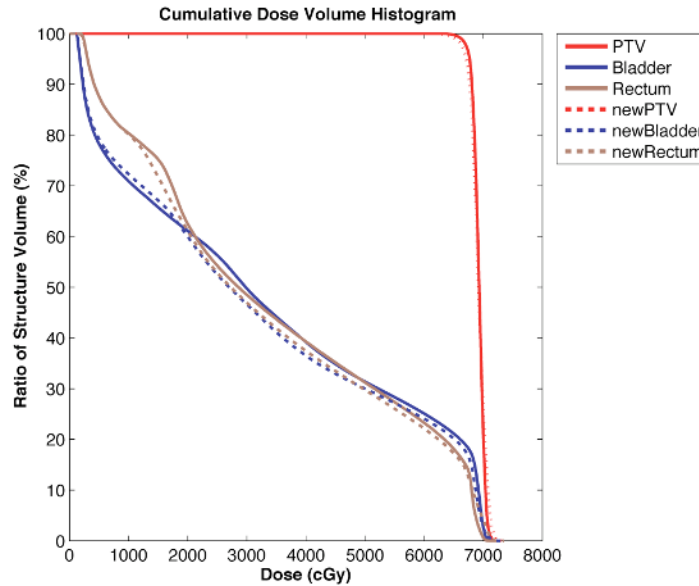


**Figure 12** Comparison of DVH for the new plan vs. original plan (case #8)

A comparison of DVHs for the new plan (dashed line) as compared to the original plan (solid lines). As one of the best of ten cases, case #8 demonstrates comparable PTV coverage and considerably greater dose sparing to both rectum and bladder.

For case #8, the PTV coverage in the new plan is approximately equal to the original plan; relative percentage differences at  $D_{98}$ ,  $D_{95}$ , and  $D_1$  were -0.5%, -0.1%, and 1.4% respectively. For the rectum, the new plan consistently demonstrated large additional dose sparing, as shown by the dashed curve of the new plan consistently shifting to the left of the solid line of the original plan. For example the new plan cut dose at  $D_{50}$  in half from 39.1% to 19.3%. Similar dose savings were also demonstrated in the bladder. Detailed cut-point values for all cases are reported below in tables. In summary, case #8 illustrates that the new plan has very comparable PTV coverage and substantial dose sparing to both the bladder and the rectum.

The DVH plot for case #7 is shown in *Figure 13*, and is representative of the majority of the ten cases, with very comparable dose distributions for the PTV, bladder and rectum.

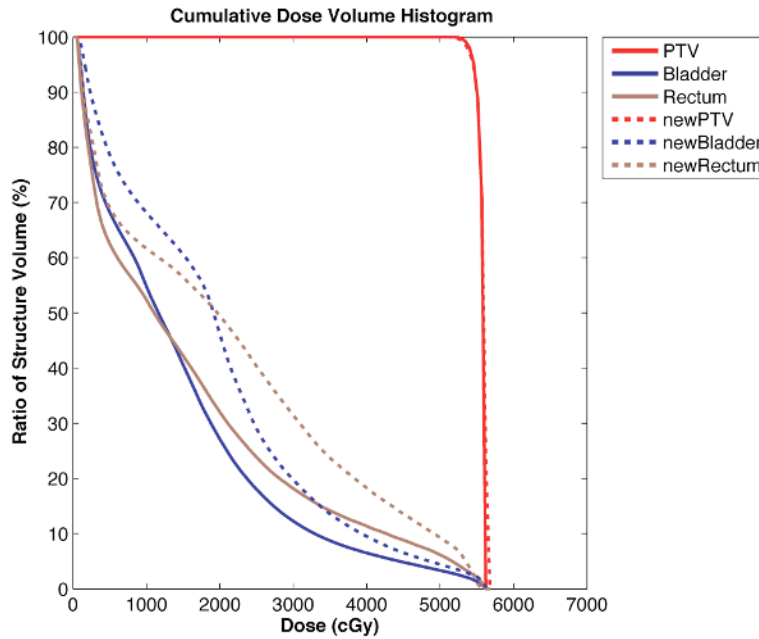


**Figure 13** Comparison of DVH for the new plan vs. original plan (case #7)

A comparison of the DVH for case #7, which is representative of the majority of the ten cases. This plot shows that dose to the PTV, rectum and bladder in the new plan (dashed) are very comparable to the original plan (solid).

Nine of the ten new plans had comparable DVH results relative to the original clinical plan. *Figure 14* illustrates the only case (case #5) which had comparable PTV coverage, but higher dose to both bladder and rectum as compared to the original plan. The PTV results are very comparable, with relative percentage differences of -0.5%, -0.2%, and 1.0% at  $D_{98}$ ,  $D_{95}$ , and  $D_1$ . However, the bladder dose at  $D_{20}$ ,  $D_{30}$ , and  $D_{50}$  values indicated decreased sparing by 25.9%, 31.2% and 64.5% respectively. Similarly, the dose to the rectum indicated decreased sparing by 36.4%, 46.5% and 78.0% respectively. Although there is substantially higher relative percentage dose to both the rectum and the bladder dose in the new plan, note that in terms of absolute dose-to-volume cut-

points, the critical structure doses would still be clinically acceptable since they are far below the dose constraints set forth in the protocol (Refer to Table I).



**Figure 14** Comparison of DVH for the new plan vs. original plan (case #5)

A comparison of DVH for the PTV, rectum and bladder of the original plan (solid) and the new knowledge-based plan (dashed) illustrates comparable PTV coverage, but higher dose to both bladder and rectum. However, these normal structure dose are still clinically acceptable since they are far below the dose-volume constraints set forth in the protocol.

A detailed comparison of the relative percentage differences (equation 2) at specific dose-to-volume values to the PTV for all ten new plans versus the original plans is shown in *Table II*. A negative value indicates that the new plan has less PTV coverage as compared to the original plan. The average percentage differences for the ten cases (mean  $\pm$  standard deviation) at  $D_{98}$ ,  $D_{95}$  and  $D_1$  are  $-0.03\% \pm 2.1\%$ ,  $0.62\% \pm 1.7\%$  and  $2.5\% \pm 2.4\%$  respectively. On average, the new knowledge-based plan is capable of

achieving very comparable PTV coverage to the original plan (within 2% as evaluated for both  $D_{98}$  and  $D_{95}$ , as well as for  $D_1$ , the dose to the highest 1% of the PTV volume).

**Table II Percentage difference in PTV coverage between new and original plan**

Case	$D_{98}$	$D_{95}$	$D_1$
1	4.3	4.6	5.6
2	-1.4	0.6	7.7
3	-0.8	0.1	0.6
4	-0.4	-0.1	0.6
5	-0.5	-0.2	1.0
6	-0.4	-0.3	1.0
7	1.8	2.2	4.2
8	-0.5	-0.1	1.4
9	1.0	1.2	1.8
10	-3.5	-1.7	1.4
Mean	-0.03	0.62	2.5
Stdev	2.1	1.7	2.4

Note:  $\% \text{ diff} = 100 * (D_{vol - new \text{ plan}} - D_{vol - original \text{ plan}}) / D_{vol - original \text{ plan}}$

*Table III* shows the percentage differences between the new and the original plan for the rectum dose at  $D_{20}$ ,  $D_{30}$  and  $D_{50}$  for the ten cases. For the normal structures, a negative value indicates an improvement, in which the new plan has greater dose sparing to the rectum as compared to the original plan. For the ten cases, the mean and standard deviation for the rectum dose percentage differences for  $D_{20}$ ,  $D_{30}$  and  $D_{50}$  are  $1.8\% \pm 8.5\%$ ,  $-2.5\% \pm 13.9\%$  and  $-13.9\% \pm 23.6\%$  respectively. Three of the ten new plans (#2, 4, 8) have greater rectal dose sparing as compared to the original plan, while five (#1, 3, 7, 9, 10) are quite comparable to the original plan. Although the percentage

differences for  $D_{20}$ ,  $D_{30}$  and  $D_{50}$  for cases #5 and #6 appear high, the absolute dose values are still clinically acceptable since they fall below the dose constraints set forth in the original protocol. (Refer to Table I).

**Table III Percentage difference in rectum dose between new and original plan**

Case	$D_{20}$	$D_{30}$	$D_{50}$
1	5.9	5.0	-4.0
2	-11.1	-22.1	-35.8
3	2.4	2.7	-19.8
4	-3.2	-14.4	-35.5
5	36.4	46.5	78.0
6	16.3	14.1	23.7
7	1.5	-0.8	-1.6
8	-7.6	-22.1	-50.6
9	10.1	13.3	6.7
10	1.7	1.7	-8.2
Mean	1.8	-2.5	-13.9
Stdev	8.5	13.9	23.6

Note:  $\% \text{ diff} = 100 * (D_{\text{vol-new plan}} - D_{\text{vol-original plan}}) / D_{\text{vol-original plan}}$

Lastly, *Table IV* shows the percentage differences between the new and the original plan for the bladder dose at  $D_{20}$ ,  $D_{30}$  and  $D_{50}$  for the ten cases. A negative value indicates an improvement, in which the new plan has greater dose sparing to the bladder as compared to the original plan. For the ten cases, the mean and standard deviation for the bladder dose percentage differences for  $D_{20}$ ,  $D_{30}$  and  $D_{50}$  are  $-5.9\% \pm 10.8\%$ ,  $-12.2\% \pm 14.6\%$  and  $-24.9\% \pm 21.2\%$  respectively. Five of the ten new plans (#2, 4, 6, 8, 10) have substantially less dose to the bladder, while four (#1, 3, 7, 9) are very

comparable to the original plan. As previously mentioned, for case #5, the percentage differences for  $D_{20}$ ,  $D_{30}$  and  $D_{50}$  may appear to be high, but the absolute dose values are still clinically acceptable since they fall below the dose constraints set forth in the original protocol. (Refer to *Table I*).

**Table IV Percentage difference in bladder dose between new and original plan**

Case	$D_{20}$	$D_{30}$	$D_{50}$
1	4.8	4.9	1.8
2	-15.1	-28.0	-50.2
3	2.8	2.6	-3.2
4	-10.6	-19.6	-32.0
5	25.9	31.2	64.5
6	-27.7	-37.8	-57.3
7	1.6	-1.3	-6.6
8	-9.5	-18.5	-34.7
9	1.2	-4.0	-11.0
10	-0.2	-8.4	-31.0
Mean	-5.9	-12.2	-24.9
Stdev	10.8	14.6	21.2

Note:  $\% \text{ diff} = 100 * (D_{vol - new \text{ plan}} - D_{vol - original \text{ plan}}) / D_{vol - original \text{ plan}}$

### 3.4 Discussion

We have found that a new IMRT treatment plan can be developed semi-automatically by importing the set of fluences from a prior optimized case selected from a knowledge-base reference library, without having to start anew from a blank slate, as is typically done in manual planning.

### **3.4.1 Size of knowledge base**

To demonstrate the feasibility of using the knowledge-base approach to develop new treatment plans, our initial study randomly selected ten test cases from a database of 100 prior cases. Recognizing that the relatively small number of test cases may have limited statistical power, we are still encouraged by the potential of utilizing a knowledge-base treatment planning approach to produce clinically acceptable treatment plans. We hope to observe useful trends as we expand our database and continue to develop more treatment plans. As we continue to increase the size of the database, we intend to study the effect of the knowledge-base diversity on the overall system performance. For example, we did not include boost plans in our initial database; although we could construct a separate database of boost plans and develop treatment plans specific to boost treatments. In later chapters, we will present studies using additional independent cases from outside institutions to investigate inter-institutional variability of treatment planning, as well as the appropriateness of utilizing our knowledge base across institutions.

### **3.4.2 Improving case similarity retrieval**

Our study implemented MI as a case-similarity approach to search for a small set of best-matched cases from the entire database, but does not rely on MI as the sole determinant of dose-constraint satisfaction. Since the primary objective of treatment planning is to provide adequate PTV coverage, we have found that by further applying deformable registration to align the imported fluences of the matched case to the PTV of the query case, it effectively aids the optimizer by providing a good starting point to help successive optimization to proceed. We found that if we were to leave out the deformable registration of the fluence map with the PTV, the treatment plan does not yield adequate PTV coverage. The third and last step involves importing the



optimization weights from the matched case, which the optimizer then utilizes to satisfy the dose constraints. We found that by using this three-step process, the optimizer required very few steps to converge.

### **3.4.3 Assessing DVH plan quality**

We intend to investigate methods to correlate the similarity metric with the overall treatment plan quality. Currently, we chose to evaluate the plan quality based on the DVH cut points specified by the protocol at our institution. We compared each new semi-automated plan with the original manually developed plan at these specified normal structure constraints. There is still considerable disagreement about the specific values for dose/volume limits for normal tissue [59]. The RTOG Prostate Group Consensus recently reported for the rectum, 50 Gy  $\leq$  50% and 70 Gy  $\leq$  20%, and for the bladder, 55 Gy  $\leq$  50% and 70 Gy  $\leq$  30%. The limits for the femoral heads were set at <5% for 50 Gy, and for the small bowel, 0% for 52 Gy [60]. Similar dose/volume limits were recently reported in the recently published QUANTEC study [61, 62], for the rectum: 50 Gy < 50%, 60 Gy < 35%, 65 Gy < 25%, 70 Gy < 20% and 75 Gy < 15%. For the bladder, the reported limits are 65 Gy < 50%, 70 Gy < 35%, 75 Gy < 25%, and 80 Gy < 15%. The protocol at our institution specifies even more conservative normal tissue constraints (Refer to Table I). Thus, we chose not to assess DVH plan quality based on absolute differences from the normal tissue constraints, as these limits are typically already very generous, are population-based and are related to the clinical end-points corresponding to late tissue toxicities. Rather, we assessed the individual DVH plan quality of a new plan by a comparison with the original plan, as that best reflects the extent to which the new semi-automated plan can strive to meet the (human) achievable limits of dose sparing.

### **3.4.4 Reduction in Treatment Planning time**

We found that the knowledge-based treatment planning approach can substantially reduce planning time by skipping past all but the last few optimization iterations of the planning process. For each of the ten test cases, approximately 50 optimization iterations were required to meet the optimization objectives, taking approximately two to three minutes on the Eclipse treatment planning workstation. Additionally, these iterations were allowed to proceed automatically, with no human input other than initiating the optimization. It is foreseeable that several treatment plans optimized for different clinical objectives (i.e. rectal sparing) can be quickly generated in this manner and made available for the clinician to choose from. Knowledge-based treatment planning offers the potential to improve the efficiency of the treatment planning process while maintaining high quality.

## 4. Evaluation of Treatment Plan Quality

### 4.1 Introduction

The inverse problem in IMRT treatment planning requires solving for the set of fluence intensity patterns that will satisfy, for a given patient geometry and the selected beam geometry, the prescribed target volume dose and the dose-limit constraints to the normal tissue. Without an exhaustive search of the solution space of all possible fluence patterns, it is not feasible to determine what the global minimum solution would be for a specific patient. It is thus very difficult to estimate *a priori* the maximum extent of dose reduction possible to critical structures (i.e. rectum, bladder) for specific patients. In fact, it is possible that there may exist several local minimum solutions that would in fact be clinically acceptable. In practice, what is selected as a clinically acceptable treatment plan may not in fact be the mathematically optimal plan. Additionally, since the established dose-limit constraints are population-based, it can be challenging to make patient-specific decisions about the trade-off between providing adequate dose coverage to the PTV while minimizing the extent of dose delivered concomitantly to each of the critical structures.

In this chapter, we seek to understand the underlying relationship between the mutual information metric described previously and the resulting clinical treatment plan quality. Whereas we previously considered only the best matching case, here we will explore a full range of matches and investigate the effect on clinical plan quality. In any treatment plan, we seek to maximally achieve the prescription dose to the planning target volume, while minimizing to the greatest extent possible, the dose to critical structures. We propose a simple metric to quantify the extent of PTV coverage and dose sparing of a given plan, and we use this metric to evaluate and compare the plan quality of a range of different knowledge-based treatment plans.

## **4.2 Material and Methods**

### **4.2.1 Knowledge-base reference library**

For the study in this chapter, we increased the size of the database to a total of 250 prostate cancer IMRT cases, all of which were previously planned, reviewed and delivered within the Duke clinic. IRB permission was obtained for the use of these existing cases, and the cases were anonymized prior to use. As before, each treatment plan consists of the CT dataset, tumor/target volume and normal structure contours, beam geometry specifications, beam intensity (fluence) maps, and dose-distributions. Similarly, the standard seven co-planar beam angles used are as follows: 25°, 75°, 130°, 180°, 230°, 280°, and 330°. We implemented the same mutual information algorithm for identifying similar patient cases based on matching 2D projections of the structure contours.

### **4.2.2 Generating new treatment plans**

For this study, 20 unique query cases were randomly selected from the knowledge base of 250 cases. The case-similarity algorithm matched the structures of each query case, per beam angle, against the entire database. For each of the twenty randomly selected query cases, we selected five matched cases, ranging from high to low overall MI score. The selected matched cases spanned the full range of overall MI scores for each query case, reflecting a range of contour similarity. The treatment parameters from the matched case were adapted to generate a total of 100 new treatment plans.

Each new plan was normalized to deliver the prescription dose to approximately 97 - 98% of the planning target volume. In this study, we utilized only the set of seven fluence maps from the matched case, deformed to fit the query case. Unlike the previous study, we did not adapt the final structure constraints and weights originally

used to optimize the matched case. By excluding the final optimization weights, we developed ‘pre-optimized’ treatment plans that more accurately reflect the plan quality that is obtained from only using the fluence maps. We could later generate optimized treatment plans by importing the optimization weights from the matched case, and thus compare the plan quality of pre- and post-optimization.

### 4.2.3 Defining a treatment plan quality metric

Assessing DVH plan quality at specific dose-volume cut-points (i.e.  $D_{20}$ ,  $D_{30}$ ,  $D_{50}$ ) can be highly subjective, even if there is agreement about the specific dose-volume constraints. How does a radiation oncologist evaluate the tradeoff in PTV coverage versus the dose sparing of critical structures for a specific patient? This assessment is at best a qualitative, and very subjective clinical decision process. We set out to establish a simple plan quality metric to evaluate and compare different treatment plans based on the DVH. Our hypothesis is that the plan quality of a matched case is correlated with the degree of case similarity as determined by the overall MI score. That is, we seek to establish whether it is possible to use mutual information as a surrogate for the DVH plan quality of a knowledge-base IMRT plan.

Keeping in mind that the primary goal of every treatment plan is to achieve the maximum coverage of the planning target volume at the prescribed dose, and that dose sparing to each of the normal structures should be achieved to the maximum extent possible, we propose a simple DVH plan quality index (PQ) based on the weighted-sum of the areas under the curve (AUC) of the PTV, rectum and bladder, as follows:

$$PQ = w_{PTV} * AUC_{PTV} + w_{rectum} * (1 - AUC_{rectum}) + w_{bladder} * (1 - AUC_{bladder}) \quad (5)$$

where  $w_{PTV}$  is the weighting factor of the PTV, and  $AUC_{PTV}$  is the area under curve of the normalized cumulative DVH for the PTV. The weighting factors  $w_{rectum}$  and  $w_{bladder}$  and the area-under-the-curve values  $AUC_{rectum}$  and  $AUC_{bladder}$  are similarly defined. These

weighting factors are arbitrary and reflect the relative priority weighting scheme that a site may choose to establish as a general guideline for their clinical practice. We chose to ignore the femoral heads in the plan quality metric since most IMRT plans utilize beam angles that more than adequately spare the femoral heads. In this study, we explored how different PTV weights can influence the rank-order of the DVH plan quality indices for the entire database. We sought a weighting scheme that sufficiently reflects the relative priority of PTV coverage and the dose sparing of the critical structures that is desired at the Duke clinic.

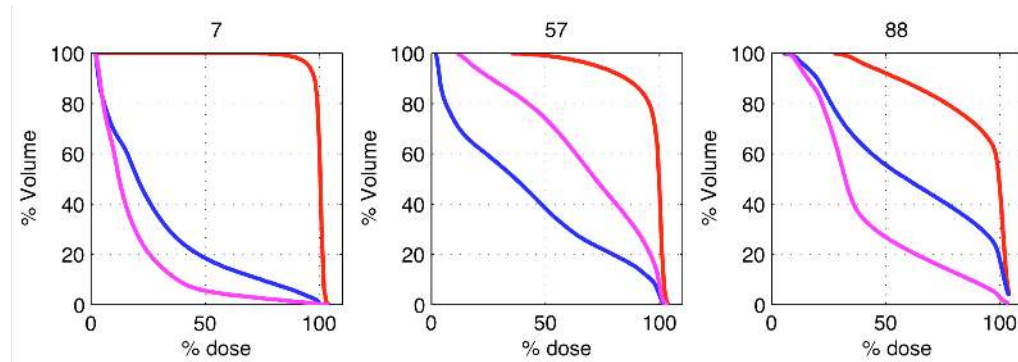
#### **4.2.4 Evaluating pre-optimized plans using the plan quality index**

For this study, we developed 100 new pre-optimized plans (20 queries x 5 matches) as described before. We then calculated the DVH plan quality score for each of the new plans. Since we were interested in determining whether we could use the overall MI score as a surrogate for DVH plan quality, we then compared the plan quality metric for the 100 pre-optimized treatment plans to the overall MI score for each match. It is important to note that we chose to compare the plan quality of the pre-optimized plans against the overall MI score, rather than to use the post-optimized plans, since the pre-optimized plans do not reflect the successive improvements in plan quality due to the last few iterative optimizations. Lastly, we also investigated the change in plan quality due to the optimization process for a small subset of 20 of the 100 cases, by comparing the plan quality score for pre-optimized plans and post-optimized plans.

### **4.3 Results**

For a single query case, there is a range of potentially acceptable matched cases, ranging from high similarity (high MI score) to low similarity (low MI score). *Figure 16* illustrates, for a given query case, the resulting DVH of the pre-optimized plans obtained from adapting parameters for a case with high similarity, medium case

similarity, and low case similarity. We seek to investigate whether a matched case of high similarity will necessarily produce a final post-optimized plan that tends to be high quality. Similarly, a matched case of fair or poor similarity, as indicated by a lower MI score, is expected to produce a post-optimized plan that is lower in plan quality than one produced from a matched case having a higher MI score.

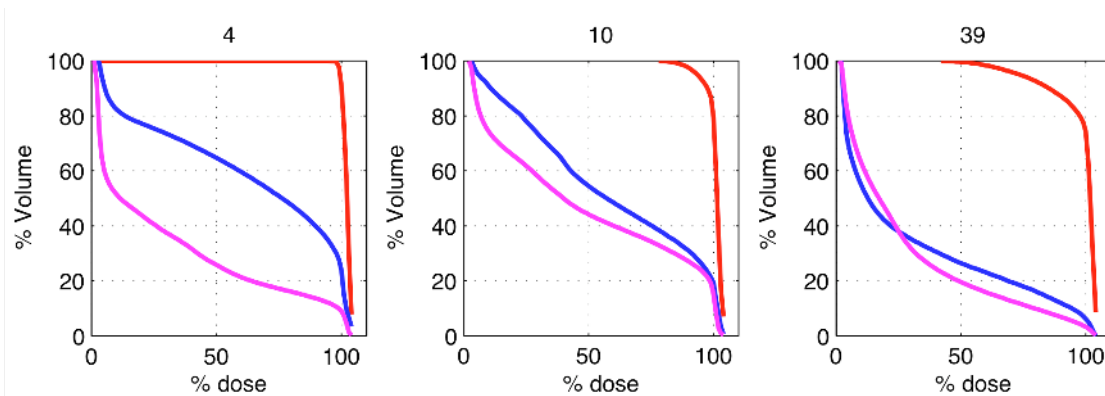


**Figure 15** Range of DVH plans using matched cases of high to low similarity

PTV is shown in (red), bladder (blue) and rectum (cyan). Range of possible DVH plan quality for a given query case, corresponding from high plan quality (left) to low plan quality (right).

Consider the scenario for which the plan quality score is derived from the non-weighted sum of the AUCs for PTV, rectum and bladder, that is  $w_{\text{PTV}} = w_{\text{rectum}} = w_{\text{bladder}} = 1$ , thus,

$$PQ = AUC_{\text{PTV}} + (1 - AUC_{\text{rectum}}) + (1 - AUC_{\text{bladder}}) \quad (6)$$



**Figure 16** Example of un-weighted plan quality scoring method, rank-ordered.

PTV is shown in (red), bladder (blue) and rectum (cyan). The three cases are rank-ordered based on  $AUC_{PTV}$  where the PTV coverage is highest for case ranked #4, and lowest for case ranked #39. However, the respective plan quality scores for cases #4, #10 and #39 are as follows: 0.753, 0.683 and 0.910.

*Figure 16* illustrates the un-weighted plan quality scoring result by a comparison of three cases, ranked in terms of descending  $AUC_{PTV}$  as #4, 10 and 39 out of 99. Note that the last case actually has the highest plan quality score of 0.910, but its  $AUC_{PTV}$  value is ranked lower than 38 other cases. Its higher overall PQ score is due to the contribution of the low AUCs for the rectum and bladder. Conversely, for the case ranked #10 out of 100, its overall PQ score of 0.683 is barely better than the mean PQ score of 0.646. It is clear that without any weighting of the structures, the plan quality score does not reflect the typical assessment of treatment plan quality, where adequate PTV coverage is of utmost importance.

We introduce a weighted-scheme to calculate the plan quality score. We recognize that the exact weighting scheme is arbitrary, and will likely vary between individual clinician’s preference, as well as institutional practice. However, we hope to implement a simple weighting scheme that could better differentiate and reflect a more appropriate rank-order of cases based on the relative importance of PTV coverage over



rectum and bladder dose sparing. We experimented with two weighting schemes, A and B:

$$\begin{aligned} \text{A)} \quad & w_{\text{PTV}} = 5, w_{\text{rectum}} = 1, w_{\text{bladder}} = 1 \\ & \text{PQ} = 5 * \text{AUC}_{\text{PTV}} + (1 - \text{AUC}_{\text{rectum}}) + (1 - \text{AUC}_{\text{bladder}}) \end{aligned} \quad (7)$$

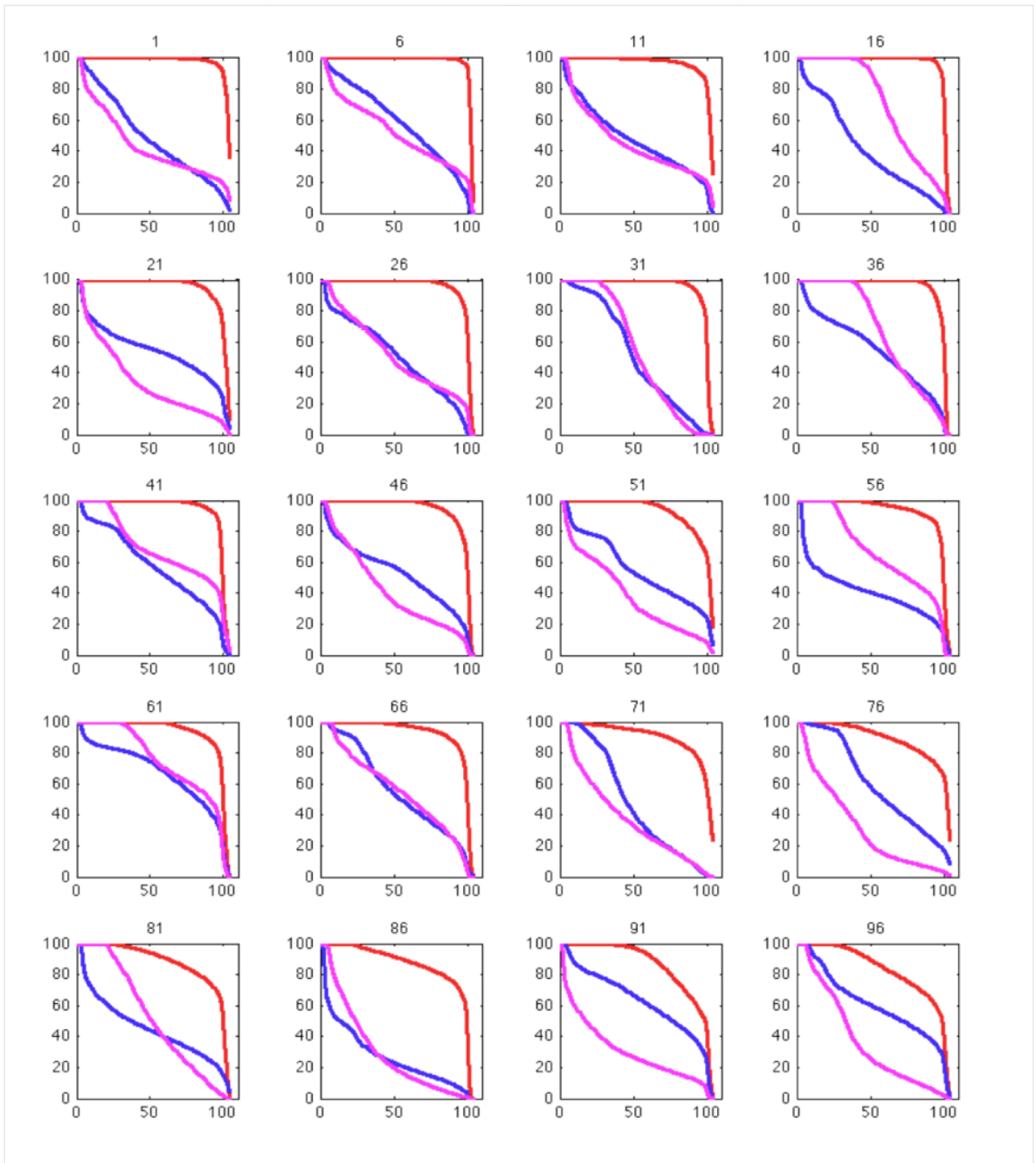
$$\begin{aligned} \text{B)} \quad & w_{\text{PTV}} = 10, w_{\text{rectum}} = 1, w_{\text{bladder}} = 1 \\ & \text{PQ} = 10 * \text{AUC}_{\text{PTV}} + (1 - \text{AUC}_{\text{rectum}}) + (1 - \text{AUC}_{\text{bladder}}) \end{aligned} \quad (8)$$

**Table V** illustrates the effect on  $\text{AUC}_{\text{PTV}}$  rank order for different weighting schemes for the three cases previously mentioned. With a PTV weight of ten, the plan quality score reflects the desired rank ordering of cases, which places priority on the PTV coverage.

**Table V** Effect of  $\text{AUC}_{\text{PTV}}$  rank order for different weighting schemes

Rank	No weighting	Rank	PTV 5 1 1	Rank	PTV 10 1 1
4	0.753	5	0.964	3	0.982
10	0.683	16	0.892	12	0.913
39	0.910	13	0.909	22	0.865

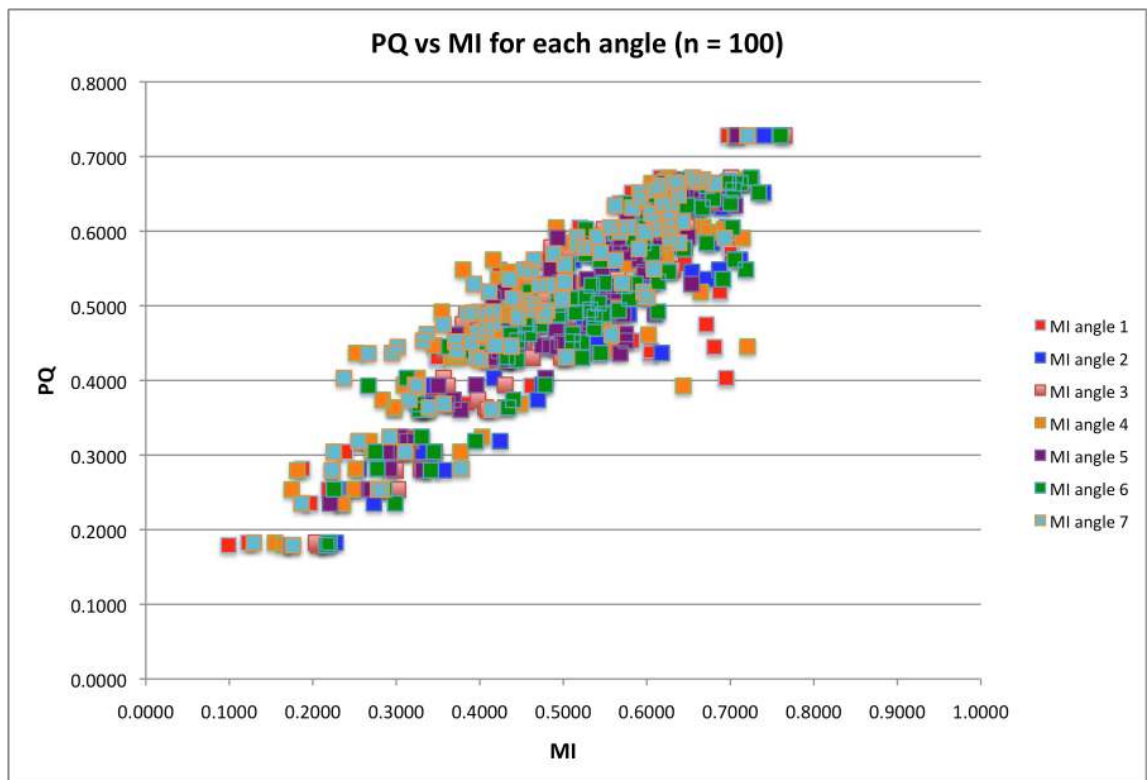
*Figure 17* illustrates the range of DVH plan quality for every 5<sup>th</sup> cases ranked #1 to #96 using a  $w_{\text{PTV}}$  equal to 10. Visual inspection of these DVH plots establishes that the quality of the PTV coverage decreases as the rank order decreases. The overall plan quality score using this weighting scheme places higher priority on the PTV, and equal weighting for the rectum and bladder.



**Figure 17** Rank order of 20 cases from #1 to #96 for PTV weighted scheme

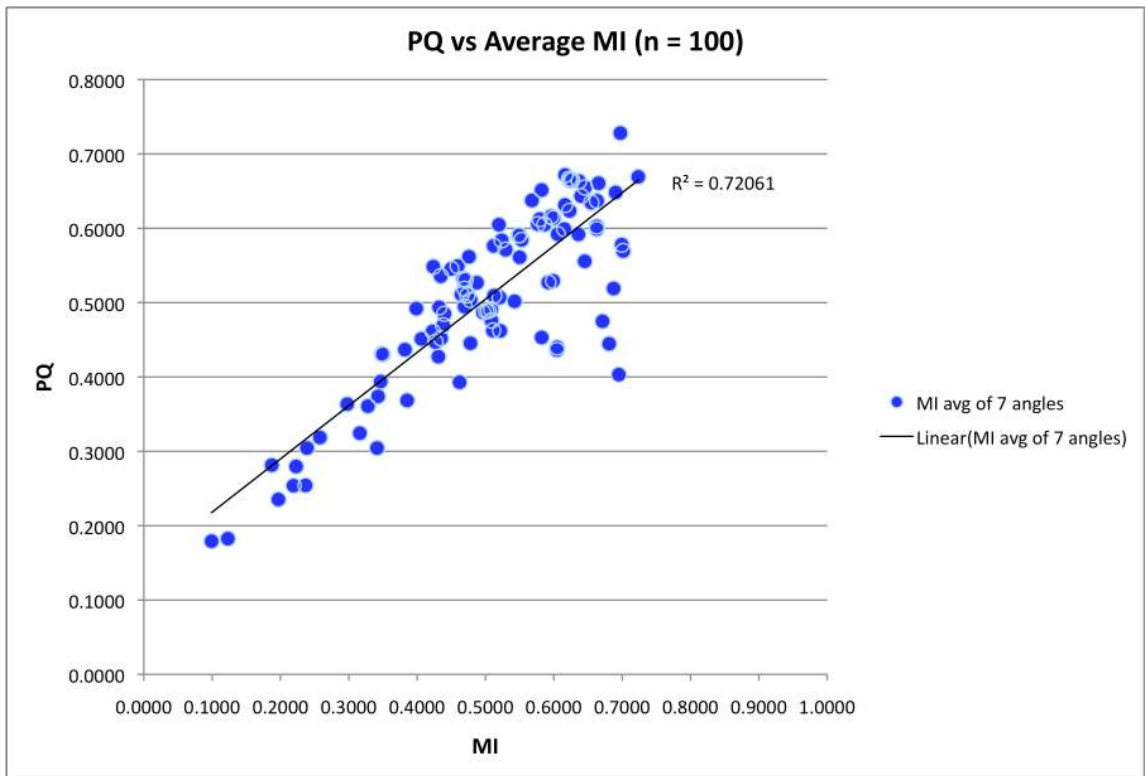
Showing every 5<sup>th</sup> case, the resulting ordering now corresponds to expected clinical plan quality.

This PTV weighting scheme was implemented in order to establish a metric to evaluate the plan quality of our knowledge-based plans. In particular, we are interested in investigating whether the overall mutual information score can be used as a surrogate of DVH plan quality. Note that for the purpose of this evaluation, we normalized the MI scores, where the maximum value of 1.0 corresponds to the self-similar comparison. *Figure 18* is a plot of the plan quality metric versus the mutual information score for each of the seven angles for the 100 generated pre-optimized treatment plans. The plot suggests that as the MI score increases, implying higher case similarity between the query case and the matched case, the DVH plan quality similarly increases.



**Figure 18** Plan quality versus MI for each angle

Note that for this plot, the x-axis represents the normalized MI value, where a maximum value of 1.0 corresponds to the self-similar value for each angle.



**Figure 19** Plan quality versus average normalized MI score

Note that for this plot, the x-axis represents the normalized MI value, where a maximum value of 1.0 corresponds to the self-similar value for each angle.

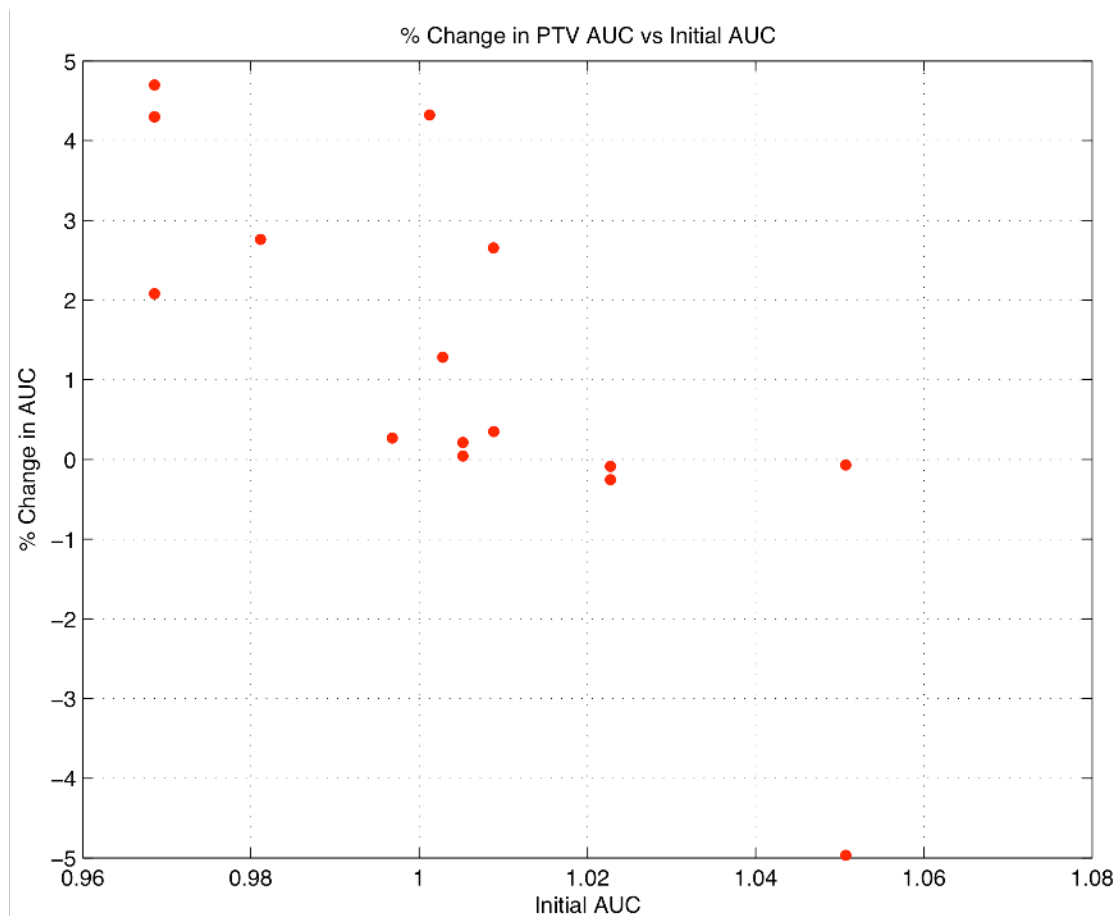
In *Figure 19*, the plan quality for each of the pre-optimized query cases is plotted as a function of the average normalized MI score for the seven beam angles. A linear fit for these 100 cases is performed, and the  $R^2$  correlation value is found to be 0.72. This result suggests that as the average MI score increases, the overall plan quality increases as well. Thus, the normalized MI score can potentially be used as a surrogate for overall plan quality of a pre-optimized treatment plan.

Although the overall plan quality is highly correlated with case similarity, we find that it is possible that a matched case of moderately high similarity, but not the highest similarity, can yield a final post-optimized plan that is superior to the best-matched case. This is due to the finding that the final plan quality is determined not

simply by the level of the initial matched case similarity, but also by the subsequent processes of adapting fluences and optimization weights from the matched case. However, based on our initial study using 20 query cases, we do establish that a case with very low case similarity simply does not appear to lead to pre-optimized plans with high plan quality. The successive steps of adapting the optimization weights is expected to improve plan quality; however, the fluences from a poorly matched case for which the structure geometry may be too different from the query case may simply be inadequate to produce a plan of high quality.

Lastly, we compared the plan quality of the pre-optimized plans and the post-optimized plans for sixteen randomly selected query cases. The optimization weights from the matched cases were imported into each query case. This final step consisted of optimizing the query case using the constraints and priorities from the reference case. The optimization was initiated and allowed to run for a total of 50 iterations (approximately 2-3 minutes on the Eclipse treatment planning system) without any manual intervention. Note that the prescription doses and constraints from the matched case were appropriately scaled to match the query case. We evaluate both the actual change in overall PQ as well as the individual AUC changes for each structure for the pre-optimization and post-optimization plan.

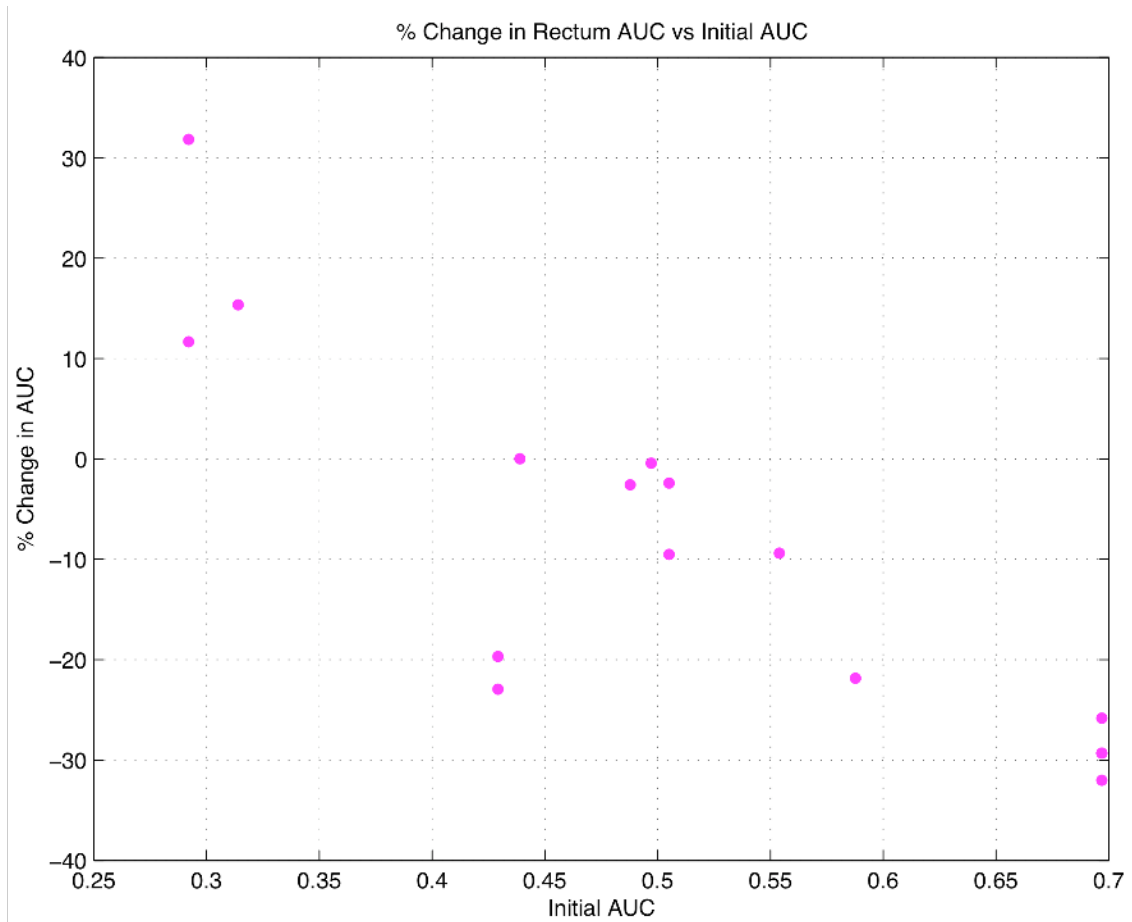
The percent change in PTV AUC is plotted for pre- and post-optimization plans is plotted in *Figure 20* against the initial pre-optimized PTV AUC. Note that optimization resulted in greater than or equal PTV AUC for fourteen of the sixteen plans. This plot suggests that the greatest percent improvement in PTV plan quality occurs for pre-optimized plans of lower PTV AUC values. The mean percentage change for PTV AUC is +1.37%, with a maximum percentage improvement of +4.7% for cases with the lowest initial PTV AUC.



**Figure 20** Percent change in PTV AUC versus the initial pre-optimized PTV AUC

We next analyzed the rectum AUC for pre- and post-optimization plans for sixteen cases. A decrease in AUC for a critical structure actually indicates an improvement, that is, greater dose sparing to the critical structure. With the exception of three cases (#6, #7, and #16) that already had very low area and thus excellent dose sparing, 13 of the 16 optimized plans resulted in greater or equal dose sparing to the rectum. The percent change in rectum AUC is plotted in *Figure 21* against the initial pre-optimized AUC. This plot suggests that the greatest percent improvement in plan quality occurs for pre-optimized plans with higher area and thus poor dose sparing to

begin with. The mean percentage change for rectum AUC is -9.15%, with a maximum dose reduction of 32% for cases with poor initial AUC value of 0.7.

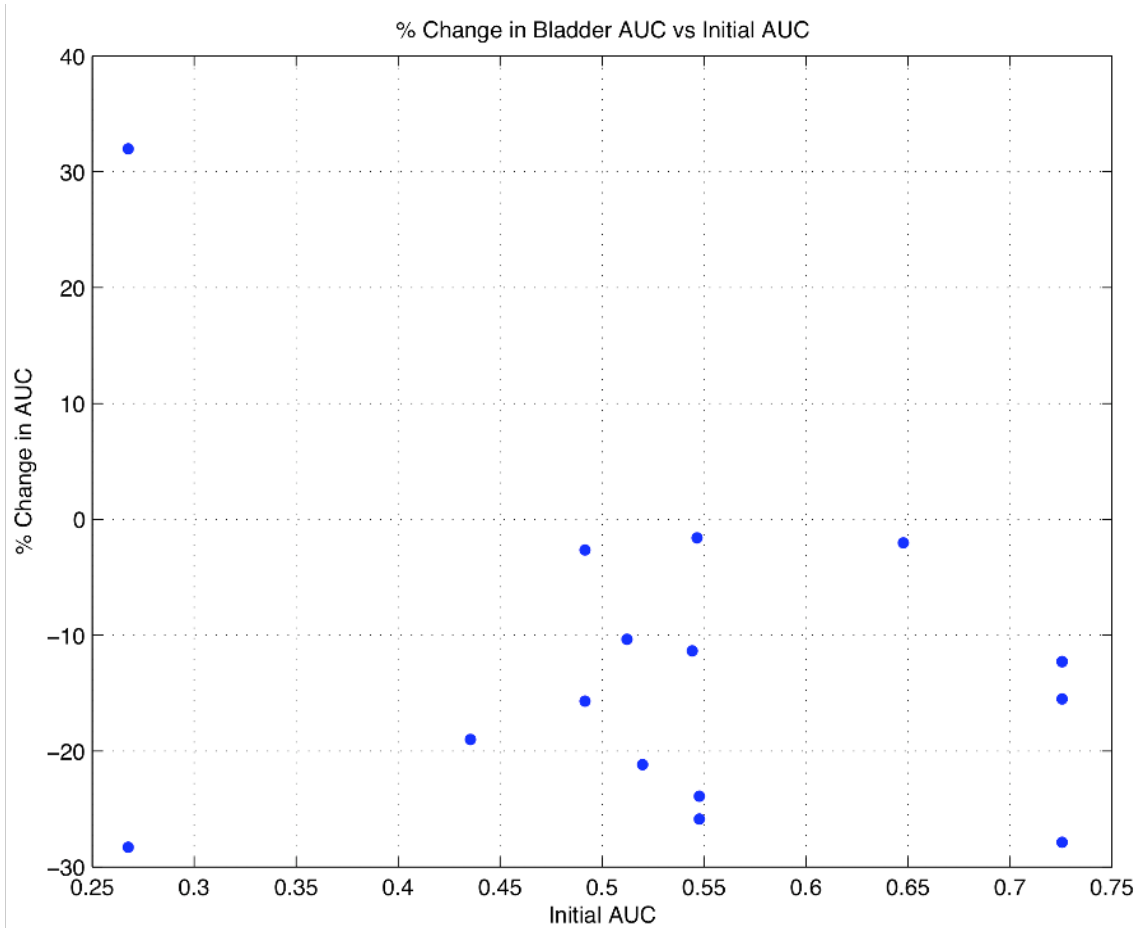


**Figure 21** Percent change in rectum AUC versus the initial pre-optimized AUC

Note: A negative percent change in AUC for a critical structure indicates an improvement, that is greater dose sparing

For the bladder, the percent change in AUC is plotted in *Figure 22* against the initial pre-optimized AUC. With the exception of one case (#6), all of the cases saw greater or equal dose sparing to the bladder after the optimization process. The mean percentage change for bladder AUC is -12.4%, with a maximum dose reduction of 28%

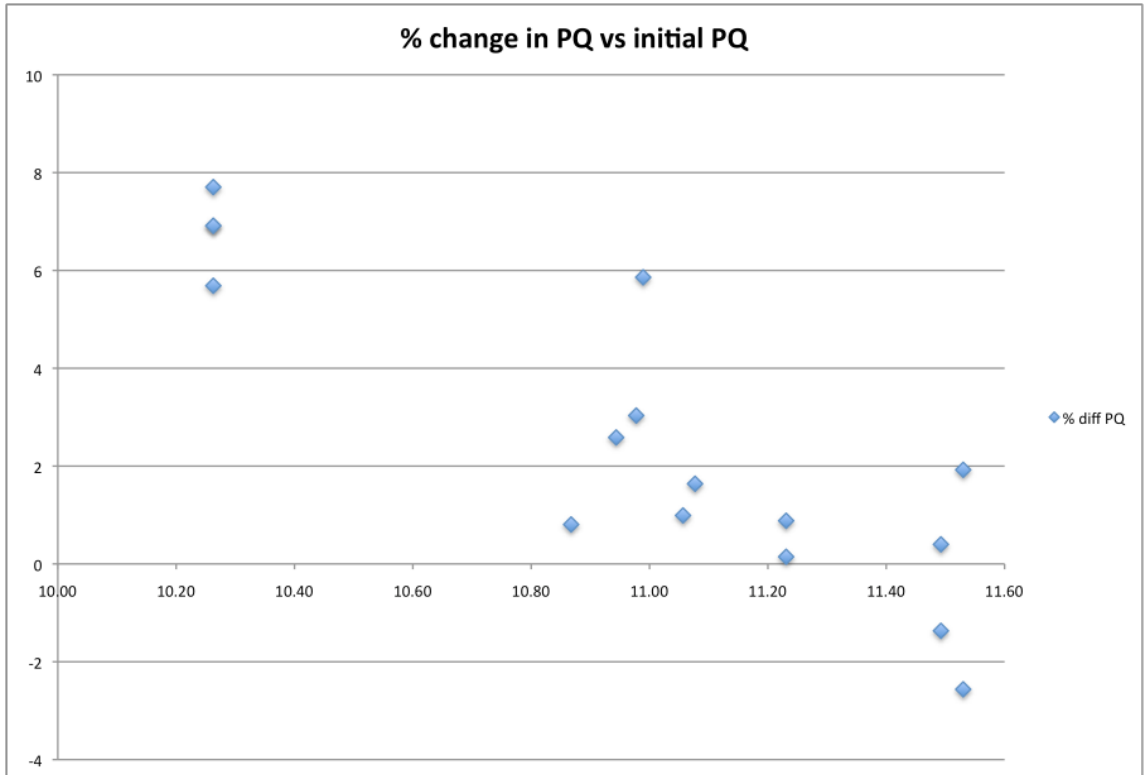
for case #7. For the bladder, there does not appear to be a dependency between the final optimized AUC and the initial pre-optimized AUC.



**Figure 22** Percent change in bladder AUC versus the initial pre-optimized AUC

In summary, we will consider the changes in overall plan quality score, which reflects the AUC for all three structures, and compare this for pre- and post-optimized plans.





**Figure 23** Percent change in overall plan quality versus the initial PQ

The percent change in overall plan quality is plotted against the initial pre-optimized PQ in *Figure 23*. This plot suggests that the greatest percent improvement in plan quality occurs for plans with lowest initial PQ, where the coefficient of determination ( $R^2$ ) is found to be 0.77. Most plans saw a slight improvement in plan quality of a few percentage points, with the mean percentage gain of +2.6%, and a maximum improvement of 7.7% for cases with the lowest initial PQ. Only two cases of the sixteen demonstrated a decrease in overall plan quality after the optimization process. What is interesting to note is that these two plans had among the highest plan quality scores to begin with. For such plans with very high PQ to begin with, the substitution of prior optimization weights from the matched cases may actually lower the final overall plan slightly. These cases suggest that it may not be necessary to

continue with the optimization process if the pre-optimized plan quality is already at a level deemed to be clinically acceptable.

#### **4.4 Discussion**

We have found that a simple plan quality metric based on the sum of the areas under the curve for the PTV, rectum and bladder of the DVH plot can be a useful summary measure for evaluating the quality of an individual plan, as well as for the comparison of a collection of treatment plans. Specifically, we established that a PTV-weighted scheme could be used to successfully rank order the plan quality of a collection of knowledge-based plans. The developed weighting scheme reflects the priority of requiring adequate PTV coverage first and foremost, followed by sufficient dose sparing of the rectum and bladder. Thus, a plan with greater PTV coverage would be assigned a higher plan quality score than a plan with lower PTV coverage. Presently, there is no established gold standard for plan quality evaluation in routine clinical practice. Typically, the evaluation of the DVH is performed by a visual comparison of various dose-volume cut-points relative to the patient-specific dose prescriptions or clinical protocols. A standard quantitative metric for comparing and evaluating different treatment plans is a pre-requisite for developing automated treatment plans. With increasing adoption of automated segmentation tools and the use of contour atlases, it is not unreasonable to expect that we could also eventually have automated systems to design treatment plans. When such systems become available, it will be necessary to have a standard quantitative metric for comparing and evaluating different automated treatment plans. This simple approach may lead to greater standardization of how treatment plans are reported and compared, thereby allowing for more meaningful outcomes study necessary to establish efficacy of different treatment

approaches (i.e. rotational IMRT vs static IMRT) and various fractionation schemes (i.e. hypofractionation).

Secondly, we found that the average MI score for the seven beam angles of a given case can be a possible surrogate for overall plan quality. We then assessed the secondary effect on plan quality of importing optimization weights from prior cases. We found that a small improvement in PTV coverage is obtained on the order of a few percentage points, and substantial gains can be realized for the dose sparing for both bladder and rectum, with average additional dose sparing of 12.4% and 9.1% respectively. For plans that have initially high plan quality to begin with, we found that the optimization process may actually reduce the overall plan quality by a few percentage points. Therefore, one may wish to establish a PQ threshold for which no further optimization would be required.

We have now demonstrated that new IMRT treatment plans can be developed semi-automatically by importing the set of fluences and optimization weights from a prior case selected from a knowledge-base reference library. This approach allows the treatment planner to forego the current manual, laborious process of having to start anew from a blank slate for each new patient. Additionally, we have shown that a simple plan quality metric can be implemented to evaluate and quantitatively compare different treatment plans, which has the potential to standardize the manner in which different treatment approaches/regimens are reported and evaluated. Lastly, we have found that the average MI score can be used as a surrogate of overall plan quality – whereby a matched case of high geometric similarity should produce a new plan of high plan quality as well. In the next study, we will conduct an inter-institutional comparison of treatment plan quality, as well as a validation of whether the knowledge-

based approach can be extended to develop clinically acceptable plans using non-Duke cases.

## **5. Inter-institutional comparison of plan quality**

### ***5.1 Introduction***

In the first study reported in Chapter 3, we demonstrated the feasibility of the knowledge-based treatment planning approach by using the best-matched case to develop a new treatment plan. The three steps of the approach include i) using the case-similarity algorithm to identify a single best-matched case based on the overall mutual information score, ii) performing image registration to adapt the fluences from the best-matched case into the query case, and iii) importing the optimization weights from the prior matched case, prior to the optimization and final dose calculation.

In the second study presented in Chapter 4, we investigated the relationship between the resulting plan quality for a knowledge-based treatment plan and the initial level of case-similarity of the matched case. For a given query case, we utilized several matched cases that span the full range in of mutual information score to develop several pre-optimized plans. We introduced a plan quality metric based on the area under the curve of the DVH plot for the PTV, rectum and bladder. This plan quality index (PQ) was then used to calculate a single score for each newly generated knowledge-based plan. Finally, we compared the plan quality score of the pre-optimized plans with the average mutual information score and determined that the resulting plan quality for a given case is positively correlated to the initial case-similarity of the selected matched case.

Lastly, in this final validation study, we explore whether it is feasible to develop a clinically acceptable plan if the initial query case was selected from an outside institution. Up to this point, we had selected retrospective query cases drawn from the Duke knowledge base and matched it against cases from our own institution. The results of this leave-one-out approach is potentially biased since we have been matching

query cases against a Duke-only database, which is comprised exclusively of cases previously developed by Duke professionals using our institutional protocols. In this final study, we assembled query cases from two outside institutions in order to investigate whether our knowledge-based approach is appropriate and extensible to designing treatment plans for an effectively 'new' patient. This validation study will examine the feasibility of matching non-Duke cases against a database comprised exclusively of Duke plans. Subsequently, the study will also investigate whether it is feasible to adapt the treatment parameters from our institution's knowledge base to develop clinically acceptable plans for new query cases drawn from outside institutions. Lastly, the study provides an opportunity to conduct an inter-institutional comparison of plan quality of the original clinical cases, as well as a comparison of semi-automated knowledge-based planning and manual planning.

## **5.2 Material and methods**

The assembled database consists of 250 retrospective prostate IMRT treatment plans collected over a three-year period at our institution. Additionally, we collaborated with two outside radiation oncology practices, which we will refer to as Institution #1 and Institution #2, to acquire 100 cases from each institution for a total of 200 cases. We obtained IRB permission for the use of these existing retrospective cases in this HIPAA-compliant study, and the cases were anonymized prior to use. At both institutions, the clinical treatment plans were developed using the Eclipse treatment planning system (Varian Medical Systems, Palo Alto, CA). Each treatment plan includes the CT dataset, tumor/target volume and normal structure contours, beam geometry specifications, fluence maps, optimization weights, as well as the final dose-distributions. At Institution #1, both five and seven co-planar beam angles are used, with the majority being seven-beam angle plans. At Institution #2, all treatment plans are developed using exclusively

five beam angles. All cases from Duke and the two outside institutions considered in our investigation were treated in the supine position.

As previously described in Chapter 3, the case-similarity algorithm uses mutual information to identify similar patient cases by matching 2D beam's eye view projection of contours. It is important to note that although the specific treatment beam angles may vary from case to case, as well as the actual number of beam angles used, (i.e. five beams versus seven beams). To account for these variability from case to case, as well as the differences in institutional practice, we chose to standardize the matching of cases at the standard set of seven beam angles used at Duke: 25°, 75°, 130°, 180°, 230°, 280°, and 330°. We believe this practice is justified since we will subsequently have to adapt the fluence maps from the matched case to the specific beam angles prescribed for the query case.

Unlike the previous study reported in Chapters 3 and 4, whereby a leave-one-out approach is implemented by using query cases selected from the Duke knowledge base, we now use the non-Duke query cases as 'new' cases for which we can test and validate the robustness of our knowledge-based algorithm. That is, by selecting query cases from an outside institution, we thus remove the intrinsic bias that potentially exists when matching query cases against a database of cases from the same institution, which have been similarly designed and implemented. Next, we registered the prior optimized fluences to the PTV and adapted the optimization weights from the best-matched cases for the query case. Since these query cases have also been previously manually optimized and clinically approved at another institution, we can also conduct a comparison of the original clinical plan quality across institutions.

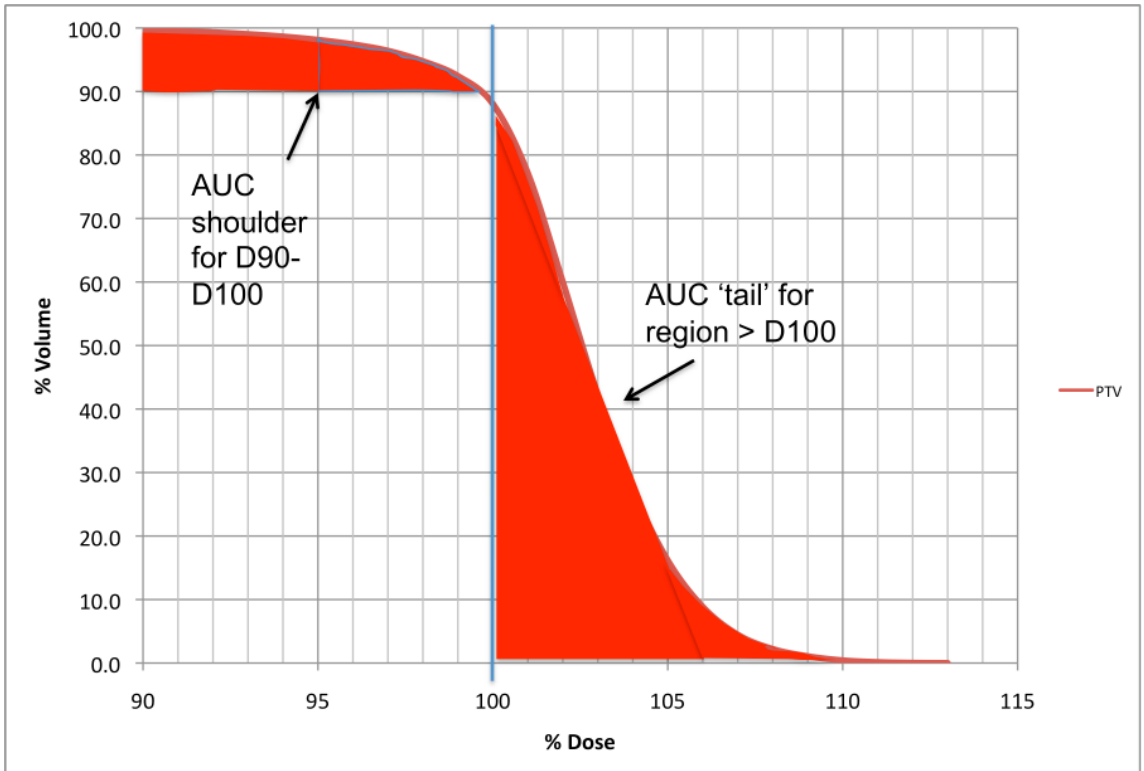
In the first part of this validation study, we compared the original, manually-derived treatment plans from the three institutions, by considering the original fluence maps as well as the DVH plots. First, we visually compared the fluence maps, looking

at the heterogeneity of the intensity distributions, as well as the size and shape of the fluence distributions.

For the second part of this validation study, we selected ten cases from PMC for re-planning. In order to simplify the comparison between five- and seven beam-angle plans, we chose to re-plan only cases that were originally manually planned as a seven-beam plan. Using the knowledge-based approach explained in Chapter 3, a single best-matched case and the corresponding fluences and optimization weights were adapted to design ten new treatment plans.

Lastly, we utilized the plan quality index described in Chapter 4 and compared the AUC for each of the structures as well as the overall plan quality for each new plan. The final part of this research study involved the inter-institutional comparison of plan quality between ten new, knowledge-based plans and the original manually-generated plans from PMC. We also examine the DVH plots of the PTV in greater detail, focusing in on the important shoulder region above 90% of the structure volume, specifically defined for the  $D_{90}$  to  $D_{100}$  region. The clinical priority of establishing PTV coverage would have i) 100% of the volume receiving 100% of the dose, and ii) a steep drop-off in coverage of the PTV above  $D_{100}$ . Any region above the  $D_{100}$  implies additional unnecessary dose exceeding the prescription limit will produce as hot spots within the PTV. For the critical structures (i.e. rectum and bladder), the same AUC plan quality metric is calculated for cases from each of the three institutions, however, there is no additional consideration of shoulder or tail regions. *Figure 24* illustrates the definition of these two regions of the PTV used in this study.





10

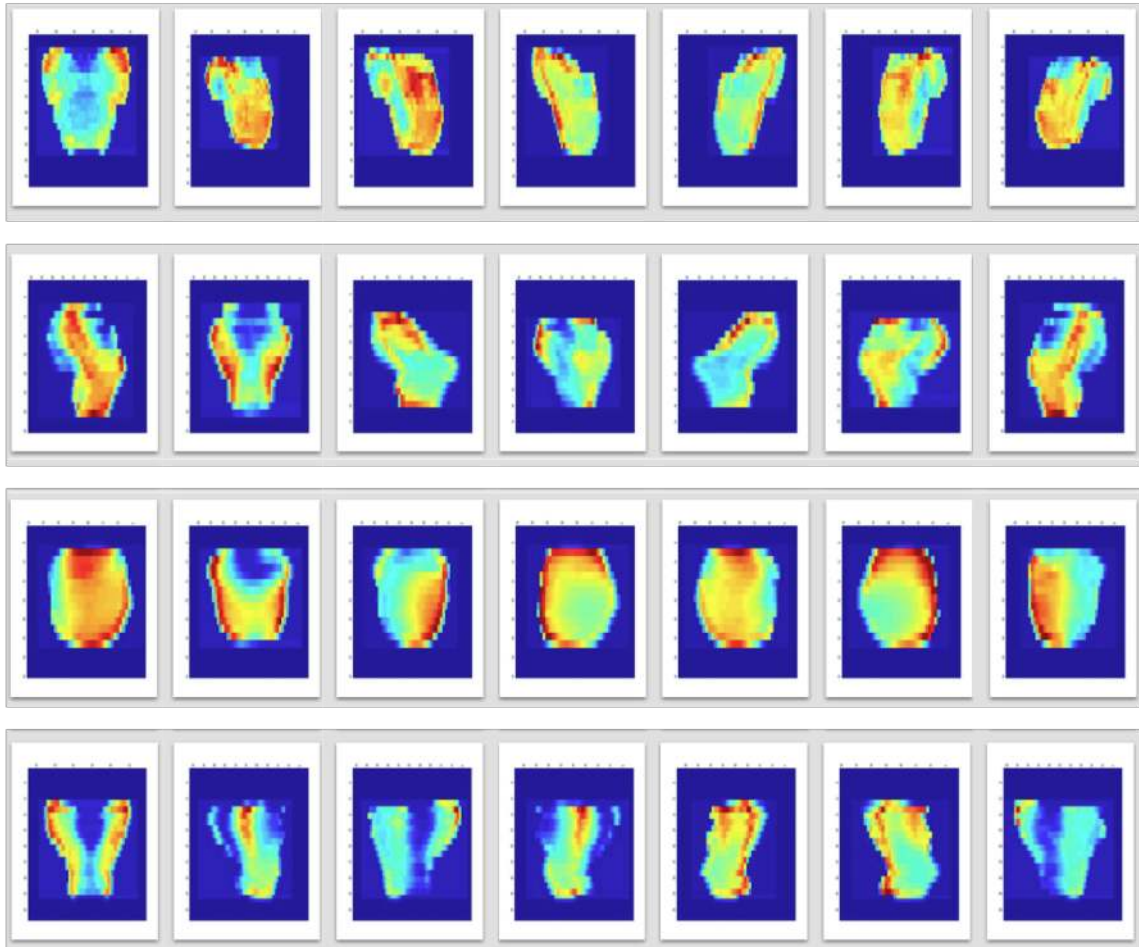
**Figure 24** Diagram of AUC shoulder and AUC tail regions of the PTV

The shoulder region is defined as the partial area under the curve, from 90%-100% of the dose, for percentage volume greater than 90%. An ideal PTV would have an AUC of 1.0 for the shoulder. The tail region corresponds to the region above 100% of the prescription dose, that is, the region where excess dose greater than the prescription can produce 'hot' spots.

**5.3 Results**

Typically, the fluence intensity maps are not visually reviewed except perhaps as an overlay on top of the beam's eye view of the structure contours for the purpose of checking alignment. It is not uncommon for the physicist or dosimetrist to do minor 'dose-painting' to manually adjust the fluence map pixel by pixel to remove undesired hot spots. In the first part of this study, we generated the set of seven fluence maps for each case, and viewed the ensemble of these fluence map sets for each institution. Interestingly, we found that the fluence maps at our institution were very different than

those from both outside institutions. *Figure 25* depicts the heterogeneous distribution of radiation intensities from Duke cases, which is what we would expect to see in a typical intensity modulated fluence map.

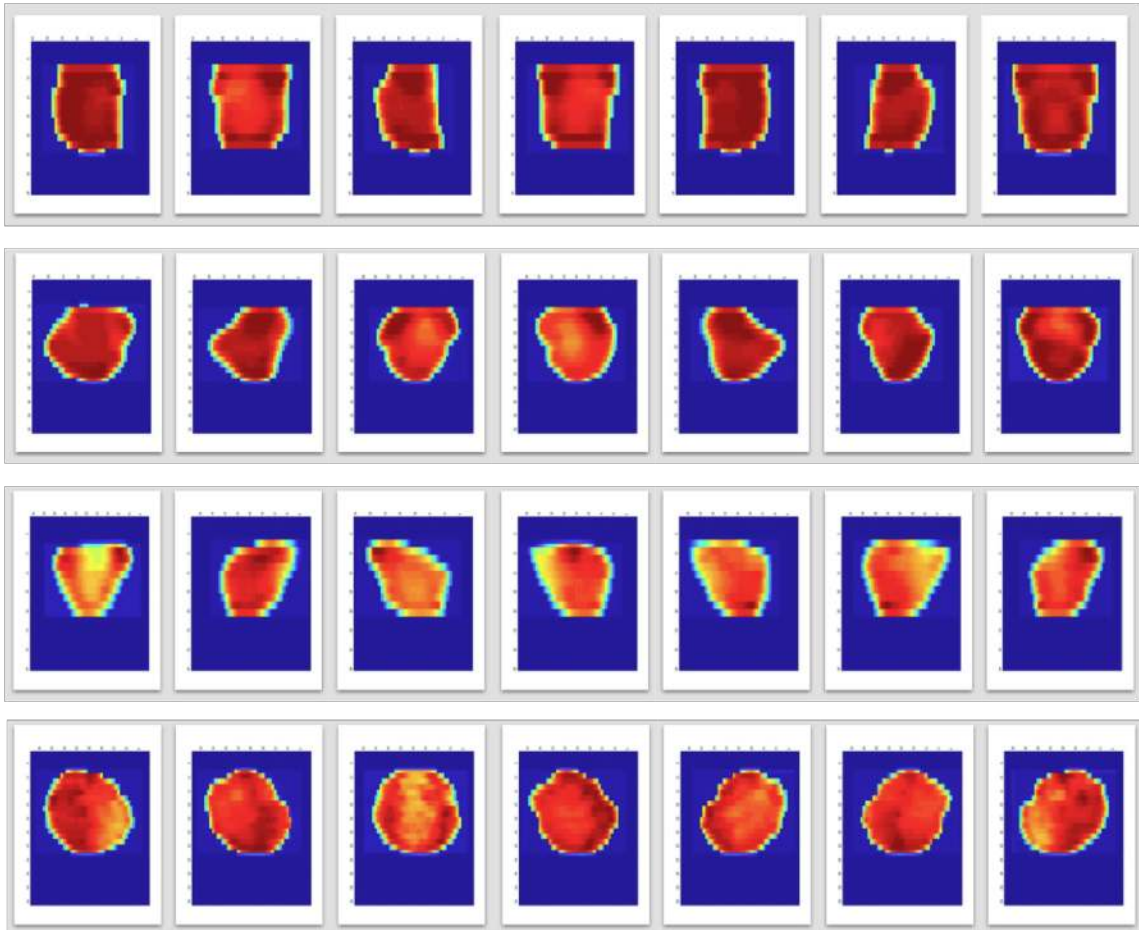


**Figure 25** Four examples of the set of seven fluence maps from Duke.

Each row corresponds to the set of seven fluence maps for a given case. Note that these fluence maps reveal the variation in the intensity distributions expected in an intensity-modulated treatment plan.

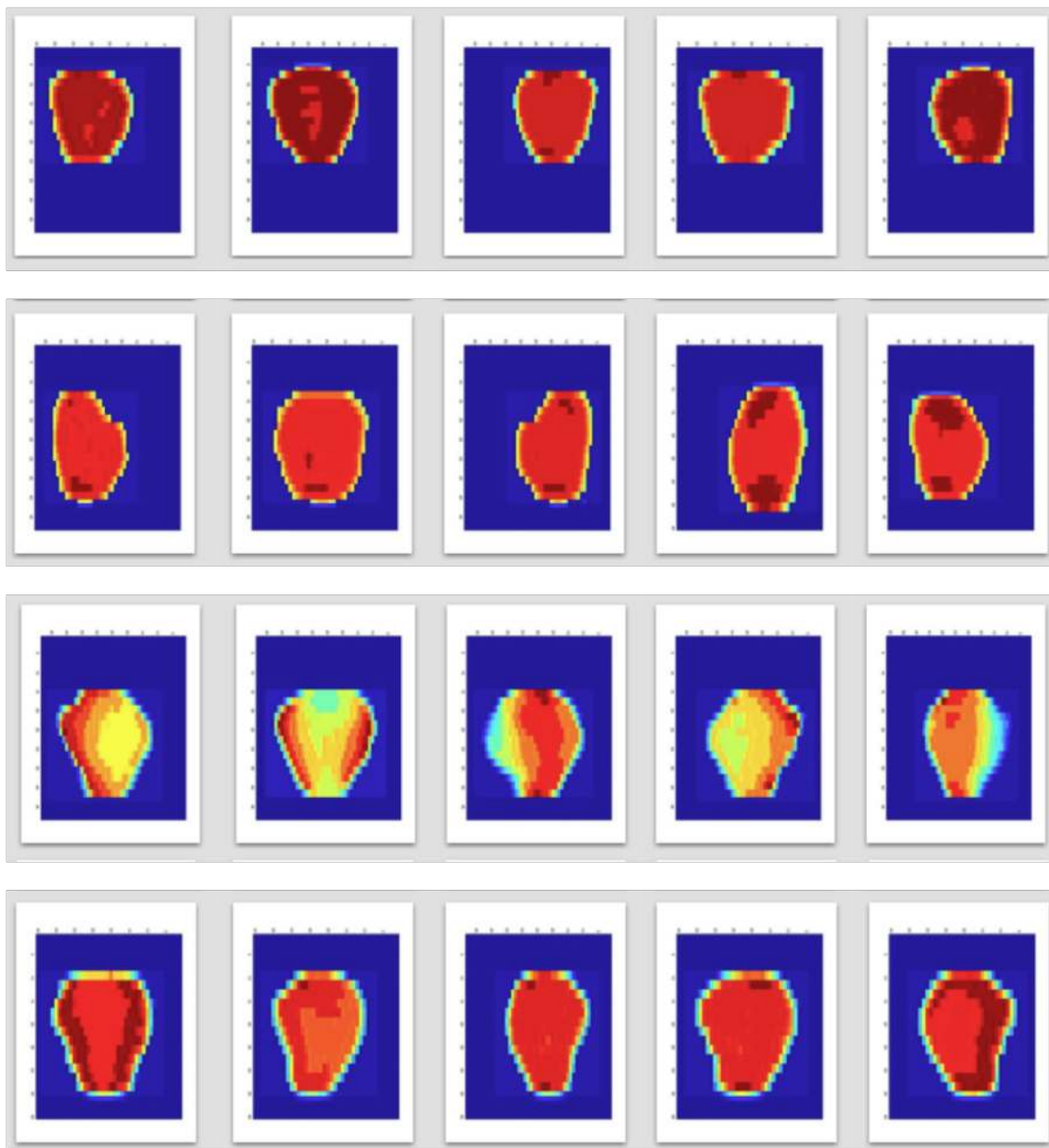
*Figure 26* and *Figure 27* illustrate the representative examples of fluence maps from Institution #1 and Institution #2 respectively. At both institutions, the fluence maps are more uniform in the intensity distributions, implying little modulation of the fluence

intensities. Additionally, the shapes of the fluence fields at both institutions tend to be more circular than those at Duke.



**Figure 26** Four examples of the set of seven fluence maps from Institution #1.

Each row corresponds to the set of seven fluence maps for a given case. Note that these fluence maps reveal very little variation in the intensity distributions, which is more reflective of a homogenous 3D conformal fluence pattern.

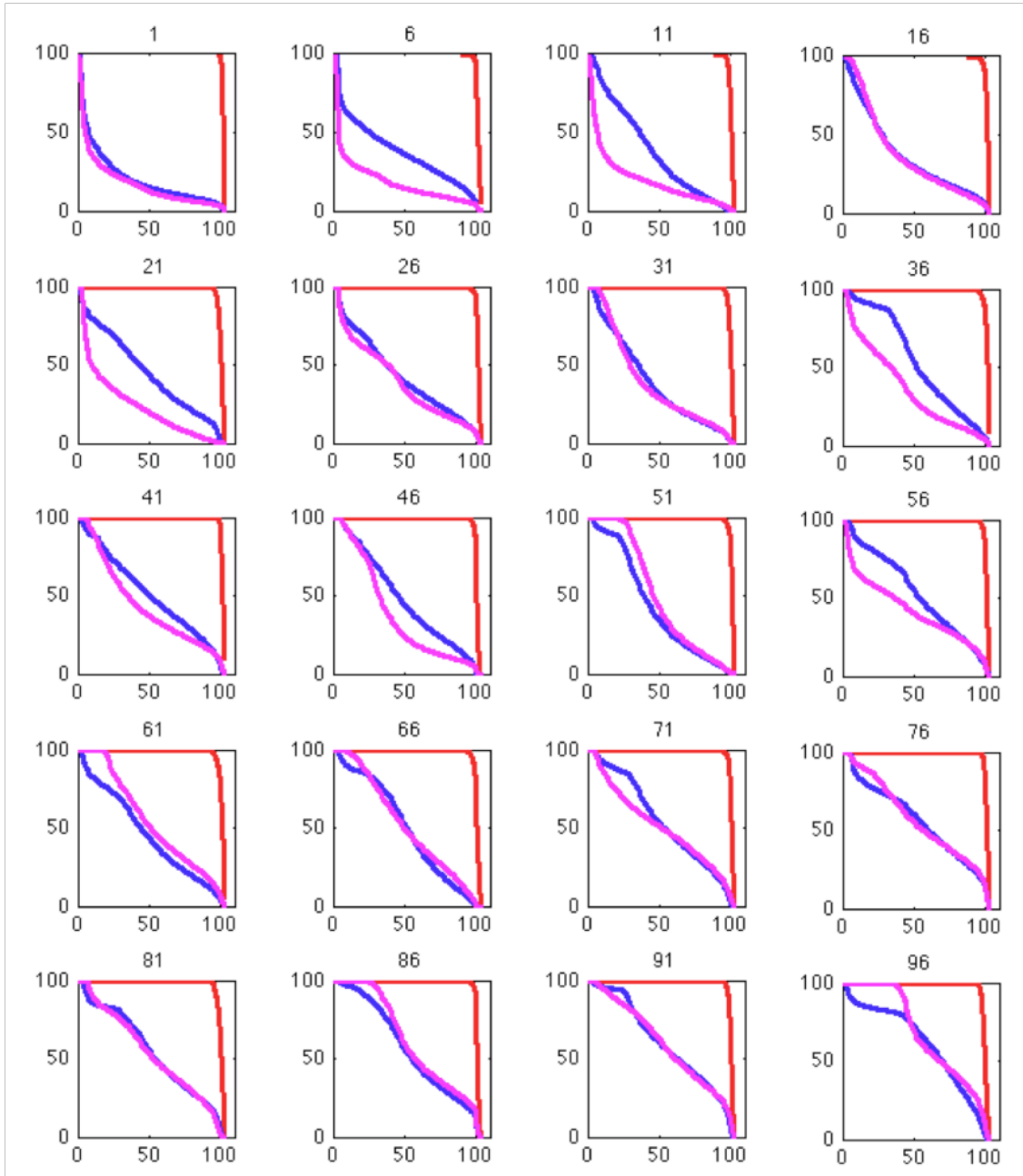


**Figure 27** Four examples of the set of seven fluence maps from Institution #2.

Similarly, these fluence maps from High Point also reveal very little variation in the intensity distributions, which is more typical of 3D conformal fluence maps.

Next, we consider the dose-volume histograms of the original clinical case for each institution. The plan quality score, based on the weighted sum of the area under the curves, was calculated for each plan and the plans at each institution were rank-

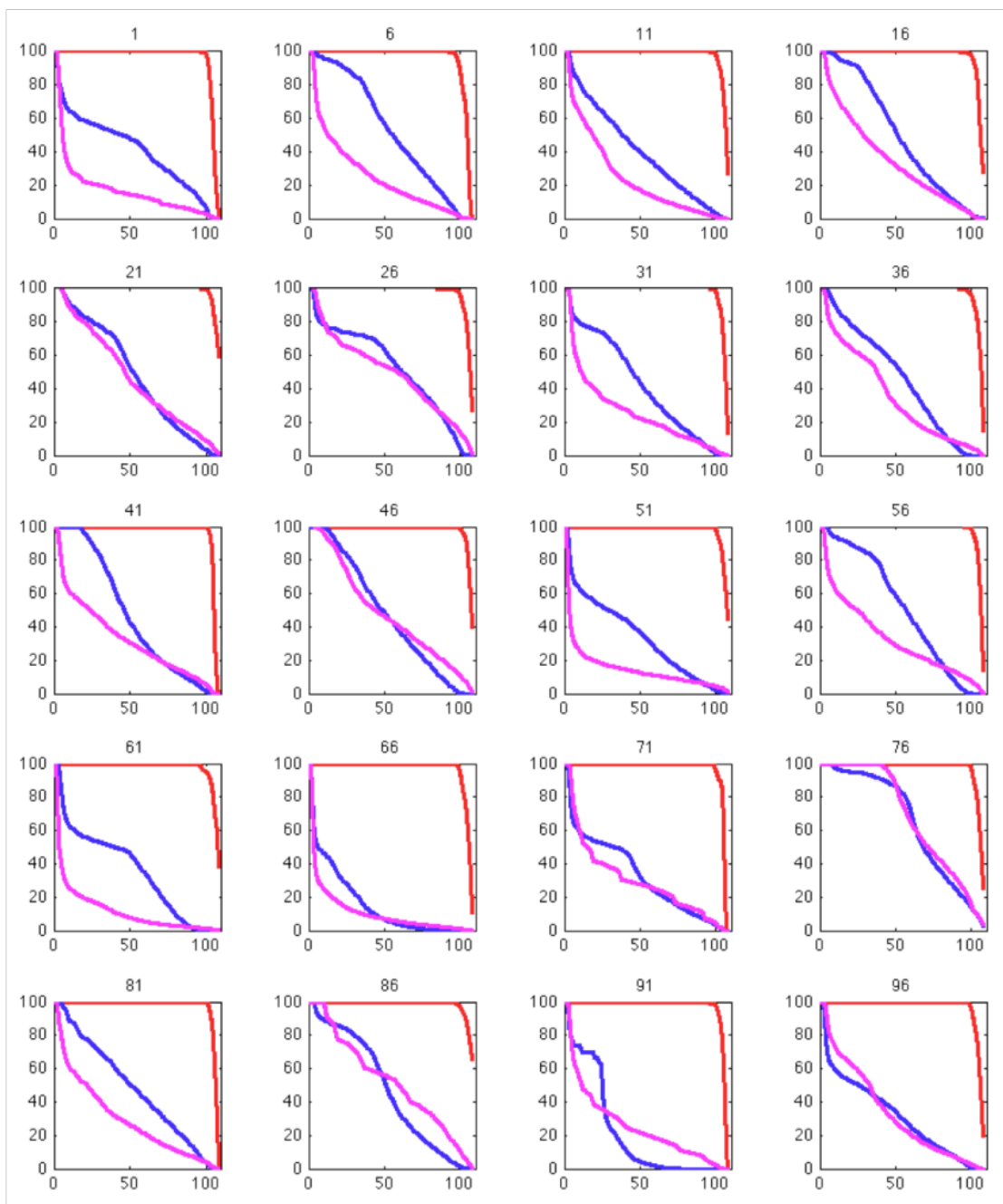
ordered. *Figure 28* depicts the dose-volume histograms for twenty Duke cases, sorted from high PQ to low PQ. The red line corresponds to the PTV, the blue line the bladder, and the cyan line represents the rectum.



**Figure 28** Twenty Duke cases sorted from high PQ (#1) to low PQ (#96)

Every fifth case from the Duke database is shown.

In examining the DVH for each case, we wish to see the full extent of PTV volume coverage, given by the shoulder region at the prescription dose of 100%. We also consider the excess dose to the PTV indicated by the area under the curve of the 'tail' region for dose greater than 100%. In general, we find that Duke cases have very good PTV coverage, as evidenced by the sharp drop-off in the shoulder region. We also see the full range of rectum and bladder dose-to-volume coverage, with the rectum dose tending to be equal to, or less than, the bladder dose. It is useful to remind the reader that the protocol at Duke (refer back to Table I) requires the dose to 50% of the rectum and bladder volume to be under 39 Gy. For a protocol of 78 Gy, this limit corresponds to 50% of the prescription dose.

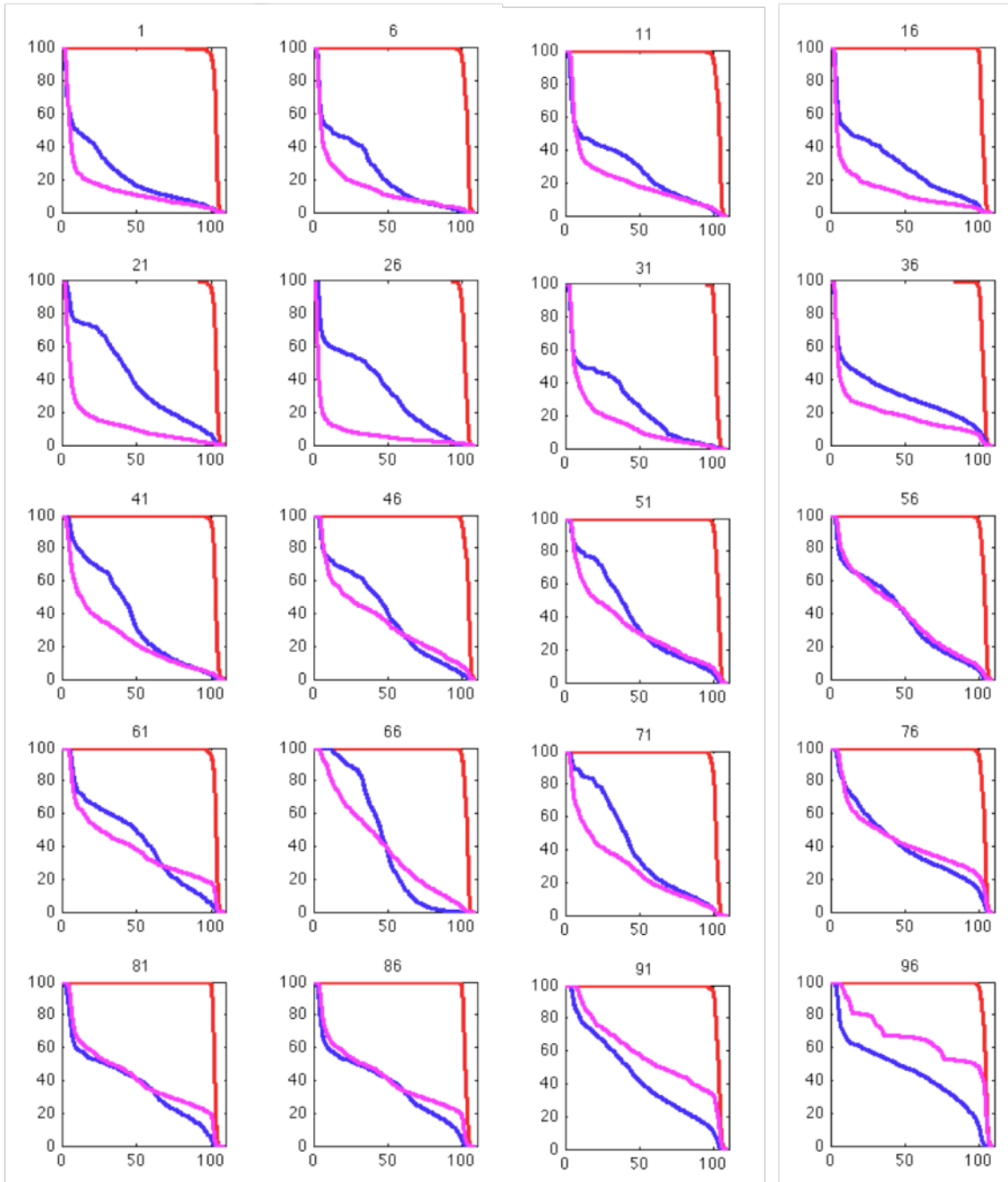


**Figure 29** Twenty cases from Institution #1, sorted from high PQ to low PQ.

Every fifth case from the database of Institution #1 is shown.

The DVH for the cases from Institution #1 are shown in *Figure 29*. It can be visually seen in many cases that the slopes of the PTV shoulder are not as sharp as for those at Duke. This implies that while there may be adequate PTV coverage at the 100% prescription dose level, there will also be additional volume receiving doses in excess of the prescription dose. We also find at Institution #1 that the rectal dose tends to be lower than the bladder dose. *Figure 30* depicts the DVH for cases from Institution #2. We find sharp drop-off in PTV shoulder region for most plans, and the tendency of relatively lower rectal dose than bladder dose for high PQ cases.





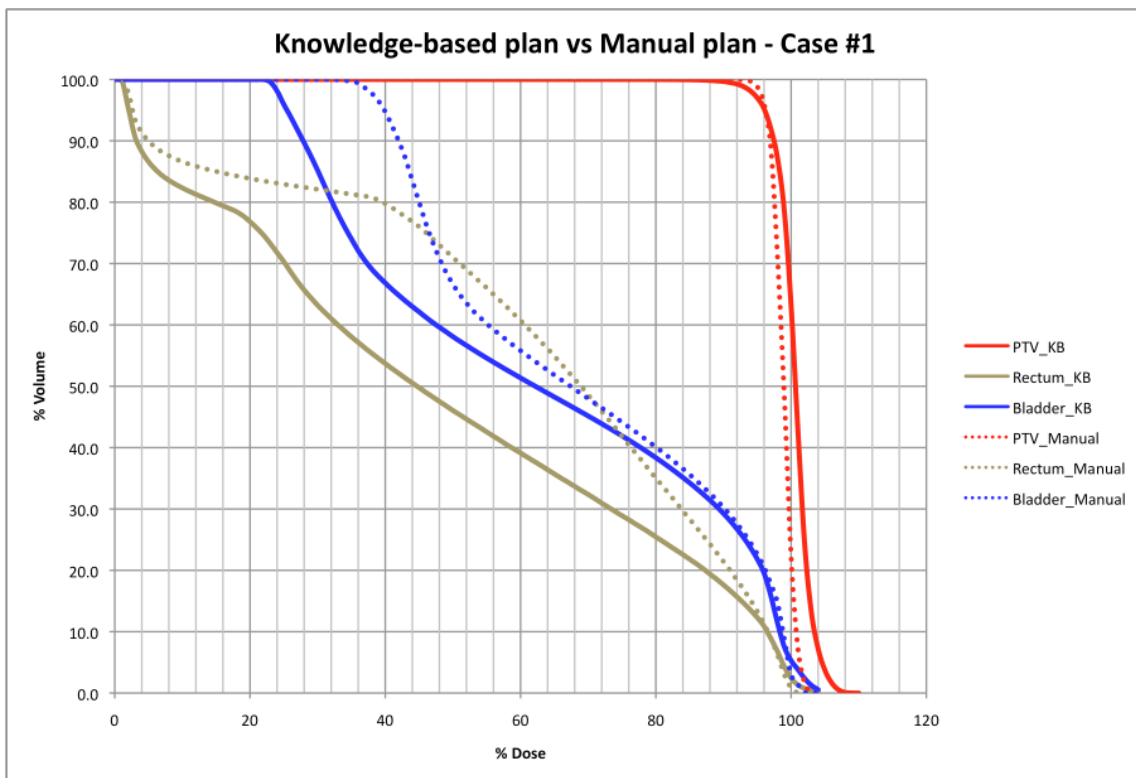
**Figure 30** Twenty cases from Institution #2, sorted from high PQ to low PQ

Every fifth case from the database of Institution #2 is shown.

Mid-range PQ cases tend to have equal bladder and rectal dose, while lower PQ cases tend to have higher rectal dose.

## Re-planning for new cases using a knowledge-based approach

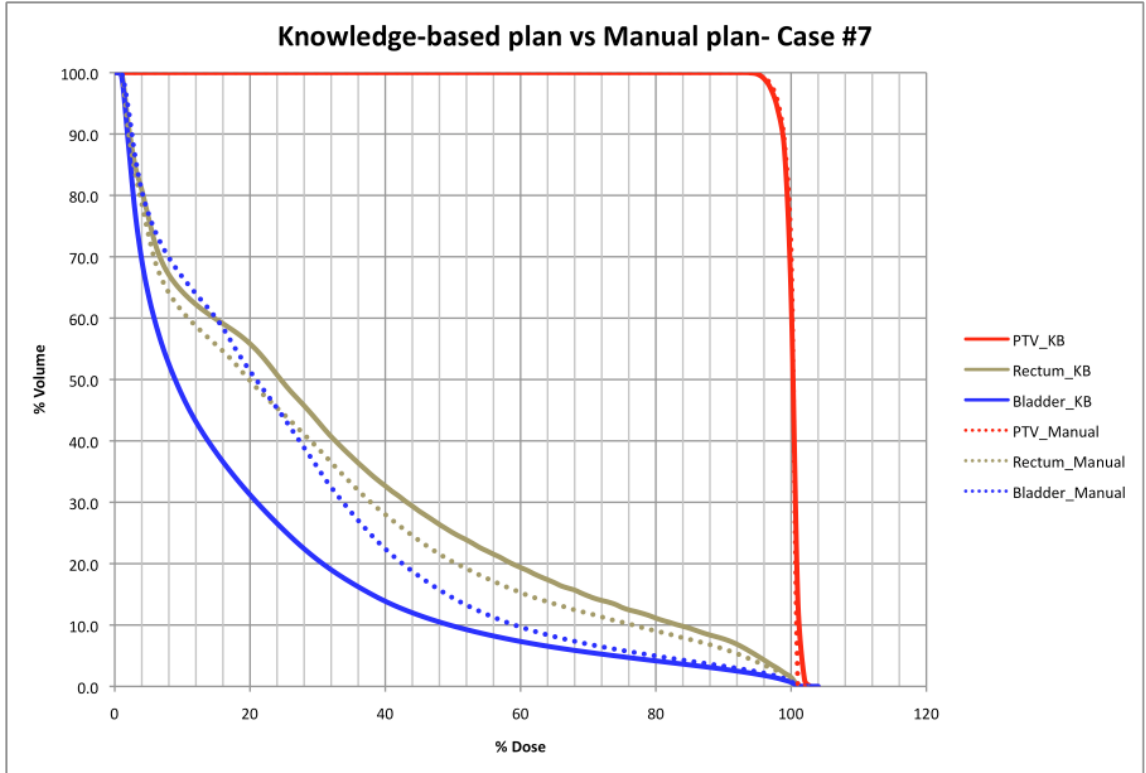
For the last part of this validation study, we adapted the prior fluence maps and optimization weights from the matched case to design new plans for the query case. We selected only seven-beam angle cases from Institution #1 for re-planning. Here, we present a summary of the DVH results of re-planning for ten query cases. We selected three cases that represent the full range of our re-planning results, i) substantial plan quality improvement, ii) comparable plan quality, and iii) slightly reduced plan quality. *Figure 31* is a comparison of the DVH for Case #1 - the new, knowledge-based plan (solid lines) and the original, manually generated clinical plan (dotted lines) from Institution #1.



**Figure 31** Comparison of knowledge-based plan (Case #1) vs. manual plan

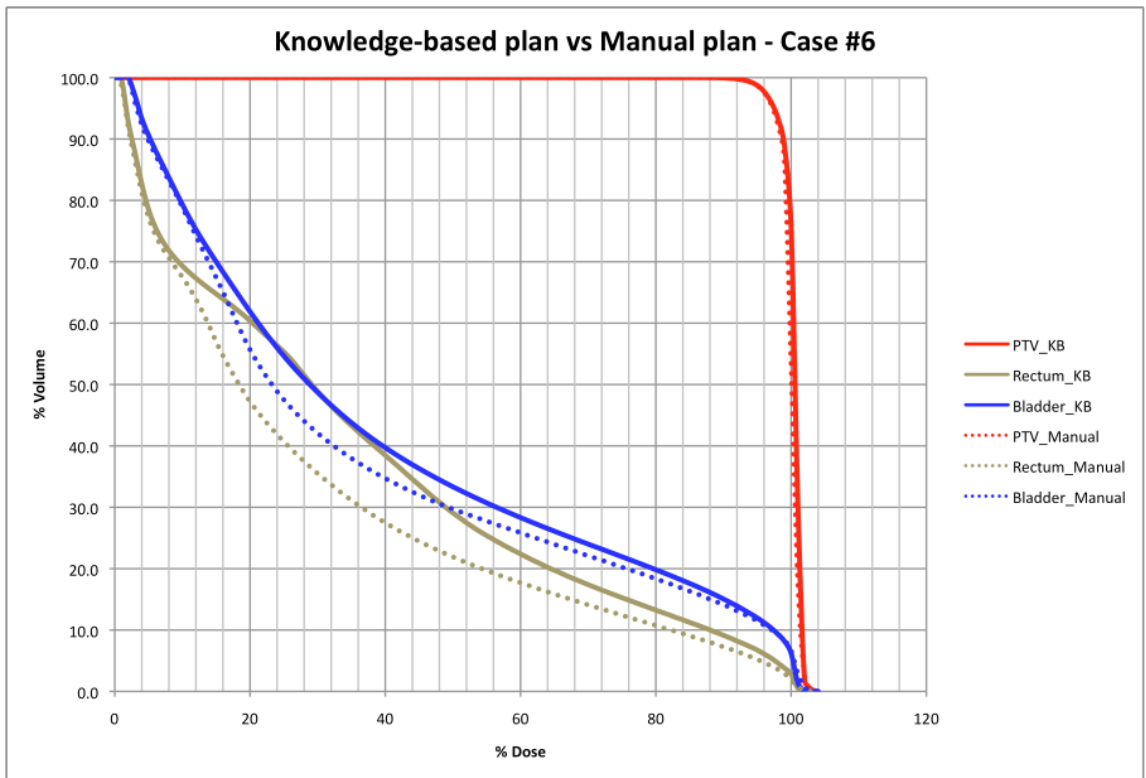
For case #1, we see that for this knowledge-based plan, the PTV coverage is slightly higher than the original manual plan, and the rectum and bladder dose are lower than

the original clinical plan – even though the rectum dose is still slightly higher than required by Duke protocol.



**Figure 32** Comparison of knowledge-based plan (Case #7) vs. manual plan

**Figure 32** illustrates the re-planning for case #7, where the PTV coverage in the knowledge-based plan is equal to that of the original clinical plan. For this knowledge-based plan, there is a slight reduction (improvement) in the bladder dose but a slight increase (degradation) in the rectum dose. However, for this case, both the knowledge-based plan and the manual plan are able to achieve relatively great dose sparing to bladder and rectum.



**Figure 33** Comparison of knowledge-based plan (Case #6) vs. manual plan

**Figure 33** illustrates the comparable dose volume histograms between the knowledge-based plan and the original manual plan. Of the ten new re-planned cases, Case #6 was the only one case where both the rectum and bladder dose were slightly higher in the knowledge-based plan. All the other nine re-planned cases were found to have comparable or slight improvement in both rectum and bladder dose. These specific values for each of the ten cases are shown in **Table XI**.

In summary, for the PTV, the mean percentage difference was found to be 0.4% indicating that the knowledge-based plan is comparable in PTV coverage as the manual plan. The mean improvement in dose sparing for rectum and bladder were 3.6% and 9.3% respectively. For nine out of ten cases, both rectum and bladder doses were comparable or slightly less for the knowledge-based plans. In fact, if we exclude the

outlier case, the average AUC decrease for rectum and bladder were 7.5% and 13.9%, respectively, which is actually quite substantial. Using the weighted plan quality metric, the mean percentage difference for the overall plan quality of -0.3% suggests that the knowledge-based plans are very comparable to the manual plan.

**Table VI** Percentage difference in AUC between knowledge-based plan vs. manual plan

	<b>% diff AUC = 100*(KB plan – Manual plan)/Manual plan</b>			
<b>Case</b>	<b>Δ AUC PTV</b>	<b>Δ AUC rectum</b>	<b>Δ AUC bladder</b>	<b>Δ PQ plan</b>
1	2.7	-0.4	-2.0	2.2
2	4.3	-21.8	-19.0	2.0
3	0.3	-9.4	-1.6	-0.3
4	-5.0	-19.7	-25.9	-6.5
5	-0.1	-22.9	-23.9	-2.1
6	0.2	31.8	32.0	1.9
7	0.0	11.7	-28.3	-0.4
8	-0.1	-2.4	-2.6	-0.3
9	1.3	0.0	-10.3	0.7
10	0.3	-2.6	-11.4	-0.4
<b>Mean</b>	<b>0.4</b>	<b>-3.6</b>	<b>-9.3</b>	<b>-0.3</b>

Note that  $PQ = 10 * AUC_{PTV} + (1 - AUC_{rectum}) + (1 - AUC_{bladder})$ . Also, for the critical structures, a negative value indicates an improvement, greater dose sparing.

## **5.4 Discussion**

In this final study, we were interested in determining the feasibility of developing knowledge-based plans using cases from an outside institution as ‘new’ query cases matched against a Duke-only database. We acquired 100 additional cases

each from two clinics, for a total of 200 cases, for the purpose of plan quality comparison and for re-planning.

A cursory review of the fluence maps show that the Duke fluence maps are far more heterogeneous in intensity distributions – as compared to the mostly uniform fluence maps from Institution #1 and Institution #2. These uniform fluence maps are more typically seen in 3D conformal plans rather than in intensity-modulated plans. However, it is difficult to say what the net effect of these uniform fluence maps by a visual inspection alone.

The dose volume histograms from all three institutions reveal adequate PTV coverage at the prescription dose level. Duke cases reveal a very sharp drop off in PTV coverage beyond the 100% prescription level, with very little volume receiving excess PTV dose. In contrast, while the PTV coverage is similarly adequate in both institutions, the excess dose beyond the prescription level is substantially higher in both Institution #1 and Institution #2. The dose sparing to rectum and bladder at all three institutions is adequate, and most cases have critical structure dose below the 50% volume < 50% Rx protocol used at Duke. The mean rectal dose for Duke cases was found to be slightly higher than at the other two institutions. However, the mean bladder dose for Duke cases was found to be substantially higher than those at Institution #1 and Institution #2. We believe this is due to the fact that these original plans are defined with different contouring protocols, i.e. expansion of PTV margin. As such, a meaningful inter-institutional comparison of treatment plan quality can only be made for re-planning of cases using the same structural contours and expansions.

An inter-institutional comparison of manual planning versus knowledge-based planning was conducted. Results of the re-planning study found that comparable PTV coverage was obtained for all ten cases, and that for nine out of ten cases, both rectum

and bladder doses were comparable or slightly less for the knowledge-based plans. These results are highly suggestive of the potential feasibility of designing clinically acceptable treatment plans using the knowledge-based approach based on a site-specific database of prior plans. The current process of importing fluence maps and optimization weights from the matched case is time consuming since the dose calculation and optimization is performed in the 'closed' commercial treatment planning system. Future work in developing an off-line dose-calculation and optimization engine should enable a more thorough investigation of a larger number of case comparisons. The potential for realizing significant clinical impact through this semi-automated treatment planning approach will come from developing an interface to automate the process of case-retrieval, registration of fluences and the importing of optimization weights.

We recognize that inconsistencies in structure contouring, particularly across different institutions, may present some challenges in the comparison of plan quality across institutions. For instance, both outside institutions created their PTV by expanding the prostate with a lower expansion margin than was done at Duke. In other words, the overlap between PTV and bladder/rectum is higher for Duke, so it would not be possible for AUC for bladder/rectum to be lower in the Duke plans. A second example is worth noting, some institutional practices include nodal involvement when defining the PTV, resulting in a larger PTV contour, which in turn may not appear in the knowledge base constructed at an institution that does not share that practice. Similarly, the contouring protocols are related to the dose prescription of that institution (i.e. posterior PTV margin adjacent to rectum is related to dose escalation protocols). However, we believe that the general framework for knowledge-based treatment planning can still be implemented across institutions using site-specific guidelines

concerning structure definition and margins. Further adoption of Radiation Therapy Oncology Group (RTOG) contouring and prescription protocols as well as increasingly robust atlas-based solutions may also help reduce the practice variability across institutions, and thus, it would be quite feasible to implement a national archive of knowledge cases for use across institutions.

In conclusion, we find that a site-specific database of prior plans can be effectively used to design new treatment plans for outside cases. It remains to be seen whether a heterogeneous database of plans drawn from multiple institutions is suitable, or perhaps superior to a site-specific database. There remains the long-standing issue of contouring variability. As more efficient algorithms for atlas-based contouring become more widely used, the role for knowledge-based treatment planning should increase, and the potential for further automation of the treatment planning process should be realizable. This would allow for greater standardization of the quality of treatment planning across institutions. Given the current time and effort expended on developing IMRT plans, a fully automated and integrated knowledge-based planning solution will likely improve treatment planning efficiency across institutions.



## **6. Summary, Conclusion and Future Work**

### ***6.1 Summary of Findings***

The clinical challenge in IMRT treatment planning requires providing an effective coverage of the planning target volume while sparing the normal structures (i.e. rectum and bladder) to the greatest extent possible. We have addressed this problem by proposing a knowledge-based solution, whereby the manually optimized prior cases that reflect the expertise and time spent by a human planner is leveraged as a semi-automated algorithm. First, a knowledge base consisting of CT data, structure contours as well as various treatment parameters is assembled from prior optimized and clinically approved IMRT cases from the Duke clinic. Second, a case-matching algorithm is developed using mutual information as the basis for identifying similar patient cases. A query case is selected from the knowledge base, and considered as if it were a new patient case for which a treatment plan was needed. The case-similarity algorithm identifies a range of matched cases similar to the query case. Thirdly, the best-matched case having the highest overall MI score is selected and its treatment parameters are adapted to design a new treatment plan for the query case. Lastly, we evaluate the resulting new semi-automated plans and compare these to the original plans developed by a human planner. We evaluate the plans based on dose-to-volume criteria that reflect the initial prescribed protocol.

In the first study described in Chapter 3, we demonstrated the feasibility of a knowledge-based approach of using prior clinically approved treatment plans. We found that the approach can substantially reduce planning time by skipping past all but the last few optimization iterations of the planning process. Specifically, we showed it is possible to generate high quality plans using a semi-automated approach, and this approach can substantially improve the efficiency of treatment planning.

In the second study described in Chapter 4, we conducted an evaluation of treatment plan quality and the relationship between plan quality and case-similarity based on mutual information. We introduced a figure of merit called the plan quality score, based on the weighted sum of the area under the DVH curve for the PTV, rectum and bladder. We found that a weighting of ten, one, one for the PTV, rectum and bladder respectively produces a monotonic rank-ordering of cases consistent with clinical plan quality expectations. The matched cases that are substituted into the query cases result in plans that have a full range in plan quality. We found that the plan quality of a pre-optimized plan will nearly always improve more after the optimization process. Even though plans with lower pre-optimized plan quality tend to improve substantially after optimization, they generally end up with lower post-optimized plan quality than plans with higher pre-optimized plan quality. On the other hand, plans with higher pre-optimized plan quality also generally had higher post-optimized plan quality. However, the plans that are initially high in plan quality do not result in much more substantial plan quality gains, and for a few instances, the plan quality actually may decrease slightly. Lastly, we found that the mutual information score can act as a surrogate for overall plan quality, which will greatly facilitate future research efforts.

In the final study described in Chapter 5, we compared the treatment plans collected from two outside institutions, in terms of their respective DVH and the fluence maps. We found that the Duke fluence maps are far more heterogeneous in intensity distributions, in contrast to the very uniform fluence maps from both of the outside institutions. However, it is difficult to determine what the net effect of these differences in uniformity alone may be. Next, the comparison of the DVH found that the PTV coverage to the prescription was adequate at all three institutions, although there was a tendency for there to be substantially higher excess dose to the PTV at the other two

institutions. For the critical structures, the dose sparing to both the rectum and bladder at each of the three institutions was found to be adequate, where most cases had normal tissue dose falling below the 50% volume < 50% Rx protocol. The mean rectal dose for Duke cases was found to be slightly higher than at the other two institutions, while the mean bladder dose for Duke cases was found to be substantially higher than those at the other two institutions. We believe this is due to the differences in contouring protocol, since the PTV margin extension at the other two institutions are lower than those at Duke. That is, the overlap of the PTV with both rectum and bladder is expected to be greater due to the larger margin extensions. A meaningful comparison of plan quality cannot be extrapolated from this if the initial contours are defined differently for each institution.

Lastly, we conducted a limited comparison study of manual planning versus knowledge-based planning using ten query cases selected from one of the outside institution. These query cases were matched against the Duke database, and new plans were generated using the knowledge-based approach. Results of the re-planning study found that the knowledge-based plans provided comparable PTV coverage to the original manual plan. With the exception of one case, the average AUC decrease for rectum and bladder was 7.5% and 13.9%, respectively, which is actually quite substantial. In conclusion, we found that a site-specific database of prior plans can be effectively used to design new treatment plans for outside cases.

## **6.2 Conclusion**

The current process of IMRT treatment planning is an iterative time-consuming process that is highly dependent on the skill and experience of the planner. It has previously been reported that there is wide variability across medical institutions between the prescribed radiation dose and the dose delivered. This is suggestive that

there may also be wide variability in treatment plan quality across institutions. We demonstrate the potential of using a knowledge base approach of identifying and adapting prior optimized treatment plans to generate clinically acceptable prostate IMRT treatment plans for new cases. This approach involves identifying a similar matched patient case from a database of prior clinically approved plans, and adapting the treatment parameters to the new case. We have demonstrated that our knowledge-based approach is a feasible semi-automated solution for IMRT treatment planning, offering the potential to improve the efficiency of the treatment planning process while ensuring that high quality plans are developed.

We have shown that a simple plan quality metric can be implemented to evaluate and quantitatively compare different treatment plans, which has the potential to standardize the manner in which different treatment approaches/regiments are reported and evaluated. Lastly, we have found that the average mutual information score can be used as a surrogate of overall plan quality – whereby a matched case of high geometric similarity should produce a new plan of high plan quality as well.

Lastly, we found that a site-specific database of prior plans can be effectively used to design new treatment plans for outside cases. It remains to be seen whether a heterogeneous database of plans drawn from multiple institutions is suitable, or perhaps superior to a site-specific database.

### **6.3 Future Work**

In this initial study, our algorithm identified the single best-matched case based on the simple sum of the seven individual MI values for each beam angle. In the future, we plan to explore alternative methods of improving the case-similarity matching, such as weighting schemes based on beam angle. It may also be possible to improve case

similarity retrieval through matching the full 3D volumetric dataset, rather than simply using the beam's eye view projection images.

Since it is not yet feasible to know *a priori* the full extent of dose sparing that can be realized for any individual patient, it would be beneficial to use a database of prior clinical plans with parameters that could reasonably estimate the possible extent of dose sparing. Further investigation should be conducted to determine the relationship between the match score and the extent of dose-sparing levels. That is, it would be tremendously useful to be able to obtain an estimate of this dose-sparing extent for a new case, extrapolated based on the similarity of the matched case, without having to actually perform the dose calculation. If the knowledge-based approach could provide a quick and accurate estimate of the achievable dose sparing, by avoiding the computationally intensive dose-calculation, this would be a valuable alternative to an exhaustive search of the entire solution space of useful fluences.

Presently, the dose calculation process for generating a knowledge-based treatment plan is still quite time-consuming. It was necessary to perform the dose-calculation and the final optimization within the commercial Eclipse treatment planning system (TPS). Since the commercial TPS is designed for clinical use, rather than as a research tool, it is presently not feasible to easily swap fluences and optimization weights from one patient case to another. This process was done off-line for this research project, and the set of fluences and optimization weights imported from the matched case had to be entered manually. Potentially, using the CERR tool to perform these automated dose-calculations off-line could solve this problem.

Lastly, there is still plenty of interesting work to be done in terms of evaluating the knowledge-base design. By continuing to add additional cases, one can investigate the robustness of the knowledge base, in terms of the number of cases, as well as the

quality and the kinds of plans that are used to assemble the knowledge base. We have shown that a site-specific database can be used to develop a treatment plan for query cases drawn from an outside institution. However, due to limited time, we were not able to investigate whether a knowledge base assembled from cases selected from multiple institutions would be more useful, and what techniques should be developed to query, and adapt fluence and optimization parameters given that they would differ from one institution to another. Further work could also extend the case-similarity algorithm to consider combining treatment parameters drawn from multiple patients, rather than the present method of using parameters taken from only a single, individual patient.

Finally, we had initially selected prostate cancer to investigate the feasibility of developing a knowledge-based treatment planning approach for two reasons: i) the high volume of prostate cancer cases presenting in radiation oncology, and ii) the relative ease of IMRT treatment planning for prostate cancer as compared to more challenging disease sites such as head and neck. It may be possible to extend this approach to other disease sites (i.e. head & neck) that are more challenging and time consuming to plan. Additionally, since head & neck cases typically present less frequently than prostate cases, many sites with limited experience and planning expertise could benefit greatly from the shared expertise of a knowledge-base system using cases collected from multiple institutions.

## References

1. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. *CA Cancer J Clin.* 2010;60(5):277-300.
2. Cancer Facts & Figures. Atlanta: American Cancer Society, 2007.
3. (ASTRO) ASTRO. Fast Facts About Radiation Therapy. 2008; Available from: <http://www.astro.org/PressRoom/FastFacts/documents/AboutRT.pdf>.
4. IMV. 2004 Radiation Oncology Market Summary Report2004.
5. Cahlon O, Hunt M, Zelefsky MJ. Intensity-modulated radiation therapy: supportive data for prostate cancer. *Semin Radiat Oncol.* 2008;18(1):48-57.
6. Webb S. Intensity-modulated radiation therapy. Bristol and Philadelphia: Institute of Physics Publishing, 2001.
7. Radiation Oncology Census Database [database on the Internet]. IMV. 2010.
8. Djajaputra D, Wu Q, Wu Y, Mohan R. Algorithm and performance of a clinical IMRT beam-angle optimization system. *Phys Med Biol.* 2003;48(19):3191-212.
9. Pugachev A, Li JG, Boyer AL, Hancock SL, Le QT, Donaldson SS, et al. Role of beam orientation optimization in intensity-modulated radiation therapy. *Int J Radiat Oncol Biol Phys.* 2001;50(2):551-60.
10. Das IJ, Cheng CW, Chopra KL, Mitra RK, Srivastava SP, Glatstein E. Intensity-modulated radiation therapy dose prescription, recording, and delivery: patterns of variability among institutions and treatment planning systems. *J Natl Cancer Inst.* 2008;100(5):300-7.
11. Philips Medical Systems website. Available from: [http://www.medical.philips.com/main/products/ros/assets/images/clinicalimages/IMRT/PROS-IMRT\\_Prostate-dose\\_747.jpg](http://www.medical.philips.com/main/products/ros/assets/images/clinicalimages/IMRT/PROS-IMRT_Prostate-dose_747.jpg).
12. Trotti A, Byhardt R, Stetz J, Gwede C, Corn B, Fu K, et al. Common toxicity criteria: version 2.0. an improved reference for grading the acute effects of cancer treatment: impact on radiotherapy. *Int J Radiat Oncol Biol Phys.* 2000;47(1):13-47.
13. Livsey JE, Routledge J, Burns M, Swindell R, Davidson SE, Cowan RA, et al. Scoring of treatment-related late effects in prostate cancer. *Radiother Oncol.* 2002;65(2):109-21.
14. Hanks GE, Schultheiss TE, Hunt MA, Epstein B. Factors influencing incidence of acute grade 2 morbidity in conformal and standard radiation treatment of prostate cancer. *Int J Radiat Oncol Biol Phys.* 1995;31(1):25-9.

15. Dearnaley DP, Khoo VS, Norman AR, Meyer L, Nahum A, Tait D, et al. Comparison of radiation side-effects of conformal and conventional radiotherapy in prostate cancer: a randomised trial. *Lancet*. 1999;353(9149):267-72.
16. Zelefsky MJ, Fuks Z, Hunt M, Yamada Y, Marion C, Ling CC, et al. High-dose intensity modulated radiation therapy for prostate cancer: early toxicity and biochemical outcome in 772 patients. *Int J Radiat Oncol Biol Phys*. 2002;53(5):1111-6.
17. Peeters ST, Heemsbergen WD, van Putten WL, Slot A, Tabak H, Mens JW, et al. Acute and late complications after radiotherapy for prostate cancer: results of a multicenter randomized trial comparing 68 Gy to 78 Gy. *Int J Radiat Oncol Biol Phys*. 2005;61(4):1019-34.
18. Michalski JM, Winter K, Purdy JA, Perez CA, Ryu JK, Parliament MB, et al. Toxicity after three-dimensional radiotherapy for prostate cancer with RTOG 9406 dose level IV. *Int J Radiat Oncol Biol Phys*. 2004;58(3):735-42.
19. Shu HK, Lee TT, Vigneau E, Xia P, Pickett B, Phillips TL, et al. Toxicity following high-dose three-dimensional conformal and intensity-modulated radiation therapy for clinically localized prostate cancer. *Urology*. 2001;57(1):102-107.
20. Brahme A. Optimization of Radiation-Therapy and the Development of Multileaf Collimation. *International Journal of Radiation Oncology Biology Physics*. 1993;25(2):373-75.
21. Brahme A. Optimization of Radiation-Therapy. *International Journal of Radiation Oncology Biology Physics*. 1994;28(3):785-87.
22. Brahme A. Optimization of the 3-Dimensional Dose Delivery and Tomotherapy. *International Journal of Imaging Systems and Technology*. 1995;6(1):1-1.
23. Brahme A. Development of radiation therapy optimization. *Acta Oncologica*. 2000;39(5):579-95.
24. Bortfeld T, Schlegel W. Optimization of beam orientations in radiation therapy: some theoretical considerations. *Phys Med Biol*. 1993;38(2):291-304.
25. Bortfeld T, Schlegel W, Dykstra C, Levegrun S, Preiser K. Physical vs. biological objectives for treatment plan optimization. *Radiother Oncol*. 1996;40(2):185-7.
26. Bortfeld T, Stein J, Preiser K. Clinically relevant intensity modulation optimization using physical criteria. *Proceedings of the Xiith International Conference on the Use of Computers in Radiation Therapy*. 1997:1-4495.
27. Webb S. Optimizing the planning of intensity-modulated radiotherapy. *Phys Med Biol*. 1994;39(12):2229-46.



28. Webb S. Configuration options for intensity-modulated radiation therapy using multiple static fields shaped by a multileaf collimator. II: constraints and limitations on 2D modulation. *Phys Med Biol*. 1998;43(6):1481-95.
29. Mohan R, Mageras GS, Baldwin B, Brewster LJ, Kutcher GJ, Leibel S, et al. Clinically Relevant Optimization of 3-D Conformal Treatments. *Medical Physics*. 1992;19(4):933-44.
30. Mohan R, Wu QW, Wang XH, Stein J. Intensity modulation optimization, lateral transport of radiation, and margins. *Medical Physics*. 1996;23(12):2011-21.
31. Webb S. The physical basis of IMRT and inverse planning. *Br J Radiol*. 2003;76(910):678-89.
32. Fox T, Applegate D, Cook W, Crocker I. Automated beam angle and energy selection for multicriteria IMRT treatment planning optimization. *Medical Physics*. 2003;30(6):1491-91.
33. Das S, Yin F. Efficient selection of beam number and orientations for intensity modulated radiotherapy (IMRT) by emulating an ideal dose distribution. *Medical Physics*. 2006;33(6):2096-96.
34. Niemierko A. Reporting and analyzing dose distributions: A concept of equivalent uniform dose - Response. *Medical Physics*. 1997;24(8):1325-27.
35. Mathworld. Available from: <http://mathworld.wolfram.com/>.
36. Varian Medical Systems. Palo Alto, CA; Available from: [http://www.varian.com/us/oncology/radiation\\_oncology/eclipse/brochures.html](http://www.varian.com/us/oncology/radiation_oncology/eclipse/brochures.html).
37. Webb S. Optimisation of conformal radiotherapy dose distributions by simulated annealing. *Phys Med Biol*. 1989;34(10):1349-70.
38. NOMOS Best. Pittsburgh, PA; Available from: [http://www.nomos.com/pdf/BST-001-Corvus\\_Brochure.pdf](http://www.nomos.com/pdf/BST-001-Corvus_Brochure.pdf).
39. Sterpin E, Tomsej M, De Smedt B, Reynaert N, Vynckier S. Monte carlo evaluation of the AAA treatment planning algorithm in a heterogeneous multilayer phantom and IMRT clinical treatments for an Elekta SL25 linear accelerator. *Med Phys*. 2007;34(5):1665-77.
40. Heath E, Seuntjens J, Sheikh-Bagheri D. Dosimetric evaluation of the clinical implementation of the first commercial IMRT Monte Carlo treatment planning system at 6 MV. *Med Phys*. 2004;31(10):2771-9.
41. Francescon P, Cora S, Chiovati P. Dose verification of an IMRT treatment planning system with the BEAM EGS4-based Monte Carlo code. *Med Phys*. 2003;30(2):144-57.

42. Liu BJ. A knowledge-based imaging informatics approach for managing proton beam therapy of cancer patients. *Technol Cancer Res Treat.* 2007;6(4 Suppl):77-84.
43. Lee KJ, Barber DC, Walton L. Automated gamma knife radiosurgery treatment planning with image registration, data-mining, and Nelder-Mead simplex optimization. *Med Phys.* 2006;33(7):2532-40.
44. Wu B, Ricchetti F, Sanguineti G, Kazhdan M, Simari P, Chuang M, et al. Patient geometry-driven information retrieval for IMRT treatment plan quality control. *Med Phys.* 2009;36(12):5497-505.
45. Wu B, Ricchetti F, Sanguineti G, Kazhdan M, Simari P, Jacques R, et al. Data-driven approach to generating achievable dose-volume histogram objectives in intensity-modulated radiotherapy planning. *Int J Radiat Oncol Biol Phys.* 2011;79(4):1241-7.
46. Han X, Hoogeman MS, Levendag PC, Hibbard LS, Teguh DN, Voet P, et al. Atlas-based auto-segmentation of head and neck CT images. *Med Image Comput Comput Assist Interv.* 2008;11(Pt 2):434-41.
47. Stapleford LJ, Lawson JD, Perkins C, Edelman S, Davis L, McDonald MW, et al. Evaluation of automatic atlas-based lymph node segmentation for head-and-neck cancer. *Int J Radiat Oncol Biol Phys.* 2010;77(3):959-66.
48. Strassmann G, Abdellaoui S, Richter D, Bekkaoui F, Haderlein M, Fokas E, et al. Atlas-based semiautomatic target volume definition (CTV) for head-and-neck tumors. *Int J Radiat Oncol Biol Phys.* 2010;78(4):1270-6.
49. Deasy JO, Blanco AI, Clark VH. CERR: a computational environment for radiotherapy research. *Med Phys.* 2003;30(5):979-85.
50. Cover TM, Thomas JA. *Elements of Information Theory.* New York: Wiley, 1991.
51. Li W. Mutual information functions versus correlation functions. *J Stat Phys.* 1990;60:823-37.
52. Studholme C, Hill DL, Hawkes DJ. Automated three-dimensional registration of magnetic resonance and positron emission tomography brain images by multiresolution optimization of voxel similarity measures. *Med Phys.* 1997;24(1):25-35.
53. Tourassi GD, Frederick ED, Markey MK, Floyd CE, Jr. Application of the mutual information criterion for feature selection in computer-aided diagnosis. *Med Phys.* 2001;28(12):2394-402.
54. Tourassi GD, Vargas-Voracek R, Catarious DM, Jr., Floyd CE, Jr. Computer-assisted detection of mammographic masses: a template matching scheme based on mutual information. *Med Phys.* 2003;30(8):2123-30.

55. Tourassi GD, Harrawood B, Singh S, Lo JY. Information-theoretic CAD system in mammography: entropy-based indexing for computational efficiency and robust performance. *Med Phys.* 2007;34(8):3193-204.
56. Tourassi GD, Harrawood B, Singh S, Lo JY, Floyd CE. Evaluation of information-theoretic similarity measures for content-based retrieval and detection of masses in mammograms. *Med Phys.* 2007;34(1):140-50.
57. Klein S, Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.W. elastix: a toolbox for intensity based medical image registration. *IEEE Transactions on Medical Imaging.* 2010;29(1):196-205.
58. RTOG 0126 - A Phase III Randomized Study of High Dose 3D-CRT/IMRT versus Standard Dose 3D-CRT/IMRT in Patients Treated for Localized Prostate Cancer 2008.
59. Milano MT, Constine LS, Okunieff P. Normal tissue tolerance dose metrics for radiation therapy of major organs. *Semin Radiat Oncol.* 2007;17(2):131-40.
60. Lawton CA, Michalski J, El-Naqa I, Buyyounouski MK, Lee WR, Menard C, et al. RTOG GU Radiation oncology specialists reach consensus on pelvic lymph node volumes for high-risk prostate cancer. *Int J Radiat Oncol Biol Phys.* 2009;74(2):383-7.
61. Jackson A, Marks LB, Bentzen SM, Eisbruch A, Yorke ED, Ten Haken RK, et al. The lessons of QUANTEC: recommendations for reporting and gathering data on dose-volume dependencies of treatment outcome. *Int J Radiat Oncol Biol Phys.* 2010;76(3 Suppl):S155-60.
62. Marks LB, Yorke ED, Jackson A, Ten Haken RK, Constine LS, Eisbruch A, et al. Use of normal tissue complication probability models in the clinic. *Int J Radiat Oncol Biol Phys.* 2010;76(3 Suppl):S10-9.

## Biography

Name: Vorakarn Chanyavanich  
Date of birth: May 20, 1970  
Place of birth: Bangkok, Thailand

## Education

Duke University, Durham, NC

- Medical Physics, Ph.D., expected spring 2011

Dissertation: "Knowledge-based IMRT Treatment Planning for Prostate Cancer"

Co-Advisors: Joseph Y. Lo, Ph.D. (Radiology)  
Shiva K. Das, Ph.D. (Radiation Oncology)

University of Colorado Health Sciences Center, Denver, CO

- Medical Physics, M.S., 1999

Thesis: "A Clinical Evaluation of Mammographic Image Quality"

Advisor: Edward Hendrick, Ph.D. (Radiology)

Stanford University, Palo Alto, CA

- Education, A.M., 1993  
Advisor: Myron Atkins, Ph.D.

Macalester College, St. Paul, MN

- Physics and Mathematics, B.A., 1992  
Advisor: Ray Mikkelsen, Ph.D.

## Honors and Awards

*Duke University*

Carey E. Floyd, Jr. Graduate Fellowship, Graduate Program in Medical Physics (2009)  
Named in honor of late program faculty member, Professor Carey E. Floyd, Jr., this is one of the program's highest awards. The Floyd Fellow is recognized as the top Ph.D. student in each class as judged by excellence in academic performance, creativity, and innovative thinking.

Director's Exemplary Service Award, Graduate Program in Medical Physics (2009)  
Graduate Fellowship, Graduate Program in Medical Physics (2006)

Research Fellowship, University of Colorado Health Sciences Center (1997-1999)  
Teaching Fellow, Ctr. For Talented Youth Program, Johns Hopkins Univ. (1994, 1995)  
Graduate Fellowship, Stanford Teacher Education Program (STEP) (1992-1993)  
Pi Mu Epsilon, Math Honor Society (1992)

## **Grants and External Support**

- Wallace H. Coulter Translational Partnership Award, 2008-2009 (\$120,000)
  - *co-PIs – Joseph Lo Ph.D., Shiva Das Ph.D.*
  - *Project: Knowledge-based Optimization of IMRT Treatment Planning for Prostate Cancer*
- Wallace H. Coulter Translational Partnership Award renewal, 2009-2010 (\$60,000)
  - *co-PIs – Joseph Lo Ph.D., Shiva Das Ph.D.*
  - *Project: Knowledge-based Optimization of IMRT Treatment Planning for Prostate Cancer*
- Renaissance Computing Institute of North Carolina (RENCI)- Applied Scientific and Information Visualization Program, 2010-2011 (\$12,000)
  - *co-PIs – Joseph Lo, Ph.D., Shiva Das, Ph.D.*
  - *Project: A Visualization Tool for Knowledge-Based Treatment Planning for Prostate Cancer*

## **Publications**

### **A. Refereed journals**

1. Vorakarn Chanyavanich, Shiva K. Das, W. Robert Lee, Joseph Y. Lo, “*Knowledge-based IMRT Treatment Planning for Prostate Cancer*”, accepted by Medical Physics, March 2011.
2. Vorakarn Chanyavanich, Matthew Freeman, Joseph Y. Lo, Shiva K. Das, “*Evaluation of mutual information as a surrogate for plan quality for knowledge-based treatment planning*”, (in preparation)
3. Vorakarn Chanyavanich, Georgia Tourassi, Joseph Y. Lo, Shiva K. Das, “*Implementation and Clinical Evaluation of Knowledge-Based IMRT Treatment Planning Across Multiple Institutions*” (in preparation)

### **B. Conference proceedings**

1. V Chanyavanich, M Freeman, S K Das, J Y Lo, “*Knowledge-Based IMRT Treatment Planning for Prostate Cancer,*” Med. Phys. **37**, 3215; AAPM 2010- Poster
2. M S Freeman, V Chanyavanich, S K Das, J Y Lo, “*A Linear Metric of Knowledge-Based IMRT Treatment Plan Quality for Prostate Cancer,*” Med. Phys. **37**, 3214; AAPM 2010- Poster