Brain Tumor Target Volume Determination for Radiation Therapy Treatment Planning Through the Use of Automated MRI Segmentation

Gloria Patrika Mazzara
University of South Florida

Follow this and additional works at: https://scholarcommons.usf.edu/etd

Part of the American Studies Commons

Scholar Commons Citation
Mazzara, Gloria Patrika, 'Brain Tumor Target Volume Determination for Radiation Therapy Treatment Planning Through the Use of Automated MRI Segmentation' (2004). Graduate Theses and Dissertations.
https://scholarcommons.usf.edu/etd/1153

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.
Brain Tumor Target Volume Determination for Radiation Therapy Treatment Planning

Through the Use of Automated MRI Segmentation

by

Gloria Patrika Mazzara

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
Department of Electrical Engineering
College of Engineering
University of South Florida

Co-Major Professor: Robert P. Velthuizen, Ph.D.
Co-Major Professor: Kenneth A. Buckle, Ph.D.
Laurence P. Clarke, Ph.D.
Stanley R. Deans, Ph.D.
John J. Heine, Ph.D.
Kent H. Larsen, Ph.D.

Date of Approval:
February 27, 2004

Keywords: automatic segmentation, glioma, magnetic resonance, knowledge guided, k-nearest neighbor

© Copyright 2004, Gloria Patrika Mazzara
DEDICATION

To Hans Christian Beyer for all your support.
ACKNOWLEDGEMENTS

This work was partially supported by 1999 RSNA (1) Seed Grant, entitled “Automatic brain tumor target volume for radiation treatment planning”.

I would like to thank Computerized Medical Systems, and especially Colin Sims, Product Manager, for providing the software and support that made this research possible. Thanks are also due to the personnel from Moffitt Cancer Center: the radiation oncologists (Dr. Pearlman, Dr. Greenberg, and Dr. Wagner) involved in the contouring process, Physicists and Dosimetrists from the Department of Radiation Oncology, Dr. Pearlman for his additional support with the clinical data, and Carol Johnson for her assistance in collecting and managing the patient data. I would like to thank Dr. Murtagh, for his help in contouring post-MR data, and the technologists from the Department of Radiology for their help collecting the MRI data.

I want to thank Larry Hall, Ph.D., Department of Computer Science, for providing access to their computer workstations and Matt Clark, Ph.D., for assisting with the KG data processing. Additionally, I gratefully acknowledge my current employer, Varian Medical Systems, and previous employer, Huff, Ferras, & Associates, since they provided me with the time and support that made this work possible. Finally, I would like to thank each of the members of my advisory committee for all of their help and support. And most importantly, I would like to thank my major professor Robert P. Velthuizen, Ph.D., my mentor, advisor, and friend, for working with me through all of these years despite the various life changes.
TABLE OF CONTENTS

LIST OF TABLES iv

LIST OF FIGURES v

ABSTRACT viii

CHAPTER I. INTRODUCTION 1

CHAPTER II. EVALUATING CURRENT PRACTICE: MANUAL OUTLINING BY RADIATION ONCOLOGISTS 8

II.1. Introduction: manual outlining in radiation oncology 8

II.2. Subjects 10

II.3. CT and MRI brain images and registration 12

II.4. Patient imaging techniques 15

II.4.a. CT scanning 15

II.4.b. MRI scanning 15

II.4.c. Image registration 16

II.5. Radiation oncologist tumor volume definition 18

II.6. Analysis of contouring variability 19

II.6.a. Intra-operator variability 19

II.6.b. Inter-operator variability 19

II.7. Results 21

II.7.a. Image registration 21

II.7.b. Operator time 21

II.7.c. Manual outline variability 21

II.8. Discussion 26

CHAPTER III. AUTOMATIC RADIATION VOLUME OUTLINES USING MRI SEGMENTATION METHODS 29

III.1. Introduction: contouring with MRI automatic segmentation methods 29

III.2. Principles of MRI segmentation methods 32

III.2.a. kNN segmentation method 32

III.2.b. KG segmentation method 33

III.2.c. MRI non-uniformity correction 34
III.3. Application of MRI segmentation methods to radiation oncology
   III.3.a. Segmentation methods
   III.3.b. Registration to CT coordinates system
   III.3.c. Contour extraction process
III.4. Accuracy of MRI automated segmentation methods
III.5. Results
   III.5.a. Computer segmentation processing time
   III.5.b. Contour extraction mechanism
   III.5.c. Automated GTV contour results
   III.5.d. Accuracy
   III.5.e. Excluded accuracy
   III.5.f. Analysis of contour variances
III.6. Discussion

CHAPTER IV. EFFECT OF SEGMENTATION METHODS ON RADIATION THERAPY TREATMENT PLANS

IV.1. Introduction: automatic segmentation and radiation therapy
IV.2. Radiation treatment procedure
IV.3. Radiation volume definition
IV.4. Clinical brain radiation protocol
IV.5. Treatment plans for volumes of physicians and segmentation methods
IV.6. Expert definition
IV.7. Treatment plan evaluation techniques in current clinical practice
   IV.7.a. Dose statistics
   IV.7.b. Cumulative dose volume histograms
   IV.7.c. Conformity index
IV.8. Treatment plan analysis methods for this study
   IV.8.a. Dose statistics
   IV.8.b. Cumulative dose volume histograms
   IV.8.c. Conformity index
IV.9. Post-treatment evaluation methods
IV.10. Results
   IV.10.a. Segmentation GTV volume expansion
   IV.10.b. Planning target volumes
   IV.10.c. Plan comparison
   IV.10.d. Dose statistics
   IV.10.e. Cumulative dose volume histogram
   IV.10.f. Conformity index
   IV.10.g. Cross-conformity indices
   IV.10.h. Post-treatment evaluation
IV.11. Discussion
LIST OF TABLES

Table 1. Patient demographics 11
Table 2. Physician intra-operator variability in percentages of total volume 23
Table 3. Physician inter-operator variability in percentages of total volume 23
Table 4. Accuracy of physicians and segmentation methods 50
Table 5. Excluded accuracy for physicians and segmentation methods 51
Table 6. Treatment planning parameters for each of the patients 75
Table 7. Patient post-treatment imaging studies 99
Table 8. Values of resulting PTVs for physicians and segmentations methods 108
Table 9. Average dosimetric values for PTV1 from the different treatment plans 109
Table 10. Average dose to normal brain irradiation for boost plans 111
Table 11. Conformity index values for physicians and segmentation methods 113
Table 12. Values of cross-conformity indices 114
Table 13. Data evaluation for post-MRI tumor volume 119
Table 14. Data evaluation for post-MRI necrosis volume 120
Table 15. Data evaluation for post-MRI damaged normal tissue volume 121
LIST OF FIGURES

Figure 1. Outline of project design 7
Figure 2. Comparison of CT and MRI tumor detection 12
Figure 3. Physician intra-operator variability 22
Figure 4. Physician inter-operator variability 24
Figure 5. KG automated segmentation process 35
Figure 6. Contour extraction process 37
Figure 7. Example of the contour extraction process 42
Figure 8. Example of KG contour extraction process 44
Figure 9. Example of contour smoothing result 45
Figure 10. GTV contours for patient 6 46
Figure 11. GTV contours for patient 3 47
Figure 12. GTV contour series from physician 1, kNN and KG for patient 4 48
Figure 13. Series of MR axial images for patient 9 53
Figure 14. 3D GTV volumes for physician 1, kNN and KG volumes for patient 11 55
Figure 15. Axial contours for physicians 1,2,3, kNN, and KG segmentation 56
Figure 16. ROC curves for all three physicians compared with kNN and KG 57
Figure 17. Linear accelerator, collimator and field size with multi-leaf collimator 68
Figure 18. Description of a 3D radiation therapy process 71
Figure 19. Illustration of volumes used in radiation therapy treatment planning 73
Figure 20. Brain clinical protocol

Figure 21. Description of initial and boost tumor volume delineation

Figure 22. Representation of typical cone down 4-field arrangement for 3D-CRT

Figure 23. Beam’s Eye View projection of two treatment fields

Figure 24. Isodose color wash on transversal, coronal and sagittal CT images

Figure 25. Representation of a 3D automatic PTV margin expansion

Figure 26. Schematic representation indicating generation of a cumulative DVH

Figure 27. DVH for different target volume coverage

Figure 28. Scale representing range of conformity index values

Figure 29. Evaluation strategy based on conformity indices

Figure 30. Target volume isodose spatial coverage

Figure 31. Schematic display of post-MRI validation process

Figure 32. GTV vertical expansion

Figure 33. Example of original GTV for kNN and KG compared to physician 1

Figure 34. Example of expanded GTV for kNN and KG compared to physician 1

Figure 35. Example of kNN segmentation PTV compared with physician 1 PTV

Figure 36. Example of KG segmentation PTV compared with physician 1 PTV

Figure 37. Comparisons of DVHs for patient 11 for PTV1 and normal brain tissue

Figure 38. Post-MRI and isodose coverage for tumor and necrosis for patient 11

Figure 39. Post-MRI and isodose coverage for tumor, necrosis and damaged tissue

Figure A1. Representation of intra-operator variability calculation
Figure A2. Representation of inter-operator variability calculation 138
Figure B1. Representation of accuracy calculation 139
ABSTRACT

Radiation therapy seeks to effectively irradiate the tumor cells while minimizing the dose to adjacent normal cells. Prior research found that the low success rates for treating brain tumors would be improved with higher radiation doses to the tumor area. This is feasible only if the target volume can be precisely identified. However, the definition of tumor volume is still based on time-intensive, highly subjective manual outlining by radiation oncologists. In this study the effectiveness of two automated Magnetic Resonance Imaging (MRI) segmentation methods, k-Nearest Neighbors (kNN) and Knowledge-Guided (KG), in determining the Gross Tumor Volume (GTV) of brain tumors for use in radiation therapy was assessed. Three criteria were applied: accuracy of the contours; quality of the resulting treatment plan in terms of dose to the tumor; and a novel treatment plan evaluation technique based on post-treatment images.

The kNN method was able to segment all cases while the KG method was limited to enhancing tumors and gliomas with clear enhancing edges. Various software applications were developed to create a closed smooth contour that encompassed the tumor pixels from the segmentations and to integrate these results into the treatment planning software. A novel, probabilistic measurement of accuracy was introduced to
compare the agreement of the segmentation methods with the weighted average physician volume. Both computer methods under-segment the tumor volume when compared with the physicians but performed within the variability of manual contouring (28% ± 12% for inter-operator variability).

Computer segmentations were modified vertically to compensate for their under-segmentation. When comparing radiation treatment plans designed from physician-defined tumor volumes with treatment plans developed from the modified segmentation results, the reference target volume was irradiated within the same level of conformity. Analysis of the plans based on post-treatment MRI showed that the segmentation plans provided similar dose coverage to areas being treated by the original treatment plans.

This research demonstrates that computer segmentations provide a feasible route to automatic target volume definition. Because of the lower variability and greater efficiency of the automated techniques, their use could lead to more precise plans and better prognosis for brain tumor patients.
CHAPTER I

INTRODUCTION

The brain is the body’s single most critical and indispensable component. Its central role in the functioning of the body explains the importance of improving treatment protocols for primary brain tumors, or gliomas, since this form of cancer results in significant neuronal losses which frequently result in the death of the patient. Brain cancer will strike down five to ten people out of every 100,000 \(^{(1)}\). An estimated 17,000 new malignant primary brain tumors were diagnosed in 2002 in the United States \(^{(2)}\). Brain tumors are the leading cause of death among childhood cancers, the second leading cause of cancer deaths in males between the ages of 20-39, and the fifth leading cause of cancer deaths in women between the ages of 20-39. The two year survival rate for patients diagnosed with glioblastoma multiforme (GBM), one of the most common malignant brain tumors, is less than 9%; and it is only 35% for patients with other types of brain tumors \(^{(3)}\). The rate of mortality due to brain malignancies remains high despite significant advances in early diagnosis brought about by the use of computed tomography (CT) and magnetic resonance imaging (MRI).

Accordingly, brain cancer is a particularly deadly disease which current treatment modalities are relatively ineffective in curing or even controlling. Current treatment protocols for gliomas generally call for tumor removal through surgical procedures
followed by irradiation of the tumor bed. This study focuses on the latter part of this common treatment regime.

External radiation therapy uses high energy x-rays externally delivered from multiple directions to damage critical biological molecules in cancerous cells. Unfortunately, when irradiating tumors, normal healthy cells may also receive damaging radiation. This irradiation could cause critical side effects, including loss of vision or reduced brain function. Thus, the goal of radiation therapy is to effectively irradiate the tumor area, or target volume, while minimizing the radiation dose absorbed by adjacent normal tissues and organs. A radiation treatment is considered successful if tumor control is achieved, i.e. cancer cells are inhibited from dividing or reproducing, without undue damage to surrounding healthy tissue. The long-term goal of radiation treatment is to prevent any tumor recurrence or growth.

Radiation treatment protocols have evolved to deliver radiation that conforms closely to the three-dimensional (3D) shape of the treatment volume; this treatment protocol is commonly referred to as three-dimensional conformal radiation treatment or 3D-CRT. The key to successful 3D-CRT treatment is achieving the highest possible degree of accuracy when defining the tumor volume (or target volume). One reason for the disappointing results of current brain radiation therapy techniques is an inability to sufficiently irradiate the target volume to kill the cancerous cells. Gliomas are relatively resistant to radiation when compared with other types of tumors (4). Research has showed that most brain tumor recurrences are located within the primary tumor area (5-8). Over the last few years, a number of studies have addressed the issue of brain radiation dose
escalation. Escalating the dose to the tumor area could increase local control thus prolonging the progression-free interval for such cases (3,9,10). This increase in dose delivery would be feasible only if the target area could be more clearly identified and segregated from surrounding healthy tissue for treatment purposes (9).

Radiation oncologists traditionally model the brain treatment target volume through a time-intensive manual procedure involving the outlining of the palpable or visible extent of the tumor (gross tumor volume, GTV) on numerous consecutive two-dimensional images, or “slices,” using either CT or MRI. While the trial and error part of modern treatment planning has, for the most part, been automated, the effort required from the radiation oncologists to produce accurate 3-D treatment plans has only been increased. Limitations remain inherent in the current methods of tumor delineation making it virtually impossible for different radiation oncologists to reproduce consistent results (11). Various studies have demonstrated the existence of this large variation in target volume definitions produced by different physicians using both CT and MR images (6,12,13). Inadequate 3-D target definition may lead to an inability to exploit the full potential of advanced treatment planning techniques or a loss of tumor control due to geographical misses with radiation doses (14). Additionally, imprecise identification of the tumor area prevent the types of dose escalations to the tumor bed which could significantly improve treatment efficacy (2,9,3).

Because of the highly technical yet repetitious nature of manual outlining, the process offers an excellent opportunity for the creation of an automated expert system utilizing a computerized system which could both reduce the subjectivity inherent in the
manual process and conserve the scarce time resources of radiation oncologists. Automatic segmentation of MR images offers the potential to accurately define complex treatment volumes, to speed the contouring process in radiation therapy treatment planning, and to provide a standardized and reproducible tumor measurement protocol that can be employed by geographically diverse facilities and physicians.

Automating the MR brain tumor segmentation process is extremely difficult due to the complexities inherent in modeling the tumor bed as well as changes in the magnetic resonance characteristics of normal tissues $^{15,16}$. Most reported MRI segmentation techniques have been pixel based, i.e., each image pixel is individually classified into a tissue class. At our institution, several techniques of MR segmentation have been developed and evaluated specifically for use with brain tumors. These methods use the information derived from several magnetic resonance contrasts (i.e. multi-spectral data). Supervised automated segmentation methods require an operator to select regions of interest (ROIs) on each slice of multi-spectral (MS) MRI data; these regions are in turn used to train the automated classifier. One of these methods, the “k nearest neighbor” (kNN) system, has been shown to perform better than other tested supervised methods and has been used by many researchers for automated brain segmentation $^{15,17,18}$. Unsupervised techniques of MR segmentation do not require operator input for the processing of each data set. To automate the tumor volume determination, Clark, Hall and Goldgof encoded knowledge of the pixel intensity and spatial relationships in the images to create a fully-automated segmentation system known as the Knowledge Guided (KG) Method $^{19,20}$. Both the kNN and KG
segmentation methods have been clinically applied as a technique for more accurately measuring tumor volume variation in the brain \(^{18,21}\). To date, no fully automatic protocol for outlining of treatment volumes for radiation therapy has been developed.

The goal of this study was to bring state-of-the-art engineering into the realm of radiation oncology treatment planning. This work evaluates the performance of kNN as a representative of operator assisted semi-automated segmentation and the KG system as a promising candidate for fully automated GTV determinations. Specifically, the goal was to ascertain if the two types of computer segmentations methods could be used for radiation therapy treatment planning, and, if not, which improvements are required to achieve this goal. This study developed a technique for incorporating the automated segmentation methods into radiation treatment planning and devised an evaluation mechanism that could be used to assess the effectiveness of these new contouring techniques. The study process is represented in Figure 1.

This dissertation is divided in three sections. The first section (Chapter II) evaluates the current practice of tumor outlining by radiation oncologists for radiation therapy treatment planning. An evaluation of the intra- and inter- operator variability was performed. The results of this evaluation form the basis for the second section of the study (Chapter III), which describes the methods for incorporating two different types of MRI segmentation techniques (kNN and KG) for radiation GTV definition into clinical radiation therapy treatment (a software toolkit was developed as part of this study to integrate the segmentation methods into radiation therapy planning). A study of the level of agreement of the automatic segmentation contours with respect to the physician
outlines was performed. Finally, a novel evaluation technique was developed to quantify results from the physician contours and perform the comparison with the segmentation contours. Additionally, this section presents conclusions on how the automated systems for generating segmentation volumes could be improved to reach the goal of producing reliable outlines of treatment volumes without any human input.

The final section (Chapter IV) uses the data and insights from the previous section to evaluate the automated segmentation routines in the clinical radiation therapy setting. A comparison of treatment plans designed using physician defined tumor volumes with treatment plans developed from computer segmentation was performed. The differences between physicians and automatic segmentations in terms of delivered dose were calculated. This final section also provides the design of a novel evaluation mechanism to determine the accuracy of treatment planning of automatic segmentations with respect to the actual patient treatment based on follow-up imaging (Figure 1). This evaluation scheme could be applied to any imaging modality and would serve the purpose of evaluating the applicability of any automatic contouring method for radiation therapy.
Figure 1.
Outline of project design
II.1. Introduction: manual outlining in radiation oncology

Three-dimensional conformal radiation treatment (3D-CRT) planning and delivery has the general goal of conforming the shape of a prescribed dose volume to the shape of a three-dimensional target volume. The success of the radiation treatment relies on the accurate determination of the target volume. Failure to assess the extent of the disease or to include too much normal tissue could lead to treatment failure or unnecessary damage of healthy tissues.

The radiation oncologist is the physician in charge of a patient’s radiation treatment. Radiation oncologists traditionally model the brain gross tumor target volume (GTV) through a time-intensive manual procedure involving the outlining of the tumor on numerous two-dimensional images or “slices”. The introduction of CT and MRI in the last few years has contributed to improvements in the accuracy of delineating tumor extensions. However, the variation in contouring of target volumes between different observers can range from 5% to 32% \(^\text{(12,13,22)}\). Therefore, significant limitations remain in both the accurate delineation of tumor volumes and in the ability of different radiation oncologists to reproduce consistent results \(^\text{(11)}\).
The aim of this chapter is to measure the variability of radiation oncologists in the manual definition of the gross tumor volume. The results of this analysis are used as the basis for comparing the results of automated segmentation methods with the physician contours. This chapter is divided into eight sections. Section II.2 gives the basis for patient selection and demographics. Sections II.3 and II.4 describe the basis of different imaging techniques used in the study. Sections II.5 and II.6 summarize the methodology used to create the contour definitions by physicians and to calculate the accuracy of the results, respectively. Section II.7 gives the results of contouring variability by radiation oncologists. Finally, Section II.8 discusses the results obtained in this chapter and their relevance to the remaining chapters of the dissertation.
II.2. Subjects

Pre-existing MRI and CT data of eleven patients with primary brain cancer (glioma) was used as the basis in this study. The study was approved under the University of South Florida (USF) institutional review board #5253 and required no patient informed consent since only existing data was used and recorded in such a manner that participants could not be identified. Patient selection was based on available cases collected over a period of one year with primary brain cancer (glioma) that had a pre-treatment MRI in our clinic and proceeded to have radiation therapy in Moffitt Cancer Center.

The demographics of the patient group are listed in Table 1. In conformance with the standard clinical protocol of Moffitt Cancer Center, these patients were imaged pre-surgery with MRI and post-surgery with both MRI and CT. The MRI was used to identify and delineate the tumor and the CT was used for radiation therapy treatment planning. Depending upon the treatment protocol selected for each individual patient, the MRI images used in connection with this study may have been taken either before or after surgery. The CT was used for 3D radiation treatment planning.
Table 1.
Patient demographics

<table>
<thead>
<tr>
<th>Case</th>
<th>Age</th>
<th>Sex</th>
<th>Diagnosis</th>
<th>Surgery (days from MRI)</th>
<th>MRI Type</th>
<th>Tumor Enhanced</th>
<th>RT (++) Start (days from MRI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>65</td>
<td>F</td>
<td>AO</td>
<td>3</td>
<td>Pre</td>
<td>No</td>
<td>165</td>
</tr>
<tr>
<td>2</td>
<td>52</td>
<td>F</td>
<td>O</td>
<td>1</td>
<td>Pre</td>
<td>No</td>
<td>70</td>
</tr>
<tr>
<td>3</td>
<td>63</td>
<td>F</td>
<td>GBM</td>
<td>1</td>
<td>Pre</td>
<td>Yes</td>
<td>28</td>
</tr>
<tr>
<td>4</td>
<td>69</td>
<td>F</td>
<td>GBM</td>
<td>3</td>
<td>Pre</td>
<td>Yes</td>
<td>19</td>
</tr>
<tr>
<td>5</td>
<td>62</td>
<td>F</td>
<td>GBM</td>
<td>-89</td>
<td>Post</td>
<td>Yes</td>
<td>16</td>
</tr>
<tr>
<td>6</td>
<td>47</td>
<td>M</td>
<td>GBM</td>
<td>-24</td>
<td>Post</td>
<td>Yes</td>
<td>10</td>
</tr>
<tr>
<td>7</td>
<td>52</td>
<td>M</td>
<td>AO</td>
<td>3</td>
<td>Pre</td>
<td>Yes</td>
<td>73</td>
</tr>
<tr>
<td>8</td>
<td>62</td>
<td>F</td>
<td>GBM</td>
<td>-17</td>
<td>Post</td>
<td>Yes</td>
<td>6</td>
</tr>
<tr>
<td>9</td>
<td>80</td>
<td>M</td>
<td>GBM</td>
<td>-2</td>
<td>Pre</td>
<td>Yes</td>
<td>23</td>
</tr>
<tr>
<td>10</td>
<td>47</td>
<td>F</td>
<td>GBM</td>
<td>-13</td>
<td>Post</td>
<td>Yes</td>
<td>7</td>
</tr>
<tr>
<td>11</td>
<td>79</td>
<td>F</td>
<td>GBM</td>
<td>-4</td>
<td>Post</td>
<td>Yes</td>
<td>7</td>
</tr>
</tbody>
</table>

(*) GBM = Glioblastoma multiforme  (+) Pre = pre-surgery MRI  
AO = Astrocytoma                  Post = post-surgery MRI  
O = Oligodendroglioma             (++) RT = radiation therapy
II.3. CT and MRI brain images and registration

Common for radiation therapy treatment planning, a series of transaxial (perpendicular to the long axis of the body) CT images are used for locating both the target volume and normal tissues, and for designing the treatment plan. However, the definition of brain tumor volume by CT studies has been found inadequate. Other imaging modalities, such as MRI, have made significant improvement in the detection and localization of brain lesions. It has been demonstrated by various researchers that MRI is superior to CT for diagnostic brain imaging \(^{(10,11,22,23)}\). MRI is more sensitive than CT in both lesion detection and in the margin delineation of gliomas, as demonstrated in the example of Figure 2.

![CT and MRI comparison](image)

Figure 2.
Comparison of CT and MRI tumor detection
However, CT is the preferred image modality for use in treatment planning since it is an x-ray based technique which captures the characteristics of the tissues (electron density) as seen by the treatment beams used in radiation oncology (6). CT images provide anatomical information necessary for verification of the treatment plans (14). Therefore, CT and MRI are complementary in providing information for brain radiotherapy treatment planning. The use of MRI for treatment volume definition together with the electron density information from the CT allows integration of anatomical information with tissue information for more accurate 3D treatment planning. Thus, the ability to correlate these two imaging modalities through an accurate system of image registration becomes essential to maximizing the effectiveness of brain treatment protocols.

Often the registration of different types of images is done manually, i.e. a radiation oncologist uses mental reconstruction to superimpose MRI information onto the CT data set. This difficult process, although in common practice, has been recognized as a source of inaccuracies and can lead to incorrect information regarding location of the tumor volume (22,24,25). Several computerized techniques have been developed to determine the transformation relating MR images to CT images (26-28). This process, known as “image registration”, is a system that determines the 3D coordinate transformation to map the two image sets maintaining optimal anatomical identity.

Frass et. al. describe some of the earliest registration work for treatment planning and the technical aspects of integration of MRI into CT-based radiation therapy planning. Their approach to image registration was with the use of external fiducials and contour
matching algorithms \(^{(26)}\). Other more recent computerized methods describe image registration based on point matching, surface matching, and interactive matching by using anatomical points, anatomical surfaces or outlines of anatomical structures \(^{(25,27)}\). The accuracy of these methods has been evaluated using phantoms and patient studies and are on the order of 1-2 mm.

The previously described registration methods are not automated. These depend on patient positioning and selection of reference points resulting in a time consuming process. A newer registration algorithm known as the Mutual Information (MI) algorithm provides a fully automated registration method without the need to define fiducial or anatomical points \(^{(29)}\). The algorithm aims to maximize mutual information of the two images. It attempts to match similarities in the overall shape of the distribution of image intensity in a data set. Quantitative accuracy of the MI registration algorithm has been validated by comparing results with registration based on external markers \(^{(29,30)}\).

Various studies have indicated that distortions in magnetic resonance images are on the order of 1-2 mm for the skull volume \(^{(31)}\). A recent study showed that errors in the location of fiducial markers on MRI after registration with CT were on the order of 1 mm \(^{(32)}\). This reported error includes components of the registration error and of the geometric distortion in MRI. These results suggest that the geometric distortions in MR images are of small magnitude and would not prevent accurate image registration.
II.4. Patient imaging techniques

II.4.a. CT scanning

The CT images were obtained using a Siemens CT HiQ spiral scanner (Siemens Medical Systems, Germany) with 512x512 pixel images taken at 4 mm spacing from the vertex (top of the head) through the treatment area and 8 mm slice thickness through the thyroid. Patients were immobilized using a customized mask together with a head rest (MedTec, Orange City, Iowa). This mask holds the patient in exactly the same position during each of the radiation therapy treatments. The CT treatment planning system includes MergeCom, the precursor to DICOM data communications.

II.4.b. MRI scanning

The MRI scan was performed on a different day than the CT scan. The patients were imaged in either a 1.5 Tesla GE Signa Horizon (General Electric Company, Milwaukee, WI) or a Siemens Magnetom Symphony with fast gradient systems using the standard multi-element head coil. The systems include perfusion imaging software as well as DICOM data communications.

The multi-spectral data set used for MRI segmentation consisted of 5 mm thick axial anatomical slices T1-weighted, proton-density- (PD) weighted, and T2-weighted images obtained with a field of view (FOV) of 220 mm and reconstructed to a 512 X 512 pixel image. The T1-weighted scans used for this study were obtained after administration of 0.1 mmol/kg body weight of gadolinium (Gd) MRI contrast material (Gd-DTPA) and using a standard spin-echo (SE) sequence with a repetition time (TR)/echo time (TE) = 400/8 or 525/17 ms. The PD images were acquired using a spin
echo PD-weighted sequence with TR/TE = 3000/7.5 ms or a fluid-attenuated inversion recovery (FLAIR) sequence with TR/TE=10002/147 ms or TR/TE=9000/110 ms. The T₂ images were acquired using TR/TE=3000/104 ms or TR/TE=4000/96 ms. Radiation oncologists used axial post-contrast T₁–weighted images to define GTV for cases of enhancing tumor or pre-contrast T₂-weighted images for cases where no tumor enhancement was seen.

II.4.c. Image registration

Both CT and MRI image sets were transferred to Hewlett Packard workstations running Computerized Medical Systems (CMS) 3-D treatment planning FOCUS software version 2.4.0. Each image set was then transferred to a Dell Inspiron 7000 laptop computer equipped with CMS software for image fusion (“Focal Fusion,” software release version 1.3) and contouring (“Focal Ease,” software release version 1.3.0). The Focal Software suite provides a fully integrated environment for PCs and the CMS planning computers, including a full featured contouring program for 3-D radiotherapy. The laptop computer was dedicated to this research.

The CT and MRI data were registered using the CMS Focal Fusion software. Registration was required to allow manual physician contouring on the MRI images. The Focal Fusion software utilizes the Mutual Information (MI) algorithm for fully automatic image registration, as described in the previous section. The software also incorporates a manual method for pre- or post- interactive adjustment of the registration (rotation and translation) to be coupled with the automated registration. The program writes out a file with a transformation matrix to convert MRI data to the CT coordinates.
The final MRI image transformation was evaluated and approved by a radiation oncologist who specialized in neuro-oncology. Qualitative assessment was achieved by overlay of CT and MR images in axial (“looking up”), coronal (front view), and sagittal (side view) planes. The pertinent interfaces (bone-tissue, tissue-air and bone-air as well as different soft tissue interfaces) were visually examined to ensure that all relevant interfaces were correctly aligned. Registration accuracy for all data sets was verified by measuring the displacement of three unique anatomical landmarks.
II.5. Radiation oncologist tumor volume definition

Three radiation oncologists were selected to perform manual contouring of the GTV. The expertise of each radiation oncologist is as follows: Physician one is a radiation oncologist specialized in neuro-oncology with ten years of clinical practice in radiation oncology and involvement with over 200 glioma brain tumor cases, including multi-center clinical trials for brain tumor treatment\(^{(4,33)}\). Physicians two and three are radiation oncologists with less specialization in brain tumors, but each with more than twenty years of experience in radiation oncology.

The reconstructed and registered MR images were used to define the GTV using the CMS Focal Ease software. The guidelines for contouring required the definition of the GTV (enhancing tumor) from which the clinical (CTV) and planning (PTV) target volumes would be expanded. The GTV was defined by the Gd contrast enhancement in T\(_1\) images or changes in the white matter (edema as defined by T\(_2\) MRI images). Each radiation oncologist performed three different GTV outlines on each image set for each of the 11 patients, resulting in a total of 33 contours. The three different outlining sessions for each physician were separated by approximately one month to prevent memory bias. The laptop computer was brought to each radiation oncologist’s location of choice. The time each of the radiation oncologists took for the outlining process was measured and recorded as part of this study.
II.6. Analysis of contouring variability

All analytical analysis and image data transformation was performed using programs developed with Interactive Data Language software version 5.4 (IDL, Research Systems Inc, Boulder, Colorado).

II.6.a. Intra-operator variability

The intra-operator variability was calculated by overlapping the three volumes defined on the same patient by the same radiation oncologist at roughly one-month intervals. The variability was then calculated as the ratio of the average disagreement, that is, the size of each volume minus the intersection of the three volumes, divided by the average size of the three volumes (see Appendix A). For example, if a radiation oncologist had identified the same target volume in the three sets of contours prepared for any single patient, then the variability for that patient would have been zero. The larger the difference between contours of the same physician, the larger this value becomes.

II.6.b. Inter-operator variability

The inter-operator variability was calculated using the nine sets of outlines for each of the eleven patients and then calculating the disagreement of each volume outline prepared by each physician for each patient with each volume outline prepared by each of the other two physicians for that same patient. This process was repeated for each patient to provide a data set comprised of the average disagreement between the three contours for each patient prepared by one physician with the other six sets of contours prepared by the other two physicians for that same patient (see details of calculation in
Appendix B). The greater the difference between the contours of different physicians, the larger this ratio becomes.
II.7. Results

II.7.a. Image registration

The CT and MRI images were registered using the automatic registration function of the software. The radiation oncologist specializing in neuro-oncology reviewed each case and, if necessary, performed an additional manual adjustment. In all cases a good visual match was achieved in the axial, coronal and sagittal views. This was confirmed by calculating the average 3D error over the eleven patients for three distinct anatomical landmarks. An average displacement of 1.5mm ± 1.5mm was found. The maximum individual displacement was out-of-image slice plane where resolution accuracy could not exceed the minimum CT slice spacing of 4 mm. These results are similar to previously reported image registration results (6,31).

II.7.b. Operator time

The time each radiation oncologist took to outline all eleven patients was recorded. It ranged from 4.0 to 6.5 hours in the aggregate, i.e., an average of approximately 30 minutes per patient. Physician one spent the most time outlining the patient contours resulting in an average of approximately 30 minutes while physician two spent the least time, averaging approximately 20 minutes per patient.

II.7.c. Manual outline variability

Reproducibility of the delineation of the GTV on the MRI scans by the same radiation oncologist (intra-operator variability) was assessed producing the results set forth in Figure 3 and Table 2.
Figure 3.
Physician intra-operator variability

Graph shows percentage of average physician intra-operator variability as a function of patient number for each physician. Information on MRI type (pre- or post- surgery) and the average tumor volume (cm$^3$) is included below each patient number.
Table 2.
Physician intra-operator variability in percentages of total volume

<table>
<thead>
<tr>
<th>Intra-operator variability</th>
<th>Physician 1</th>
<th>Physician 2</th>
<th>Physician 3</th>
<th>Average</th>
<th>Volume (cm$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>13 $\pm$ 5</td>
<td>26 $\pm$ 19</td>
<td>22 $\pm$ 23</td>
<td>20 $\pm$ 16</td>
<td>63 $\pm$ 33</td>
</tr>
<tr>
<td>Median</td>
<td>13</td>
<td>22</td>
<td>15</td>
<td>16</td>
<td>61</td>
</tr>
</tbody>
</table>

The intra-operator variability averaged 20% $\pm$ 16% over all 33 contour sets of the eleven patients. From Figure 3 it can be noted that the average of the reproducibility of the delineation of target volume was better in pre-operative cases (eighteen sets of contours: 15%) than in post-operative cases (fifteen sets of contours: 27%).

The difference among GTVs identified by the three radiation oncologists (inter-operator variability) was also assessed resulting in a total average of 28% $\pm$ 12% (Table 3 and Figure 4). The variability in the six pre-operative cases was 24% with a higher average ratio obtained for the post-operative cases, i.e., 32%.

Table 3.
Physician inter-operator variability in percentages of total volume

<table>
<thead>
<tr>
<th>Inter-operator variability</th>
<th>Physician 1-2</th>
<th>Physician 1-3</th>
<th>Physician 2-3</th>
<th>Average</th>
<th>Volume (cm$^3$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>30 $\pm$ 11</td>
<td>23 $\pm$ 11</td>
<td>30 $\pm$ 14</td>
<td>28 $\pm$ 12</td>
<td>63 $\pm$ 33</td>
</tr>
<tr>
<td>Median</td>
<td>26</td>
<td>21</td>
<td>27</td>
<td>23</td>
<td>61</td>
</tr>
</tbody>
</table>
Figure 4.
Physician inter-operator variability

Graph shows percentage of average physician inter-operator variability as a function of patient number for each physician. Information on MRI type (pre- or post- surgery) and the average tumor volume (cm$^3$) is included below each patient number.

It should be noted that one out of the eleven patients (i.e., patient 6) had a tumor which proved abnormally difficult to contour and represented a case of total resection.

From Figure 3 and 4, a large variation can be observed for this patient. The MRI used
was post-operative and the enhancement boundaries were not clear due to cystic formation inside the resected area. This case was kept in the study since it represents cases encountered in the clinic and will serve to test if segmentations could have any potential for contouring all types of brain tumors treated with radiation therapy. Notice from the variability analysis than even though there was a large variation for Patient 6, the median is close to the average and is within the standard deviation for both intra- and inter-operator variability (see Tables 2 and 3).
II.8. Discussion

In this section of the study it was found that the radiation oncologist who took the most time for outlining achieved the smallest intra-operator variability, i.e. 13% (the average for all of the oncologists was 20%), and the one who took the least time in outlining produced the largest intra-operator variability, 26%.

The variation between different radiation oncologists, or the inter-operator variability, ranged from 11% to 69% with an average variability rate of 28%. These results mirror previously published results\(^{(12,6,13)}\) and show that there is significant uncertainty in target volumes definition even when such volumes are determined by a single radiation oncologist observing the same set of data on multiple occasions. The variability in delineation of GTV was about 10% larger in post-operative cases than in pre-operative cases. In post-operative cases, the margins of residual tumor are unclear making the identification of the GTV a difficult task. Previous studies confirm similar results\(^{(6,12)}\).

The presented results confirm published findings that variability in tumor contouring by human experts is high. Ten Haken and coworkers ran a simple test to assess the variation in the definition of contours by a team of physicians (neuro-radiologist and radiation oncologist) when defining tumor volume in CT, MRI, and MRI-CT registered images for 15 patients. It was found that after two iterations of the contouring, the definition of tumor volumes were smaller and averaged just 75% of average volumes indicated in the first set of contours\(^{(22)}\). This difference was obtained by comparing the second set of volumes to the first set of volumes. Notice that only two
contouring iterations were used to quantify the intra-operator variability, compared to three used in this study. Yamamoto and coworkers measured intra- and inter-operator GTV brain contouring variability that averaged 8% and 15%, respectively (12). These outlines were performed by four radiation oncologists on nine (pre- and post-surgery) brain CT images. Their study goal was to assess the error in radiation field definition based on the target reconstruction. Therefore, the intra-operator variability was obtained by comparing the smallest rectangular field that encompassed the two physician outlines. Similar to the previous study, only two iterations were considered to quantify the intra-operator variability. Using the same evaluation mechanism, the target inter-operator variability was obtained by the ratio of the rectangular projection surrounding a single physician contour divided by the smallest rectangular field surrounding the projection of the intersection of all the physician contours. This approach would underestimate the actual error in contouring since it doesn’t consider the outlining, but its rectangular enclosure. Nevertheless, it showed the implications of contour variability in the definition of radiation treatment fields.

A recent study reported very large inter-observer variations in brain GTV delineation for nine physicians of different specialties (radiation oncologist, neurosurgeon and radiologist) performing contours in both CT alone and CT with MRI for five patients (13). The variability in their study was obtained from the standard deviation based on the mean of the volumes delineated by different observers. This approach is similar to the variability method of this study. Their results found that the volumes varied between ±30% of the mean volume. This contouring error is larger than the set-up variations and
organ motions that are traditionally taken into account in radiation therapy planning. All of these results demonstrate the need for a method of tumor volume outlining that is more consistently reproducible.
CHAPTER III
AUTOMATIC RADIATION VOLUME OUTLINES USING MRI SEGMENTATION METHODS

III.1. Introduction: contouring with MRI automatic segmentation methods

As previously noted, one of the critical factors to influence the success of radiation therapy three-dimensional treatment planning is the accurate determination of the target volume. The current practice involves the manual contouring of the treatment area by radiation oncologists. This manual method of separating image data is operator-dependent, relies heavily on human judgment, and is very time consuming. As shown in the preceding chapter, it is virtually impossible for different radiation oncologists to reproduce consistent tumor volume results using this methodology, despite accurate image registration\(^\text{(11,12)}\).

Average inter-operator variability in outlining gross tumor volume (GTV) by radiation oncologists was found to be 28% ± 12% (see Chapter II). Similarly, various studies have found contouring variation between different physicians of up to 32%\(^\text{(22,12,13)}\). The incorporation of automatic segmentation methods to determine the target volume could reduce this inherent variability in contouring. This would aid in multi-center treatment trials since it would prevent physician- and center- bias from affecting trial outcomes.
Various supervised and unsupervised MRI segmentation methods for brain tumor volume definition have been developed and tested for various clinical applications, including the measurement of brain tumor volume changes during radiation therapy. These methods have not been applied in the delineation of tumor volume for radiation therapy treatment planning. The objective of this chapter is to present the application of two MRI segmentation techniques, k-nearest neighbor (kNN) and Knowledge Guided (KG), in delineating brain tumor volume for use in radiation therapy treatment planning. The specific goals are to find the level of agreement of the segmentation methods with respect to the outlines developed by the radiation oncologists and to identify shortcomings and recommend improvements necessary to reach the goal of fully automatic GTV determination for radiation therapy.

This chapter is divided into six sections. Section III.2 presents the basic principles for the MRI segmentation methods under study. Section III.3 describes the steps followed to apply the MRI segmentation methods to radiation therapy treatment planning, including the development of custom software programs and algorithms. These custom applications included: extraction of volumes of interest, manipulation of image data and segmentation results to match treatment planning image coordinate system, extraction of contour points, and file formatting to adapt results of segmentation to the treatment planning software. Section III.4 describes the methodology used to calculate the accuracy of the segmentation methods in delineating GTV for use in radiation treatment planning. Since it is customary for radiation oncologists to delineate the GTV used in radiation therapy, all comparisons were performed against the physician outlines.
presented in Chapter II. Due to the variation in physician contouring encountered, a novel, probabilistic measurement of accuracy was introduced for this evaluation. Section III.5 presents graphical and analytical results of the contours produced by the automated segmentation methods as compared with those generated by the physicians. True-positive and false-positive ratios for the computer segmentations were obtained and compared to those found for the radiation oncologists. Section III.6 discusses the results obtained from this analysis and gives the conclusions on the application of segmentation methods for contouring in radiation therapy. Additionally, information is presented on the necessary modifications to these computer techniques to be applied in clinical practice.
III.2. Principles of MRI segmentation methods

Each square on an image matrix is called a pixel. The goal of segmentation methods is to group pixels into specific tissue regions or classes. There have been various algorithms or classifiers developed to identify and separate different brain anatomic structures from MRI images\textsuperscript{(15,34)}. Segmentation methods can be termed as supervised or unsupervised. Supervised methods require an operator to select regions of interest (ROIs) on each slice of image data, which are then used to train the classifier. One of these methods is the “k nearest neighbor” (kNN) approach. Unsupervised segmentation methods are classifiers that do not require the use of training data, such as the “knowledge guided” (KG) algorithm.

Both of these methods use the information derived from several magnetic resonance contrasts (i.e. multispectral data). This higher dimensional feature space improves discrimination for tissues with similar relaxation times and improves the image segmentation.

III.2.a. kNN segmentation method

The kNN is a supervised segmentation method that uses operator selected training data to identify each tissue type. It requires the labeling of a certain number of pixels in each slice of multi-spectral MRI. The kNN algorithm starts a pixel classification by finding the k labeled pixels from the ROIs closest to that pixel. It then classifies it into the same class as the majority of the k-closest training data. The kNN method uses pattern recognition for tissue classification, i.e. resolves the signal magnitude distribution as displayed in the different MR images into a probability map of tissue types\textsuperscript{(35)}. It is
considered a nonparametric classifier because it makes no underlying assumption about 
the statistical structure of the data\textsuperscript{(34)}.

Previous studies have shown that the success of supervised segmentation methods 
depends on the classifier and the training data. The kNN method performed better than 
other tested supervised methods and has been used by many researchers for MRI 
segmentation\textsuperscript{(15,18)}. Results from these studies indicate that relative measurements of 
tissue volume are feasible by independent segmentation of serial MRI data sets to within 
a maximum error of ± 6\% for the kNN method. Additionally, the intra- and inter-
operator variation arising from the training data selection was found to be 6\% and 4\%, 
respectively\textsuperscript{(17)}.

III.2.b. KG segmentation method

The kNN method discussed above is strictly based on pattern recognition in 
feature space. To fully automate the tumor volume determination, Clark, Hall and 
Goldgof encoded knowledge of the pixel intensity and spatial relationships in the images 
to create a fully-automated segmentation system known as the Knowledge Guided (KG) 
method\textsuperscript{(19,20)}. First, the KG expert system was trained to identify slices of MR images of 
the brain that contain pathology from slices that do not contain pathology\textsuperscript{(19)}. This work 
was successfully extended to provide volume measurements of normal white matter and 
gray matter. The KG system’s current incarnation is able to identify glioma tumor tissue 
after gadolinium enhancement in the brain\textsuperscript{(20)}.

The KG system has five steps to perform the segmentation process (see Figure 5): 
(a) detect abnormal slices (i.e. containing tumor) by finding deviations from expected
normal properties within the slice; slices free of abnormalities are not processed any further; (b) extract the intracranial region from the rest of the MR image based on information provided in the first step; (c) initial tumor segmentation produced by a combination of adaptive histogram thresholds in the T₁ and PD feature images, (d) removal of non-tumor pixels by a density screening operation on the basis that pixels of normal tissues are grouped more closely together in feature space than tumor pixels; and (e) complete tumor segmentation by analyzing each spatially disjoint region in image space separately, removing regions free of tumor and those regions remaining labeled as tumor (20).

III.2.c. MRI non-uniformity correction

Many contributions to the literature on MRI segmentation have stressed an artifact of magnetic resonance imaging that the intensity of image pixels not only depends on the tissue type, but also on the location in the image, or rather in the radio-frequency (RF) antenna (coil). Several methods have been described to correct for the resulting image nonuniformity, mostly based on separation of low-frequency information in the image through smoothing or filtering. A recent study has shown that these corrections do not separate the nonuniformity from the image information (36). More importantly, it was found that although correction techniques have a significant effect on pixel intensities, the effect on brain segmentation results is non-significant. The tumor segmentation methods considered in this research are robust for RF nonuniformity effects since brain tumors are confined to small volumes and intensity nonuniformity is a slowly varying function of location.
Figure 5.
KG automated segmentation process
Reprinted with permission by IEEE Transactions on Medical Imaging (20).
III.3. Application of MRI segmentation methods to radiation oncology

III.3.a. Segmentation methods

The kNN segmentation method was run on UNIX workstations (Sun Microsystems, Mountain View, CA) and the KG system was run on a SPARCstation Ultra 20 computer (Sun Microsystems, Mountain View, CA). Both systems were purchased around 1996. A network and software environment was used in which MRI, CT, Focus treatment planning systems, the laptop computer and the image processing lab were all integrated to allow for convenient flow of images and other data between platforms.

The kNN segmentation method requires the user to select training data from each MRI slice. Since research has already been done on the variation arising from data selection by multiple users, this variable was not further evaluated in this study (17). In the present study, the researcher, a certified medical physicist, selected the training data. The KG system requires no user input when performing the segmentation process to extract the tumor pixels.

The results from the segmentations included scattered tumor-labeled pixels in addition to the main body of pixels identified as “the” tumor. Consistent with the previously reported studies, pixels from the segmentation results that were clustered together were selected for tumor classification and the scattered individual pixels were discarded.
III.3.b. Registration to CT coordinates system

Final results from both MRI segmentation methods were transformed to the CT coordinate system using the transformation matrix produced by the Focal Fusion Registration software to allow comparison with the GTV outlines prepared by the radiation oncologists. The use of the same transformation matrix for both physicians and segmentations eliminates any error arising from the image registration.

Figure 6.
Contour extraction process
III.3.c. Contour extraction process

The results from the segmentation methods several pixels labeled as “tumor”. A software toolkit was developed to extract contours from these results. Various routines were developed to perform this contour extraction, some of which can be found in Appendix C. The goal was to find a closed smooth contour that tightly enclosed all the tumor pixels from the segmentation, as presented in the previous figure. Figure 6 summarizes the different steps or programs used in obtaining automated tumor volume contours. All programs were developed with Interactive Data Language software version 5.4 (IDL, Research Systems Inc, Boulder, Colorado).
III.4. Accuracy of MRI automated segmentation methods

The difference between brain GTV delineation for different physician specialties (radiation oncology, radiologists, neurosurgeons) has been reported by Weltens et al. \(^{(13)}\). The goal of this chapter is to evaluate the segmentation methods as possible tools for delineating GTV for treatment planning of brain tumor volumes. Since it is customary for radiation oncologists to delineate the GTV used in radiation therapy, the contours obtained from the physicians in Chapter II were used as the basis for comparison. Due to variation in contouring, a probabilistic interpretation of the volumes delineated by the radiation oncologists provided the basis for comparing both the individual physician and the automated segmentation systems.

Specifically, the probability that a pixel in an image is properly classified as part of the tumor volume was determined by the number of times that pixel was included in the nine outlines prepared by the three physicians. Every pixel in the image volume was labeled with an integer value corresponding to the number of physician contours in which it was included, e.g., if a pixel was never included in any physician outline its corresponding value would be zero whereas a pixel included in every physician outline would have a value of nine. The resulting composite physician GTV is comprised of pixels labeled with values from zero to nine that define the probability of finding tumor volume.

The pixel label provided the weight for measuring accuracy. This analysis was done on a pixel-by-pixel basis. Thus, the loss of accuracy associated with a failure to classify a pixel as being part of a tumor volume would decrease in proportion to the
weight associated with that pixel. For example, the failure to include a pixel that was assigned a label of 9, i.e. a pixel selected every time, would reduce accuracy more than missing a pixel that was assigned a label of one, i.e., a pixel selected only once by the physicians.

Final accuracy for the computer segmentation is then defined as the ratio of the total sum of weights contained within the computer segmentation volume to the total weights generated from the nine volumes produced by the physicians (see Appendix B). This approach measures the level of agreement between the automated systems and the physicians on whether any given region of tissue should be characterized as part of the gross tumor volume, i.e., it provides the study’s true positive rate for the automated contouring systems. The same protocol was used to determine the accuracy of each contoured volume produced by each individual physician.

To estimate the false positive rate, i.e., the level of agreement between physicians on healthy tissue which was incorrectly characterized as constituting part of the GTV, the study calculated the excluded accuracy volume in an analogous manner (Appendix B). All analytical analysis was performed using programs developed with Interactive Data Language software version 5.4.
III.5. Results

III.5.a. Computer segmentation processing time

The time to perform a kNN segmentation averaged 20 minutes per patient with some variation based upon the number of slices evidencing enhanced tumor and the difficulty of selecting the training data for the kNN segmentation algorithm. For the KG system, the only time required was in the preparation of the MRI scans for segmentation. Specifically, in the case of the KG system, approximately 1.5 hours of operator time was required to prepare the necessary data for automatic segmentation.

It should be noted that the data preparation task for the KG system could be substantially or fully automated in the future. This would reduce the human operator time currently required by the segmentation system. The process of automatic segmentation using the KG system required approximately thirty minutes of computer time on for all patients and required no user input. Note that computers have increased in speed the last few years, which would result in faster processing times for the computer segmentation methods.

III.5.b. Contour extraction mechanism

The computer segmentations resulted in pixels from which contours were extracted following the technique described in Figure 6. The following figure represents in detail the process of contour extraction mechanism.
Figure 7.
Example of the contour extraction process
Figure 7 shows the steps involved in the contour extraction process for an image of patient #3 using the kNN process. The top (a) images show the segmentation results and resulting GTV outline. This figure shows the segmentation case which resulted in “sections” of tumor labeled pixels. It can clearly be noted that the resulting contour from the pure segmentation could not be used for radiation therapy because of the resulting multiple small contours instead of a single contour enclosing the pixels. The second set of images (b) shows a polar coordinate system representation of that segmentation. This section of the algorithm used the center of mass of the image pixels to calculate the distance of each boundary point. This is represented as black dots in the figure. For each radian angle section, the farthest point was found. An interpolation was performed to obtain the missing radian angle points. An averaging of the maximum radius value for each fraction of angles was obtained to smooth the contour. The resulting contour in polar coordinates is displayed as a solid line in red. The conversion to the image coordinate system of this contour is represented next to the polar figure. The third set of figures (c) show the resulting area labeled tumor as a result of filling the image with the polar coordinate process, followed by the contour extraction and the integration of the outline in the treatment planning image. The last set of images (d) show one of the outlines for each physician (from left to right, Physician 1,2,3). The automatic extraction of contour as the result of the software program created in this research resulted in a contour extraction that is comparable to the contours produced by the physicians. Similar results were obtained for the KG process, which are demonstrated in Figure 8.
Figure 8.
Example of KG contour extraction process

Figure shows the steps for contour extraction process using KG for an image of patient #5. (a) Top images show the segmentation results and resulting GTV outline, (b) second set of images show the results from filling the image with the polar coordinate system, extracting the contour and integrating the outline in the planning image, and (c) show one of the outlines from each physician (from left to right, Physician 1,2, and 3).
It can be observed from the previous figures that these routines also create a smoothing effect for the contours. This result can be seen in more detail in Figure 9. This figure shows an example of a case where the tumor segmentation resulted in a continuous section of tumor pixel classification. The application of the contour extraction routine served as a smoothing mechanism.

Figure 9. Example of contour smoothing result

Figure shows the smoothing effect resulting from implementing the subroutines for contour extraction. First image (left) shows the result from the segmentation, continuing with the contour extraction, filling the image and final smoothed contour (image on right). See Appendix C for details on the contour extraction algorithm.

III.5.c. Automated GTV contour results

For the automated segmentation methods it should be noted that the KG algorithm was not designed to evaluate non-enhancing tumors (tumors not visible on T1-weighted MRI) such as those encountered in connection with Patients 1 and 2. For Patients 3 and 6, the KG method identified tumors in very few of the slices that had physician outlines for tumor. Patient 6 had cystic formation inside a partly enhanced area so the margins were not clear (see Figure 10). The MRI used was post-operative and the enhancement boundaries were not clear due to cystic formation inside the resected area.
Figure 10.
GTV contours for patient 6

MR image of patient 6 showing a GTV contour from physicians 1 (a), 2 (b), and 3 (c); and kNN segmentation (d). Notice the variability between all physician GTV contours and close agreement of kNN segmentation with Physician 1. The MR image shown is a T1 axial scan after application of Gd contrast.

The kNN performed well for this difficult case involving a cystic formation inside a partly enhanced area since the user selects the training data. Notice in Figure 10 that all three physicians contoured an area beyond the contrast enhancing area. This is due to information from the previous image slice. The kNN method misses additional tumor volume since it does not use 3D information, i.e. volume information from previous and following images.
Patient 3 had dentures and implants that caused artifacts in the images, making the automatic segmentation difficult (Figure 11). In this case, there was also cystic formation inside the enhanced area. The kNN method was able to segment this contour and obtain results within the contours produced by the physicians. The KG was unable to segment most of these slices.

Figure 11.
GTV contours for patient 3

MR image of patient 3 showing a small teeth artifact effect next to tumor containing Gd enhancement with non-enhancing cystic necrotic centers. Contours shown are GTV contours of (a) physician 1, (b) physician 2, (c) physician 3, and (d) kNN segmentation.
Figure 12.
GTV contour series from physician 1, kNN and KG for patient 4

Series of contours from physician 1 (a), kNN (b) and KG (c) segmentation. Notice the excellent results for the segmentation methods at the middle slices and the failure found at inferior slices.

It was additionally found that the KG segmentation method performed poorly for sections of the tumor that were located in the lower part of the brain. This was the case for Patients 4 and 9 where the lower axial scans were not identified properly regardless of the enhancing property of the tumor. For both of these patients, the KG missed the last slices.
of the enhancing tumor. This effect can be observed in Figure 12, which shows a series of lower brain axial slices spaced 4 mm. Note also that the enhancement and tumor volume margins are not as clear in the lower slices. Knowledge from previous slices is desirable to correctly identify tumor volume.

III.5.d. Accuracy

As described in section 4 of this chapter, the results from the physicians were used as the basis for assessing the accuracy of the radiation oncologists and the MRI segmentation methods to determine GTV for brain radiation therapy brain. It was previously noted that the same transformation matrix was used for comparing contours generated by the physicians and the segmentation methods therefore eliminating any errors arising from the image registration.

The results of the accuracy calculation in percentages of total volume are tabulated in Table 4. The kNN method gave an average accuracy of 59% for pre-operative scans compared to the 52% average obtained for post-operative scans. For the KG method three of the pre-operative cases were non-enhancing tumors or had cystic formation (i.e., Patients 1, 2, and 3) and were unable to be segmented, as discussed previously. The average accuracy of the three physicians is 85% ± 7%, compared to 56% ± 6% for the kNN method and 52% ± 7% for the KG method, resulting in a difference from the physicians’ contours of 29% and 33% for the kNN and KG method, respectively. Comparing this difference with the average inter-operator variability of 28% ± 12% for all eleven cases (with a range of variability between physicians from 17
percent to 60 percent (Figure 4), the automated segmentation methods are within the variability range of the physicians.

Table 4.
Accuracy of physicians and segmentation methods

<table>
<thead>
<tr>
<th>Patient #</th>
<th>MRI Type</th>
<th>Average Volume (cm³)</th>
<th>Accuracy Physician 1</th>
<th>Accuracy Physician 2</th>
<th>Accuracy Physician 3</th>
<th>Accuracy kNN method</th>
<th>Accuracy KG method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pre</td>
<td>28</td>
<td>86</td>
<td>87</td>
<td>89</td>
<td>57</td>
<td>n/a</td>
</tr>
<tr>
<td>2</td>
<td>Pre</td>
<td>90</td>
<td>92</td>
<td>82</td>
<td>94</td>
<td>67</td>
<td>n/a</td>
</tr>
<tr>
<td>3</td>
<td>Pre</td>
<td>98</td>
<td>92</td>
<td>95</td>
<td>91</td>
<td>62</td>
<td>n/a</td>
</tr>
<tr>
<td>4</td>
<td>Pre</td>
<td>40</td>
<td>86</td>
<td>92</td>
<td>89</td>
<td>58</td>
<td>48</td>
</tr>
<tr>
<td>5</td>
<td>Post</td>
<td>108</td>
<td>86</td>
<td>90</td>
<td>91</td>
<td>62</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>Post</td>
<td>73</td>
<td>71</td>
<td>63</td>
<td>57</td>
<td>52</td>
<td>n/a</td>
</tr>
<tr>
<td>7</td>
<td>Pre</td>
<td>100</td>
<td>89</td>
<td>94</td>
<td>88</td>
<td>57</td>
<td>63</td>
</tr>
<tr>
<td>8</td>
<td>Post</td>
<td>63</td>
<td>84</td>
<td>88</td>
<td>89</td>
<td>51</td>
<td>54</td>
</tr>
<tr>
<td>9</td>
<td>Pre</td>
<td>36</td>
<td>74</td>
<td>91</td>
<td>78</td>
<td>52</td>
<td>43</td>
</tr>
<tr>
<td>10</td>
<td>Post</td>
<td>17</td>
<td>79</td>
<td>87</td>
<td>88</td>
<td>48</td>
<td>60</td>
</tr>
<tr>
<td>11</td>
<td>Post</td>
<td>39</td>
<td>91</td>
<td>89</td>
<td>86</td>
<td>46</td>
<td>48</td>
</tr>
</tbody>
</table>

Average: 63 ± 33  84 ± 7  87 ± 9  85 ± 10  56 ± 6  52 ± 7

It is important to note that the design of this study defined “true” volume by using the GTV generated by the same three radiation oncologists to whom the automated
systems were compared; accordingly, it would not be possible within this conceptual framework for the accuracy of the automated systems to have exceeded that of their human counterparts. This limitation of the study is explored further as part of Future Work in the Discussion Section, below.

Table 5. review pacer for updated docket (0.3); review pleadings received and update files accordingly for attorney use (1.6); Excluded accuracy for physicians and segmentation methods

<table>
<thead>
<tr>
<th>Patient #</th>
<th>MRI Type(+)</th>
<th>Physician 1</th>
<th>Physician 2</th>
<th>Physician 3</th>
<th>kNN method</th>
<th>KG method</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Pre</td>
<td>8</td>
<td>20</td>
<td>15</td>
<td>7</td>
<td>n/a</td>
</tr>
<tr>
<td>2</td>
<td>Pre</td>
<td>13</td>
<td>6</td>
<td>19</td>
<td>4</td>
<td>n/a</td>
</tr>
<tr>
<td>3</td>
<td>Pre</td>
<td>5</td>
<td>18</td>
<td>4</td>
<td>2</td>
<td>n/a</td>
</tr>
<tr>
<td>4</td>
<td>Pre</td>
<td>5</td>
<td>25</td>
<td>9</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>Post</td>
<td>7</td>
<td>18</td>
<td>15</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>Post</td>
<td>34</td>
<td>58</td>
<td>32</td>
<td>41</td>
<td>n/a</td>
</tr>
<tr>
<td>7</td>
<td>Pre</td>
<td>6</td>
<td>23</td>
<td>7</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>Post</td>
<td>8</td>
<td>19</td>
<td>17</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>9</td>
<td>Pre</td>
<td>5</td>
<td>62</td>
<td>11</td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>Post</td>
<td>6</td>
<td>26</td>
<td>24</td>
<td>12</td>
<td>23</td>
</tr>
<tr>
<td>11</td>
<td>Post</td>
<td>14</td>
<td>16</td>
<td>8</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Average 10±9  26±17  15±8  8±11  8±8
III.5.e. Excluded accuracy

The accuracy measure we introduced favors larger volumes since there is no penalty in the measure for “false positive” pixels. This false positive effect is expressed as a ratio based on the volume included by the computer segmentations and excluded by the physicians’ modeling volume; these results are set forth in Table 5. All values shown are in percentages of total volume. Interestingly, the false positive rate of the kNN method was 8% ± 11% and of the KG method was 8 ± 8% while the false positive rate of the physicians was 17% ± 11%. Thus, the automated segmentation methods have been shown to err on the side of underestimation of tumor volume when compared with the physicians.

III.5.f. Analysis of contour variances

The two segmentation methods were assessed visually and quantitatively to evaluate where the major volume differences occurred between the contours delineated by the physicians and the automated systems. This analysis should prove useful in suggesting further studies with larger sample sizes, which could significantly improve the accuracy of automated contouring systems for radiation oncology. In general, the largest variations between the contouring of the physicians as a group and those produced by the automated systems were found at superior and inferior edges of tumor. This effect can be seen clearly in Figure 12, which shows enhancing tumor in the most inferior slice which was not identified properly by the kNN method and which the KG method completely failed to detect. For slices in the central sections of tumors, both segmentation methods provided contours that were much closer to those created by
human experts. Another example of this effects can be seen in Figure 13, which shows the three superior slices for Patient 9 and contours drawn by the physicians in the top slice which shows small tumor enhancement.

Figure 13.
Series of MR axial images for patient 9

Images (a), (b), and (c) show the top three slices of tumor, 4 mm spaced. Notice how image (c) does not show much enhancement, but by comparing to previous images, some tumor volume can be seen. The tumor volume drawn the first time by each physician on the MR slice shown on (c) is displayed in images (d), (e) and (f) for physician 1, 2, and 3 respectively.
The images (a) through (c) of Figure 13 demonstrate that the contour on image (c) was included by the radiation oncologists mainly due to the physicians’ knowledge that there is tumor in the previous slices at the corresponding locations. Notice the variations in the contours between the physicians. Interestingly for this specific case, one physician did not draw any GTV on a slice he had marked on first contouring as containing tumor the next two times he was faced with identical data. The segmentation methods did not identify tumor volume on image (c). The kNN segmentation program is currently limited to 2-D identification protocols, i.e. each individual slice is analyzed for tumor without considering adjacent slices. The KG method, while taking information from prior slices forward, is essentially also based on 2-D analysis. The drawing of contour edges by human experts is a very subtle and subjective activity blending scientific training with heuristics developed through experience with a variety of tumors contoured over many years.

Figure 14 shows a 3-D reconstruction of the tumor volume drawn by physician 1 (outer yellow volume) and the GTV estimated by the kNN (Figure 14a) and the KG (Figure 14b) systems. It can be seen clearly that the physician volume contains the GTV contours produced by the segmentation methods and that the segmentation methods fail to identify tumor most frequently in the superior and inferior edges of a tumor. The kNN and KG methods have a larger agreement with the physician towards the center of the tumor compared with the superior and inferior borders. The physician volume
contains the segmented GTV. It can be noticed that the segmentation methods under-
segment the GTV volume compared to the radiation oncologist.

Figure 14.
3D GTV volumes for physician 1, kNN and KG volumes for patient 11

The 3D reconstructed images for patient 11 showing contours of physician 1 (outer yellow volume), kNN (inner red volume in figure 14a), and KG (inner red volume in figure 14b).
In addition to the differences found at the superior and inferior edges of the tumor volume, the contours prepared by the segmentation systems diverge from those of the human experts to an increasing degree the further from the center of a tumor the contour lines are compared. For the areas where the oncologists agreed all the time (pixels labeled 7, 8 or 9), towards the center of the tumor, the accuracy for kNN and KG was 75% and 72%, respectively. Figure 15 represents the general effect encountered of decreasing agreement between the contours produced by the automatic segmentation methods and the physicians as compared from the center of the tumor towards its outside borders. Additionally, it demonstrates that computer segmentations tend to agree with the contours prepared by radiation oncologists when the radiation oncologists agree with each other.

Figure 15.
Axial contours for physicians 1, 2, 3, kNN, and KG segmentation

Axial slice shows contouring of physicians 1, 2 and 3 (outer contours represented in red, orange and pink, respectively), and kNN segmentation (innermost contour represented in blue), and KG segmentation (inner contour represented in light blue).
Figure 16 shows a section of the Receiver Operating Characteristic (“ROC”) curves for all three physicians compared with those for the kNN and KG systems. The ROC curve is a plot of the true positive rate (TP) against the false positive rate (FP). It shows the tradeoff between sensitivity (portion of accurate TP) and specificity (portion of accurate true negative (TN)) since any increase in sensitivity will be accompanied by a decrease in specificity. The closer a curve follows the left-hand border and top border of the ROC space, the more accurate the test.

Figure 16.
ROC curves for all three physicians compared with kNN and KG
True tumor (TP) for this study is based on the times a pixel was included in an outline by the physicians. The curves for the segmentation methods must necessarily be below those of the physicians since the latter defined “truth” for purposes of this study. It can be seen that automated segmentation systems tend to fail in sensitivity but have a high degree of specificity, that is, most pixels identified are within the volume of the tumor defined by the physicians. This is also evidenced by the data summarized in Tables 4 and 5, and Figure 15.
III.6. Discussion

The purpose of this study was to evaluate KG and kNN as potential “cyber colleagues” for radiation oncologists in determining tumor volume definition for treatment planning. A probabilistic measure of accuracy was proposed accounting for the inherent variability in operator judgment. It was found in the previous chapter the average intra- and inter- operator variability in outlining by radiation oncologists to be $20\% \pm 16\%$ and $28\% \pm 12\%$, respectively. The underlying ranges of data displayed a high level of variability, a range of 8% to 91% and 11% to 69% for the intra- and inter-operator variability, respectively. The average accuracy of the kNN segmentation method was found to be $56\% \pm 6\%$ for all 11 cases while the KG method yielded accuracy of $52\% \pm 7\%$ for the seven cases compared with the physician contours. Based on the range of variability found for the physicians, it is clear that both segmentation methods performed within the variability of the outlining by radiation oncologists. Even without any of the system enhancements suggested herein, the automated segmentation methods could qualify as independent experts since they perform within the large range of inter-operator variability found among radiation oncologists. Combining the information obtained from the segmentation methods could help radiation oncologists, especially those with limited experience, identify the target volume with greater accuracy.

The variability in physician delineation of GTV was found to be about 10% larger in post-operative cases than in pre-operative cases (Chapter II). Similarly, a larger variation was found for post-operative cases using the automatic segmentation methods.
In post-operative cases, the margins of residual tumor are unclear making the identification of the GTV a difficult task for both physicians and automated segmentation systems.

The segmentation methods were less accurate at the superior and inferior edges of the tumor volume compared to the middle sections. A greater difference was found, at least in part, from the small enhancement found at the edges of the tumor. The drawing of contour edges is very subtle and subjective. The radiation oncologists use a 3-D method of contouring, that is, one in which the previous and subsequent 2-D slices are used to predict the presence of tumor volume on the slide in question. This knowledge needs to be included in the segmentation methods to improve accuracy, i.e., a 3-D segmentation method is needed, which uses knowledge and pixel information of tumor from adjacent 2-D slices. While the KG method has such encoding, the subtleties of partial volume effects at the edges are currently problematic.

Compared to the kNN method, the KG method performed poorly for glioma cases that show Gd enhancement with non-enhancing cystic necrotic centers. The margins of the tumor are not clear for these cases and even the physicians’ contours show a larger intra- and inter-operator variability for these cases. A possible solution is to use more knowledge from the T2 MRI images in the automatic segmentation process. Additionally, the KG method failed to detect tumor volume located in the lower part of the brain. The KG system performs differently in different areas of the brain because it has rules describing the anatomy at the various levels through the brain. Anatomical structures are simpler in superior areas of the brain and increase in complexity towards
inferior sections. The kNN method was able to give better results to these cases since there is some user input selecting the initial tumor pixels and slices from which the kNN method began its segmentation analysis.

Two cases that showed non-enhancing tumor volumes were not segmented by the KG method. It is necessary to incorporate automatic segmentation of non-enhancing brain tumors in the knowledge guided technique. Some work has been performed in developing an automatic method that separates non-enhancing brain tumors from healthy tissues in MRI images showing promising results (37).

In summary, there is need for more work on the KG method to make it fully compatible for use in radiation therapy; this work would include modifications to permit contouring of partially enhancing tumors, resection cavities and non-enhancing tumors. The user guided kNN method performed better under these special circumstances due to its use of user input for initial selection of training pixel data.

Future work should concentrate on optimizing the segmentation techniques to improve the accuracy of their results, especially with respect to the definition of the inferior and superior borders of the tumor volume. Note that different glioma types were included in this research to generate the basis of possible future applications of brain segmentation methods. It would be of interest to perform more in-depth analysis on the variability of segmentation methods based on the type of brain tumors by selecting a specific type of tumor patients (glioblastoma, astrocytoma or oligodendroglionomas) and study its tumor specific segmentations results. Both segmentation methods considered herein should also be enhanced to allow them to identify edema and structures at risk.
This would permit future incarnations of these automated systems to provide outlines for such structures and to assist physicians in automatic detection of normal tissues.

Additional research should also be performed which analyzes the results of the automated segmentation methods in a way that does not favor the physicians involved in the comparison. It was previously noted that the contours from the radiation oncologists used for comparison were also used to define the “true” volume. This posed a limitation in that the accuracy of the computer segmentation methods would always fall below that of the physicians. A recent study recommends cooperation with a radiologist and/or neurosurgeon to reduce the variability in tumor volume definition \(^{(13)}\). Incorporating a second group of physicians (preferably radiation oncologists, neurosurgeons and radiologists) as experts working together to define “true” gross tumor volume based on their mutual consensus would allow subsequent studies to fairly compare the accuracy of automated segmentation systems with that of radiation oncologists evaluating the same data.
CHAPTER IV

EFFECT OF SEGMENTATION METHODS
ON RADIATION THERAPY TREATMENT PLANS

IV.1. Introduction: automatic segmentation and radiation therapy

The aim of three-dimensional conformal radiation therapy (3D-CRT) is to give a high dose of radiation to a carefully delineated target volume, i.e., to the tumor, while minimizing the dose received by the adjacent healthy tissues. This is feasible only if the region of interest (i.e., the target volume) can be clearly differentiated from the surrounding healthy tissue (9). The patterns discernable from failed treatment protocols for brain gliomas utilizing high-dose 3D-CRT have shown that over 89% of tumor recurrences were located within the primary tumor bed (5,6,8). Escalating the dose to this viable tumor area might increase survival rates for such cases by reducing the incidence of local tumor recurrence. The rationale for such dose escalations in 3D-CRT brain tumor treatment relates to the dual goal of simultaneously sparing normal brain tissue while delivering high target volume doses of radiation. Accordingly, the success rates achieved through high quality radiation treatments can be increased through more precise tumor localization. The goal of implementing automatic tumor segmentation methods in radiation therapy planning is to accurately determine the target area to allow dose escalation if necessary and to reduce the amount of physician time required to create superior treatment plans.
It was previously found that the kNN and KG computer segmentation methods under-segment the tumor volume when compared with the weighted average of outlines produced by radiation oncologist but are within the variability of the contouring performed by experienced radiation oncologists based on the same data (see Chapter III). Additionally, the automated segmentation methods show a large degree of both specificity and accuracy in identifying central areas of the GTV.

The general aim of this chapter is to study the results of applying the two automated MRI segmentation techniques, kNN and KG, to the creation of radiation therapy treatment plans. The first step in this process was to integrate the knowledge obtained from the previous chapter to modify the GTV generated by the automated segmentation methods. Following this modification, a margin was applied to create a planning target volume (PTV). The PTV in radiation oncology results from the application of a margin to the GTV to account for patient motion and variations in daily treatment setup. This volume served as the basis for the creation of treatment plans for the contours produced by the automated segmentation methods.

The evaluation of the resulting treatment plans was performed starting from the following hypothesis: Radiation treatment plans developed from computer segmentations that are modified with a vertical extension are comparable in accuracy to the plans developed from physician contours. This hypothesis was tested by comparing hypothetical radiation therapy treatment plans based on volumes generated by computer segmentation methods to the plans generated using volumes produced by experienced radiation oncologists.
This chapter is divided into eleven sections. Sections IV.2 and IV.3 outline common concepts in the practice of radiation oncology, specifically brain radiation treatment. Section IV.4 describes the brain radiation technique that was used to actually treat the patients whose data underlies this study. Section IV.5 defines the methodology used herein to create hypothetical treatment plans for physician and the respective automated segmentation methodologies. This section also explores the application of insights gained in the previous chapter to modify the GTV resulting from the segmentation methods for application in radiation therapy treatment planning. Section IV.6 provides information regarding the selection of an expert physician to be used as the basis for comparison to test the above hypothesis. Sections IV.7 and IV.8 explore the different treatment evaluation techniques used to validate the foregoing hypothesis. They also introduce a unique evaluation scheme developed in connection with this study to evaluate the performance of any automated segmentation methods in radiation therapy planning. Section IV.9 evaluates the methods used to test the accuracy of incorporating automatic segmentation methods for radiation treatment planning by using existing post-treatment MR patient images. The results from the original patient treatment were compared with the hypothetical treatment plans based on physicians and segmentation methods. An evaluation technique that could be applied to any imaging modality was developed to perform this comparison. The results are presented and evaluated in Sections IV.10 and IV.11.
IV.2. Radiation treatment procedure

The field of radiation therapy is diverse and complex. This section is intended to familiarize the non-specialist reader with sufficient elementary information pertaining to radiation therapy to permit understanding of the design and results of this study. The references cited herein may be consulted for more in-depth knowledge on various aspects of the study.

The principal application of radiation therapy is the use of radiation to treat cancer. Radiation, if carefully controlled, can cure or control cancer by inhibiting the cancer cells from dividing or reproducing, thus achieving tumor control. The primary goal of radiation therapy is to produce the greatest possible control of the tumor while minimizing to the greatest extent possible all undesired side effects. Early research established that different organs and tissues react differently to irradiation. The radiation dose that should be directed toward the defined target is determined after considering several factors. Important factors include whether the treatment is to be curative (return the patient to health) or palliative (reduce disease discomfort, increase patient’s quality of life), the tumor’s location, the tumor’s histology, and the risk of side effects from exposure to radiation. Radiosensitivity of organs near the tumor is also of major importance since it determines how much radiation can be safely delivered to the area (38). For example, brain irradiation is limited when giving dose to areas close to the optical structures (such as optic chiasm, optic nerve, or eye lens). These structures can be damaged if irradiated to the total radiation doses used to treat gliomas.
Radiation therapy works best when radiation is administered through relatively small doses delivered over the course of several sessions. In this way, it can kill the tumor cells and yet allow sufficient time for the normal healthy cells around the tumor to repair any damage caused by the radiation. A typical course of radiation treatment lasts from two to eight weeks, depending on the total dose of radiation to be delivered. The treatment dose is defined in unit of “Gray” (Gy), which represent the amount of energy (1 Joule) absorbed per kilogram of tissue. For example, a regular CT exam delivers approximately 0.02 Gy. The radiation dose administered during a typical course of brain treatment ranges from 45 to 60 Gy with daily doses ranging from 1.8 to 2.0 Gy.

Radiotherapy for the brain is most commonly administered in the form of x-rays generated by linear accelerators. These machines are designed to produce different x-ray energies (between 6 and 22 MV) for use in cancer treatment. Different angles of treatments can be achieved by rotating the machine stand or “gantry”. Linear accelerators produce radiation that exits outside a “collimator” (see Figure 17). The collimator system is designed to vary the shape of the beam by using two pairs of heavy metal blocks that can be moved independently to produce a rectangle-shaped radiation field or “field size.” The rectangular treatment area can be further modified with the use of “blocks” or “multi-leaf collimators”. Blocks or MLCs are used to cover up normal tissues inside the treatment field size to insure that radiation is being delivered principally to the tumor area. Additionally, special filters may be placed in the path of the radiation beam to modify its dose distribution. The most commonly used device is the “wedge”, which is a wedge-shaped absorber that causes a progressive decrease in the intensity of
the radiation beam \(^{(39)}\). This modifier is used to create dose uniformity over slanting target surfaces.

Figure 17.
Linear accelerator, collimator and field size with multi-leaf collimator
During the initial patient consultation the radiation oncologist determines the area of treatment by using diagnostic images, such as x-rays, CT, or MRI. Prior to the widespread availability of x-ray computerized tomography (CT), whole-brain radiation therapy (RT) was routinely used to treat brain tumors, since conventional imaging techniques were unable to accurately assess the extent of tumor⁵. Three-dimensional (3D) treatment planning evolved through technological advancement in computers that allowed for improved target modeling in three-dimensions. Three-dimensional (3D) conformal radiation treatment (3D-CRT) planning and delivery seeks to conform the shape of a prescribed dose volume to the shape of a 3-dimensional planning target volume (PTV) while simultaneously limiting the radiation dose to normal tissues. 3D-CRT achieves this goal through the use of an external beam radiation therapy modality. Radiation treatment planning involves the use of highly specialized software and computers in designing the various treatment beams and shields to precisely localize and focus the radiation administered to the tumor. All such treatment planning systems have to be approved by the Food and Drug Administration (FDA) prior to their use in a clinical setting. Modern treatment planning systems are capable of preparing a patient treatment using a 3-D model of the patient, reconstructed from CT information¹⁴.

The treatment procedure for a patient undergoing 3D brain radiation therapy treatment is demonstrated in Figure 18. The procedure typically commences with the definition of the treatment volume by a radiation oncologist. A MRI together with a CT is used to outline the GTV brain volume by the radiation oncologist. Additionally, the physician contours the organs at risk of receiving radiation (such as the optic chiasm and
brain stem), and other structures of interest. A CT based radiographic volume in the
treatment planning system provides the foundation for all contour definitions. The
treatment plans are designed and calculated by a “dosimetrist”.

The dosimetrist selects the appropriate gantry angles, collimator angles, field size
and field modifiers in order to ensure coverage of the target volume with the dose
prescribed by the radiation oncologist. Amongst the software tools that have been
developed to aid in this process is the “beam’s eye view” display. It consists of a
projection of the patient anatomy as seen from the location of the radiation source. From
the beam’s eye view, it is possible to identify the best angles from which to irradiate the
target and to avoid irradiating adjacent sensitive normal tissue (14). Once the beam
directions and field sizes have been selected, a block, or MLC aperture, is defined to
protect normal tissues. Several studies have identified the tolerance levels of various
normal tissues to therapeutic irradiation (40,38). These factors are considered when
designing the aperture of the treatment beam. Various treatment planning systems offer
the ability to automatically design this aperture. Once the ideal arrangement of treatment
beams is directed to the patient, the radiation oncologist evaluates this plan to ensure that
the dose coverage is adequate for the intended treatment. This review is performed using
the graphical display that shows the different lines of equal dose (“isodoses”) and the
quantitative dose evaluation tools of the treatment planning software. A simulation is
then performed to verify the feasibility of the proposed treatment plan and to document
the beam portals using radiographs. Following a successful simulation, the patient starts
the actual radiation treatment.
Figure 18.
Description of a 3D radiation therapy process
IV.3. Radiation volume definition

In 3D-CRT treatment planning, the treatment volume is divided into various subsections. These different volumes often have to be irradiated using different dose levels. To avoid ambiguity in the definition of the target volumes, the International Commission on Radiation Units and Measurements (ICRU) created the ICRU 50 report that defines the different treatment volumes to be used for radiation treatment planning (41).

The palpable or visible extent of the tumor constitutes the Gross Tumor Volume (GTV). As previously discussed in Chapter II, the radiation oncologist manually outlines this tumor volume. In some cases, a margin is added around the GTV to include direct, local sub-clinical spread, i.e. an area that may not visibly show tumor but could contain cancerous cells. This volume constitutes the Clinical Target Volume (CTV). The delineation of the GTV and CTV are based purely on clinical considerations without regard to technical factors of treatment. To obtain the final volume used in treatment planning, margins have to be increased around the GTV or CTV to account for variation in size and position of tissues relative to the treatment beams due to patient positioning, internal organ motion, and to variations in the daily treatment set up (see Figure 19). This volume is the Planning Target Volume (PTV). In addition to the treatment volumes, normal critical structures, i.e. anatomical structures that can be damaged by a certain amount of radiation, must be defined in total to ensure that their radiation tolerances are not exceeded by the proposed treatment protocol (38).
Figure 19.
Illustration of volumes used in radiation therapy treatment planning

(Adapted from ICRU Report 50\textsuperscript{(41)})
IV.4. Clinical brain radiation protocol

This dissertation is based on data from eleven brain cancer patients that were under the care of physicians at Moffitt Cancer Center. This section describes the actual radiation treatment received by these patients. Section 5, below, includes the description of the “virtual” treatments developed to estimate the efficacy of the computer segmentations (see Figure 1, Chapter I).

The eleven patients under consideration in this study underwent 3-D radiation treatment following surgery. Since 3D-CRT has been used at Moffitt Cancer Center for more than 10 years, a certain consistency in target definition and planning has been developed and refined over this period. The following figure summarizes the radiation treatment received by these patients.

Radiotherapy for brain tumor patients usually begins within four weeks following brain surgery. Radiation treatment was delivered with 6 MV or 15 MV photons from one of three Siemens linear accelerators (M6700, KD or Primus, Siemens Medical Systems, Malvern, PA). One treatment of 1.8 Gy or 2.0 Gy was given daily, 5 days per week, for a total irradiation in the range of 54.0 – 60.0 Gy over six weeks. The patients were initially treated for 7 to 10 fractions with opposed lateral fields of 6 MV and 15 MV to deliver a total dose in the range of 12.6-20 Gy to the initial planning volume (PTV1).
Subsequently, a 3D conformal plan was prepared with two to four fields conformed the PTV1. A total treatment dose in the range of 45.0 – 46.8 Gy was delivered to the PTV1. Finally, a similar arrangement was used for the cone down treatment volume (PTV2) in the final 5 to 8 fractions to provide a dose in the range of 9.0 – 14.4 Gy. This final dose is usually referred to as “boost” (see Table 6).

Table 6.
Treatment planning parameters for each of the patients

<table>
<thead>
<tr>
<th>Patient #</th>
<th>PTV margins (cm)</th>
<th>No. of boost fields</th>
<th>Boost dose (Gy)</th>
<th>No. of treatment fractions</th>
<th>Total boost dose (Gy)</th>
<th>Total treatment dose (Gy)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.5</td>
<td>3</td>
<td>1.8</td>
<td>7</td>
<td>12.6</td>
<td>59.4</td>
</tr>
<tr>
<td>2</td>
<td>1.0</td>
<td>3</td>
<td>1.8</td>
<td>5</td>
<td>9.0</td>
<td>54.0</td>
</tr>
<tr>
<td>3</td>
<td>1.0</td>
<td>4</td>
<td>2.0</td>
<td>7</td>
<td>14.0</td>
<td>60.0</td>
</tr>
<tr>
<td>4</td>
<td>2.0</td>
<td>3</td>
<td>2.0</td>
<td>7</td>
<td>14.0</td>
<td>60.0</td>
</tr>
<tr>
<td>5</td>
<td>2.0</td>
<td>3</td>
<td>2.0</td>
<td>7</td>
<td>14.0</td>
<td>60.0</td>
</tr>
<tr>
<td>6</td>
<td>2.0</td>
<td>4</td>
<td>2.0</td>
<td>7</td>
<td>14.0</td>
<td>60.0</td>
</tr>
<tr>
<td>7</td>
<td>1.0</td>
<td>4</td>
<td>1.8</td>
<td>8</td>
<td>14.4</td>
<td>59.4</td>
</tr>
<tr>
<td>8</td>
<td>2.0</td>
<td>4</td>
<td>2.0</td>
<td>7</td>
<td>14.0</td>
<td>60.0</td>
</tr>
<tr>
<td>9</td>
<td>1.5</td>
<td>2</td>
<td>2.0</td>
<td>7</td>
<td>14.0</td>
<td>60.0</td>
</tr>
<tr>
<td>10</td>
<td>2.0</td>
<td>4</td>
<td>1.8</td>
<td>8</td>
<td>14.4</td>
<td>59.8</td>
</tr>
<tr>
<td>11</td>
<td>1.5</td>
<td>4</td>
<td>2.0</td>
<td>7</td>
<td>14.0</td>
<td>60.0</td>
</tr>
</tbody>
</table>
A pre- or post- surgery MRI together with a CT scan were used by the radiation oncologist to identify the GTV volumes for the eleven patients in this study. The GTV was manually outlined on the digital CT images with the aid of the hardcopy MRI. The radiation oncologist defined the PTV from a margin around the CTV or GTV, following the recommendations in ICRU Report 50 (as discussed above). The radiation oncologist decided upon the margin definition used to create the final PTV after considering actual tumor pathology, tumor location, and treatment protocol. For the patient treatment, the PTV is defined and changed as appropriate throughout the course of treatment. The initial Planning Target Volume (PTV1) for the first part of the radiation treatment (23-26 treatment sessions or “fractions”) included the contrast-enhancing lesion (Gross Tumor Volume, GTV) and surrounding edema (CTV) demonstrated on the MRI plus a 1.0 cm to 2.0 cm margin. For the last 5 to 8 fractions, a cone down target volume (PTV2) was used, which included the GTV plus a 1.0 cm to 2.0 cm margin. The cone down treatment (boost) aims at giving dose to the main tumor only, resulting in the GTV plus a margin (see Figure 21).

Additionally, several structures of interest and organs at risk were outlined to aid in the definition of treatment fields and to limit the dose to critical organs. The dosimetrists defined the skull, eyes, optic nerves, optic chiasm, brain stem, and normal brain volume for all of the patients in the study. These contours were checked and approved by the radiation oncologist prior to finalizing the planning. The dose received by these areas is documented and used for treatment planning decisions.
Figure 21.
Description of initial and boost tumor volume delineation

(a) Green (inner) contour: initial tumor volume (CTV=GTV plus edema) and
   Yellow (outer) contour: PTV1 (initial CTV plus margin for treatment)
(b) Blue (inner) contour: cone down tumor volume (only GTV) and
   White (outer) contour: PTV2 (final boost volume GTV plus margin)

The patient treatment plans were created on a FOCUS treatment planning system
(Versions 2.4-2.5, Computerized Medical Systems, St. Louis, MO). All PTV volumes
were generated from the CTV or GTV using the 3D automatic margin growth function
from the treatment planning system software. Research has shown that the PTV can
automatically be derived from manually drawn volumes to aid in the treatment planning
process and to increase the reproducibility and accuracy of the treatment\(^{42,43}\). The use
of automated volume expansion has been proven to be an invaluable aid for the final
estimation of the treatment volume by physicians. Brain PTV for the patients was defined by entering the desired margin around the GTV. The automatic growth was limited or stopped at the skull bone. If necessary, the automatically generated PTV volume was trimmed manually by the radiation oncologist around critical structures such as the optic chiasm and optic nerves.

The 3D radiation treatments were performed using static multiple, non-coplanar 6 MV or 15 MV photon beams. Beam modifiers, such as wedges, were used when necessary to create dose uniformity around the target (see Figure 22). The beam field shaping was accomplished through a multileaf collimator (MLC) to block normal structures from the projection of the target in the beam’s eye-view.

Figure 22.
Representation of typical cone down 4-field arrangement for 3D-CRT
The planning process relied on the included beam’s eye view (BEV) facility to position and design each beam automatically. Blocking was performed automatically using a standard 0.8 cm margin around the PTV. Figure 23 shows the BEV projection of two treatment fields with standard MLC blocking that conforms to the planning target volume (solid volume displayed in red).

Figure 23.
Beam’s Eye View projection of two treatment fields
For some clinical cases, the blocking was manually modified such that critical structures would not receive doses beyond their radiation tolerance. Several studies have identified the tolerance levels of various normal brain tissues to be within the range of 30-55 Gy\(^{40,38}\). Based on these tolerances, the dose to at least one eye was limited to 50 Gy if this could be accomplished without shielding the gross tumor volume. The dose to the optic chiasm was limited to 54 Gy, except for cases where the gross tumor volume would be significantly under-dosed. This was done to reduce any significant risk of visual problems arising from the radiation treatment. In the cases under consideration, the optical structures were blocked at about 52 Gy. Sometimes a portion of the planning target volume was intentionally under-dosed due to this additional blocking.

Three dimensional dose distributions were calculated with 0.45 cm resolution using the Clarkson algorithm and Convolution algorithm (algorithms implemented by CMS treatment planning software) on HP workstations (Hewlett Packard, Palo Alto, CA). It is important to note that the dosimetric effects of tissue in-homogeneities were not taken into account for these plans in any of the brain dose distributions calculated. The treatment plans were created to achieve the dual aims of insuring that the minimum dose was at least 95% of the prescribed dose and that the maximum dose did not exceed 107%. In practice, dose variation within the target volume is usually held to no more than 10% of the prescribed dose.

The inhomogeneity across the treatment volume was thus kept to a minimum. If this goal could not be achieved, then other beam arrangements were proposed to ensure that the over- and under-dosage volumes were minimized. Isodose surfaces were
computed and evaluated for each plan to ensure adequate coverage of the target volume and compliance with the treatment protocol, as shown in Figure 24. Dose volume histograms (DVH), as described in section 7 below, were calculated and evaluated to ensure proper dose coverage and appropriate dose limits on critical organs.

Figure 24. 
Isodose color wash on transversal, coronal and sagittal CT images

Patient shown is #11 with original boost treatment plan. Treatment GTV and PTV volumes are shown in white, inner and outer contour, respectively. Boost prescription dose was 200 cGy for 7 fractions for a total of 14 Gy. Total treatment dose delivered was 60 Gy. Dose color scale is shown on top right corner, from 20% (36cGy, blue) to 103% (205 cGy, red). Notice that 95% of the prescription coverage (red line) covers the GTV target volume and due to blocking of optical structures slightly misses bottom portion of PTV volume. Notice that global dose maximum is 208 cGy or 104% of the prescription dose (value shown top right).
IV.5. Treatment plans for volumes of physicians and segmentation methods

The efficacy of automated segmentation methods in radiation therapy can be evaluated by comparing hypothetical radiation therapy treatment plans based on computer segmentations with those generated by radiation oncologists. For this purpose, the CT images that were used for the planning of the actual treatment were also used to independently generate new, hypothetical treatment plans. However, only the cone down treatment plan was evaluated since the treatment volume generated for this plan is based solely on the gross tumor volume. This means that the computer segmentation methods cannot be used to define the initial brain treatment unless there is a method to properly identify edema to be included in the treatment volume. Nevertheless, the use of the GTV volume is appropriate since the GTV receives the greatest radiation dose and is the most important volume in radiation therapy planning. As mentioned before, over 89% of tumor recurrences have been found to be located within the GTV \(^{(5,6,8)}\).

Since the intra-operator variability has already been evaluated, the physician treatment plans used for comparison were developed based upon the average GTV created by each of the radiation oncologists. A software program to generate these average volumes was written in Interactive Data Language software version 5.4 and can be found in Appendix C. The use of one GTV volume per physician resulted in one treatment plan for each radiation oncologist. Similarly, a radiation therapy treatment plan was created to correspond to each of the computer segmentation generated tumor volumes. The computer segmentation volumes were transferred to the CT image coordinate system used for treatment planning by means of the CT-MRI image
registration as described in Chapter III. A software program was written to merge the new computer segmentation gross tumor volumes into the treatment planning files as described before (Figure 6).

It is important to note that one of the findings from chapter III shows that the automated segmentation methods consistently under-segment significantly in the superior and inferior borders of the tumor volume. Using this insight, some novel modifications have been incorporated in the segmentation volumes to be applied in radiation therapy. The understanding of the radiation oncologists contouring described in the previous chapter were translated to the segmentation methods to improve the resulting computer generated PTV volumes. A margin was added in the vertical direction to the gross tumor volume obtained by the kNN and KG methods. Copying the most superior and inferior contour produced by the segmentation method to the immediately adjacent CT slice simulated the process used by radiation oncologists to identify the most superior and inferior GTV contours, as described in the previous chapter. This procedure added a 4 mm margin in each vertical direction. A visual comparison of the resulting GTV as compared to the physician GTV provided visual evidence of the appropriateness of this expansion technique. This modification could automatically be applied to the results of the current segmentation methods.

As previously noted, the definition of the PTV is crucial to effective plan evaluation and comparison. For a fair comparison, the PTV for both physicians and the automated segmentation methods were determined using the 3D automatic margin expansion function of the treatment planning system with automatic exclusion of the
skull bone as defined in the previous section. These contours are created such that each point is located no closer than the specified distance from the initial contours in three-dimensional space (see Figure 25).

Figure 25.
Representation of a 3D automatic PTV margin expansion

Figure shows a 3D automatic PTV (outer outline shown in green) margin expansion in a MRI axial view, CT axial view, CT reconstructed coronal view, and CT reconstructed sagittal view for the GTV (inner outline shown in red). Patient 8 is shown with a 2 cm automatic expansion around Physician 1 average GTV.
All commercial 3D treatment planning systems perform the margin expansion function automatically. For the Focus system a version of the 3D Rolling Ball algorithm is used. Various researchers\(^{(43,44)}\) have previously defined the principles of the algorithm.

The margin of the PTV remained the same (1-2 cm) as used for the actual patient treatment (see Table 6). No additional trimming around critical structures or manual modifications was performed to eliminate human bias when comparing treatment plans. The critical structure definitions remained the same as those used in the preparation of the actual patient treatment plans. This eliminated the introduction of additional variations in size and location of critical structure during plan evaluation.

Based on the PTVs obtained, five new treatment plans were created for each patient, one for each of the contours by the radiation oncologists and the computer segmentations. The FOCUS treatment planning system with upgraded software was used to create these new patient treatment plans (Versions 4.0.2-4.0.3). These plans had the same basic treatment techniques as used in actual patient treatments described in the previous section. The number of beams, the dose prescription and fractionation, and the positioning of the patient all remained unchanged. All calculation parameters (resolution, inhomogeneity, dose calculation algorithm, etc.) were also kept the same as those of the actual patient treatments discussed in the previous section.

The multi-leaf collimator for each beam used to block normal tissue was created using the automatic MLC margin function. There was slight variation in beam apertures and MLC definition based on the new PTV volumes. The margin remained consistent
with what was used for the actual patient treatment (0.8 cm around the PTV). No additional blocking to the visual structures was added in the interest of avoiding additional bias in the comparison. This ensured that the actual and hypothetical treatment plans were consistent and that the resulting data could be meaningfully compared.

Isodose surfaces were computed for these plans to ensure adequate coverage of the target volume while complying with the applicable brain treatment protocol. As with the standard brain treatment protocol of the facility, coverage of the PTV volumes with the 95% isodose line was considered acceptable. The resulting hypothetical treatment plans were then transferred to the laptop computer for further evaluation and subsequent comparison to the actual plans (as modified to make comparison more meaningful).
IV.6. Expert definition

For the purpose of data analysis in this chapter, Physician 1 was selected as the “baseline expert” in delineating brain gross tumor volume. This decision was based on the following factors: (1) he achieved the smallest intra-operator variability (See Chapter II); (2) he is a radiation oncologist specializing in neuro-oncology with over ten years of clinical practice experience treating over 200 glioma brain tumor patients; and (3) he has participated in several multi-center brain study trials resulting in publications and treatment plans which involved peer review\(^{(4,33)}\). Additionally, Physician 1 was responsible for defining and approving the actual radiation treatment for all eleven patients evaluated in this study.

Based on the above definition of the study’s baseline expert radiation oncologist, the average planning target volume generated by Physician 1 (PTV\(_1\)) was used as the reference volume against which the treatment plans generated from the tumor volumes defined by physicians two and three, as well as the automated MRI segmentation methods, were compared.
IV.7. Treatment plan evaluation techniques in current clinical practice

Many methods have been used to quantitatively evaluate treatment plans by reducing the vast amount of 3D dose data to a single number or a few numbers. Some of the quantitative methods most commonly used are dose statistics, dose-volume histograms (DVHs), and conformity indices.

IV.7.a. Dose statistics

Dose statistics are single numbers that characterize some aspect of the 3D dose distribution. A dose estimate can be determined for a particular volume of interest (VOI), for the treatment volume (PTV), or for the entire patient. While dose statistics are simplistic, they remain an important tool in comparing dose coverage for a region of interest between different treatment plans. The most common dose statistic values are minimum dose ($D_{\text{min}}$), maximum dose ($D_{\text{max}}$), and average dose ($D_{\text{ave}}$). The first two quantities are defined as expected. The average dose is based on the calculation of the dose at each one of a large number of discrete points (grid points), uniformly distributed in the volume of interest. The average dose can be expressed by the equation:

$$D_{\text{ave}} = \frac{1}{N} \sum_{i,j,k} D_{i,j,k}$$ \hspace{1cm} [Equation 1]

where $N$ is the number of grid points, $i,j,k$ are the column, row and level index, respectively, and $D_{i,j,k}$ is the dose at the grid point $i,j,k$ located inside the volume of interest $V^{(41)}$.

IV.7.b. Cumulative dose volume histograms

Another accepted way to summarize the 3D dose distribution to a manageable size while retaining important information about dose variations is the cumulative dose
volume histogram (DVH). DVHs provide a way of presenting the user with a synopsis of the dose to a given structure. They are currently and prominent features of modern treatment-planning systems and were first introduced in order to compare rival treatment plans. A cumulative DVH (in practice referred as DVH) is a graph of volume plotted against dose\(^{45}\). The interpretation of the graph is that the ordinate of a point on the graph represents the volume which receives at least the dose associated with the point’s abscissa i.e., receives that dose or more (Figure 26). By definition, then, the value at the dose origin will be the full volume of the VOI because the entire volume receives at least zero dose. Even though they are called histograms, it is important to note that cumulative dose volume graphs represent a continuous distribution of dose.

A DVH gives the percent of a structure that receives at least a certain dose as a function of dose. This display can quickly demonstrate the proportion of a tissue that is under- or over- dosed with respect to a prescribed dose. It can also demonstrates whether dose differences between two plans are at a critical or non-critical dose. For illustration, Figure 27 schematically shows different DVHs for three possible target dose distributions.
Figure 26.
Schematic representation indicating generation of a cumulative DVH

The ordinate of a point on the graph represents the volume which receives at least the dose associated with the point’s abscissa. That is, receives that dose or more. By definition, then, the value at the dose origin will be the full volume of the volume of interest because the entire volume receives at least zero dose.

(From Goitein 1992\(^{(45)}\), reproduction permitted by Elsevier)
IV.7.c. Conformity index

Although DVHs provide very accurate information about the dosimetric properties of treatment plans, it may be difficult to infer the clinical significance of differences in rival treatment plans using DVHs alone. Conformity indices were introduced as one of the guidelines for radiosurgery treatment planning by the Radiation Therapy Oncology Group (RTOG) to evaluate the dosimetric properties of a treatment plan \(^{(46)}\). Conformity indices have been used extensively as a tool to define the conformation of a 3D treatment plan, that is, how close the prescription isodose covers the defined target volume PTV. It has also been used extensively to compare different treatment plans or treatment planning modalities \(^{(47-49)}\).

The *conformity index* (CI) of a treatment plan is defined as the ratio of the prescription isodose volume \(V_{PX}\) to the volume of the target \(V_{PTV}\):

\[
CI = \frac{V_{PX}}{V_{PTV}} \tag{Equation 2}
\]
where $V_{PX}$ is the isodose volume within the percent isodose surface that encompasses the target, usually the prescription isodose. CI evaluates the dose distribution with emphasis on the size of the target volume compared to the size of the prescribed treatment volume.

The conformity index is unity for a perfect hypothetical fully conformal treatment, i.e. the prescription isodose surface coincides exactly with the boundaries of the PTV. An increase in the value shows that the isodose surface is larger than the PTV, which results in undesirable radiation to normal tissues. A decrease in value to less than one represents under-treatment of the PTV. According to the Radiation Therapy Oncology Group (RTOG), a value between 1.0 and 2.0 is assumed to be within the treatment protocol for brain radiosurgery. A value less than 1.0 but greater than 0.9, or between 2.0 and 2.5, are classified as minor deviations. All other values are classified as major deviations (see Figure 28).

![Figure 28. Scale representing range of conformity index values](image)

Note that values within 1.0-2.0 are considered acceptable. Values within 0.9-1.0 or 2.0-2.5 are considered minor deviations. Any value beyond this range is considered an unacceptable conformity of the treatment plan. 

92
IV.8. Treatment plan analysis methods for this study

This study compares the two treatment plans created from the segmentation methods with the three treatment plans created from the physicians’ outlines based on the dose statistics, the dose volume histograms and the conformity indices. As described in Section IV.6, the average volume determined by Physician 1 is used as the reference. Therefore, all comparisons were based on exploring the coverage of the plans created on the volumes of the physicians and segmentation methods as compared to the coverage given to the volume created by Physician 1 (PTV<sub>1</sub>). All plan comparisons were performed qualitatively, by visual inspection of the dose distribution; and quantitatively, by the use of various dose statistics, dose volume histograms, and treatment plan conformity indices.

IV.8.a. Dose statistics

All dose statistics for this analysis were generated using the Focal Vue treatment planning evaluation software (Computerized Medical Systems, St. Louis, MO) running on the Dell Inspiron 7000 laptop computer (Dell, Round Rock, TX). The basic dose statistic, i.e., average dose, minimum dose, and maximum dose were determined for each treatment plan created with reference to the PTV of Physician 1 using a 2 mm calculation grid. Since the minimum dose could be misleading in interpreting dose uniformity in the target volume, an additional calculation of the percent volume that receives less than the prescription isodose (95%) was also determined.
IV.8.b. Cumulative dose volume histogram

An analysis of the cumulative dose volume histogram for each of the plans was performed in a manner similar to the comparison proposed by Verhey and co-workers \(^{(49)}\). The amount of normal brain tissue that would be treated with the different physician volumes and computer segmentation was calculated for all treatment plans and compared with the standard reference (Physician 1). DVHs were obtained using a calculation grid of 2 mm.

IV.8.c. Conformity index

The conformity index was used to compare the treatment plans in a manner similar to that set forth in previously published research \(^{(48,49)}\). The degree of conformity of each individual treatment plan was evaluated by the radiation conformity index (CI) defined as:

\[
CI_i = \frac{(V_{95})_i}{(V_{PTV})_i} \quad \text{[Equation 3]}
\]

where \(V_{95}\) represents the volume of the 95% isodose (prescription isodose) and \(V_{PTV}\) is the planning target volume for the treatment plan being evaluated, as defined in section 3 of this chapter. The volume of the 95% isodose coverage was derived from the absolute dose volume histogram information produced by the treatment planning evaluation software. The conformity index of each of the treatment plans was calculated and denoted as CI\(_1\), CI\(_2\), and CI\(_3\) for the three physicians, CI\(_{knn}\) for the plan generated by the kNN segmentation method contours, and CI\(_{kg}\) for the plan generated by the knowledge guided system contours.
The conformity index of Eq. 3 is used in clinical practice to compare plans generated from the same planning volume (PTV). However, in this study it was of particular interest to compare plans developed from different PTVs. In particular, it was intended to evaluate if the plans based on computer segmentations are acceptably conforming to the expert physician. To perform this evaluation a novel system of cross-conformity indices was developed based on each treatment plan but using the Physician 1 target volume (PTV\textsubscript{1}) as the standard.

Figure 29.
Evaluation strategy based on conformity indices

PTV\textsubscript{1} represents the average PTV of the expert physician; PTV\textsubscript{2} and PTV\textsubscript{3} represent the average PTV of physicians 2 and 3; PTV\textsubscript{kNN} and PTV\textsubscript{KG} represent the PTV of resulting from the computer segmentation methods. The cross-conformity indices (CCI\textsubscript{i-1}) are calculated by taking the dose distribution from each plan based on the outline of the expert physician (PTV\textsubscript{1}).
The treatment plan evaluation strategy is schematically outlined in Figure 25 above. For example, CCI_{21} is the conformity index obtained using the volume dose distribution of the *plan* based on the outlines of physician 2, but comparing the resulting isodose volume to the *planning target volume* of physician 1. It is important to note that the original description of CI compares the size of the prescription treatment volume to the size of the PTV since it assumes that the plan is evaluated and approved. This process assumes that the prescription isodose covers the actual target volume (Figure 30(a)). For the case of cross-conformity indices, the evaluated plan coverage is based on an entirely independent target volume (PTV_1) from the original plan and does not consider spatial location of the isodose volume evaluated since no plan approval was performed. Figure 30(b) shows an example of a conformity index of 1 but with an isodose volume that would miss much of the target.

(a) ![Conformity Index](image1.png)  
(b) ![Conformity Index](image2.png)

Figure 30.
Target volume isodose spatial coverage

Figure represents PTV (solid volume in blue) and reference isodose volume (represented by outline in red). Case (a) represents the ideal case and (b) the case of isodose spatial misplacement with respect to the PTV. Note how both cases would yield a conformity index of about 1 since the isodose volume is equal to the actual PTV.
Therefore, the definition of conformity index was modified similar to Van’t Riet et al. \(^{50}\). A correction factor is introduced that accounts for coverage of the target:

\[
CF = \frac{V_{\text{PTV}}}{(V_{\text{PTV}})_{\geq px}} \quad \text{[Equation 4]}
\]

where \(V_{\text{PTV}}\) is the planning target volume and \((V_{\text{PTV}})_{\geq px}\) is the volume of target receiving a dose equal to or greater than the reference or prescription dose. The correction factor increases as the proportion of PTV covered by at least the prescription isodose decreases. This is based on the original definition of conformity index that evaluates coverage based on the minimum prescription dose that encompasses the target \(^{47,50}\). Therefore, the conformity index is modified to include specifically the target volume enclosed by the prescription isodose selected:

\[
CI = \frac{V_{px}}{(V_{\text{PTV}})_{\geq px}} \quad \text{[Equation 5]}
\]

Using the spatial co-location correction factor (Eq. 4) and above definition of conformity index (Eq. 5) results in a cross-conformity index definition for physician or segmentation plan, \(j\), with respect to Physician 1 and 95% prescription isodose as:

\[
CCI_{j-1} = \frac{(V_{\text{PTV}})_1}{(V_{\text{PTV}1})_{95}} \cdot \frac{(V_{95})_j}{(V_{\text{PTV1}})_{\geq 95}} \quad \text{[Equation 6]}
\]

The average and range of cross-conformity indices of cross-operator plans (CCI\(_{2-1}\) and CCI\(_{3-1}\)) was calculated and compared to the average and range of cross-conformity indices of the computer segmentation plans (CCI\(_{knn-1}\) and CCI\(_{kg-1}\), respectively).
IV.9. Post-treatment evaluation methods

The validation technique described above compared an expert treatment plan with other hypothetical (research) treatment plans. This section introduces a novel approach to evaluate treatment plans based on the outcome of the actual clinical radiation treatment as seen on post-treatment images. An analysis was performed to identify radiation damage in tumor and normal tissue, as well as treatment failure in active tumor. These results were used to compare what would have been obtained by using the hypothetical treatment plans.

The clinical facility’s routine follow-up procedures after completion of a radiation treatment course include the administration of an MRI. The imaging protocol for the post-treatment MRI is similar to the pre-treatment MRI described above (see Chapter II). The primary study set used for this evaluation consisted of $T_1$ scans obtained after administration of 0.1 mmol/kg body weight of gadolinium (Gd) MRI contrast material (Gd-DTPA) and using a standard spin-echo (SE) sequence with a repetition time (TR)/echo time (TE) = 400/8 or TR/TE=525/17 ms. For cases of non-tumor enhancement, $T_2$ axial images were used as the primary study set. $T_2$ images were obtained with a TR/TE=3000/104 ms or TR/TE=4000/96 ms. Additional sagittal and coronal images were available to help identify structures of interest in the axial scans. Only eight of the eleven patients participating in this study returned to the University of South Florida to have their post-treatment MRI performed. Therefore, the results of these eight patients with follow-up MRI studies were used to develop and analyze the proposed treatment evaluation tool (see Table 7).
Table 7.  
Patient post-treatment imaging studies

<table>
<thead>
<tr>
<th>Case</th>
<th>Post-MR available</th>
<th>Tumor Enhanced</th>
<th>Post-treatment MR (days from end of radiation treatment)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yes</td>
<td>No</td>
<td>502</td>
</tr>
<tr>
<td>2</td>
<td>Yes</td>
<td>No</td>
<td>378</td>
</tr>
<tr>
<td>3</td>
<td>Yes</td>
<td>Yes</td>
<td>146</td>
</tr>
<tr>
<td>4</td>
<td>No</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>5</td>
<td>No</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>6</td>
<td>Yes</td>
<td>Yes</td>
<td>363</td>
</tr>
<tr>
<td>7</td>
<td>No</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>8</td>
<td>Yes</td>
<td>Yes</td>
<td>112</td>
</tr>
<tr>
<td>9</td>
<td>Yes</td>
<td>Yes</td>
<td>64</td>
</tr>
<tr>
<td>10</td>
<td>Yes</td>
<td>Yes</td>
<td>437</td>
</tr>
<tr>
<td>11</td>
<td>Yes</td>
<td>Yes</td>
<td>48</td>
</tr>
</tbody>
</table>

The evaluation procedure is presented graphically in Figure 31 below. Additionally, Figure 1 from Chapter I shows a flowchart of the complete design of the project, including the evaluation mechanism to determine the accuracy of treatment planning based on follow-up imaging.
An expert neuro-radiologist identified three different tissue types on the post-treatment MRI: residual or recurrent tumor tissue (tumor), radiation damage in the tumor (necrosis), and radiation damage in normal tissue (normal tissue damage). These three
areas were contoured by the neuro-radiologist on the axial primary study set (post-contrast T\textsubscript{1} or T\textsubscript{2}) axial MRI scans in its original MRI imaging position using the Focal Ease contouring software. The neuro-radiologist utilized digital displays of all the other MRI post-treatment scan types performed on the patient, which include coronal and sagittal scans, to aid in identifying these three different tissue types. Because this precise contouring is very time consuming, it was performed in three slices through the central area of the tumor volume for each of the 8 patients with post-treatment MRI.

The following is a note regarding the limitations of the post-treatment MRI available in this study. New and improved imaging modalities are constantly being adopted to improve tumor volume and tissue classification protocols. Recent studies have shown that several different imaging modalities must be used to maximize accuracy in the differentiation of real tumor tissue from regular brain tissue or necrosis\textsuperscript{(2,51)}. When MRI is used to define recurrence, it is often difficult to be certain whether the enhancing abnormality represents true tumor recurrence or radiation-induced necrosis\textsuperscript{(5)}. Until newer imaging techniques become routinely available, MRI scans, despite their limitation, offer the best assessment of tumor progression for the purposes of tumor follow-up. Therefore, the application of the proposed evaluation technique to the post-treatment image set may not yield the most accurate results. However, this method is used here as a tool to analyze the relative efficacy of the radiation treatment. The proposed evaluation technique could be applied to any imaging modality, resulting in more accurate evaluation results and possible wider medical applications.
The primary follow-up MRI study set was registered to the CT coordinate system used for treatment planning with the Focal Fusion registration software. The expert radiation oncologist verified the registration results in a similar manner as described in Chapter II. The resulting transformation matrix was applied to transform the post-MRI contours into the pre-treatment CT image coordinate system. Software programs were created using IDL to perform these contour transformations and to incorporate the new contours into the original patient contour dataset as described in Chapter III. This procedure permitted the comparison of the original patient treatment plan with the hypothetical plans; the actual patient results were then used to evaluate the treatment results in a novel manner.

Based on the tissue classification obtained from the post-treatment follow-up MRI, dose volume histogram distributions were calculated for the actual treatment plan and for the hypothetical treatment plans created from the tumor volumes created by the physicians and the automated computer segmentation. A calculation grid of 1 mm was used for structures less than 25.0 cm$^3$ to increase accuracy. Cumulative dose volume histograms were specifically calculated for the following regions as delineated by the neuro-radiologist: (a) tumor, (b) necrosis, and (c) damaged normal tissue. These different areas represent tissue that received radiation from the actual patient treatment. Additionally, graphical evaluation of isodose coverage was compared for the areas of interest irradiated under the different treatment plans.
IV.10. Results

IV.10.a. Segmentation GTV volume expansion

As described in section 5, the resulting kNN and KG volumes were modified with the addition of a margin in the vertical direction of the GTV.

(a) Original GTV

(b) Extended GTV

Figure 32.
GTV vertical expansion
Figure 32 shows the result of implementing a vertical expansion for the GTV of patient 9. The original and expanded GTV resulting from kNN (inner contour displayed in magenta) and KG segmentation method (inner contour displayed in cyan) are displayed as compared to the physician 1 average GTV (outer contour displayed in orange). A visual comparison of the resulting expanded GTV as compared to the physician 1 for this example in coronal views shows the close approximation of the volume in the vertical direction. Additionally, a 3D representation of the reconstructed volumes shows a closer agreement with physician 1.

Figure 33.
Example of original GTV for kNN and KG compared to physician 1
Figures 33 and 34 show another example of the results of expanding the GTV for kNN (inner contour displayed in magenta) and KG segmentation method (inner contour displayed in cyan) for patient 11. A visual comparison of the resulting GTV as compared to the physician 1 average GTV (outer contour displayed in orange) for this example shows the close approximation of the volume. Similar results were obtained for the other ten patients. No additional analysis was performed since the GTV modification accounts for less than 5% of the total volume size.
IV.10.b. Planning target volumes

The PTV for both physician and segmentation methods was determined using the automatic margin expansion technique as previously defined. Figure 35 shows the PTV resulting from the modified kNN segmentation for patient 11 as compared to the PTV resulting from the average physician 1 GTV. Notice the close alignment and match of the volume resulting after the expansion compared to original GTV (Figure 33).

Similarly, Figure 36 shows an example of the results for the KG volume expansion.

Figure 35.
Example of kNN segmentation PTV compared with physician 1 PTV

Figure shows PTV resulting from kNN segmentation method (inner contour displayed in red) compared to PTV of Physician 1 (outer contour displayed in yellow) for patient 11. Images displayed show axial, coronal and sagittal MR views plus a 3D representation of the reconstructed volumes.
Figure 36. Example of KG segmentation PTV compared with physician 1 PTV

Figure shows PTV resulting from KG segmentation method (inner contour displayed in red) and compared to PTV of Physician 1 (outer contour displayed in yellow) for patient 11. Images displayed show axial, coronal and sagittal MR views plus a 3D representation of the reconstructed volumes.

Table 8 shows the size of the resulting PTVs in cm$^3$. Note similar values obtained for the volumes from the kNN and KG segmentation as compared with physician 1.

Paired t-test with Bonferroni correction for multiple tests showed that PTV$_2$ and PTV$_{\text{kNN}}$ are significantly different from the PTV$_1$. 
Table 8.
Values of resulting PTVs for physicians and segmentations methods

<table>
<thead>
<tr>
<th>Patient #</th>
<th>PTV&lt;sub&gt;1&lt;/sub&gt; (cm&lt;sup&gt;3&lt;/sup&gt;)</th>
<th>PTV&lt;sub&gt;2&lt;/sub&gt; (cm&lt;sup&gt;3&lt;/sup&gt;)</th>
<th>PTV&lt;sub&gt;3&lt;/sub&gt; (cm&lt;sup&gt;3&lt;/sup&gt;)</th>
<th>PTV&lt;sub&gt;kNN&lt;/sub&gt; (cm&lt;sup&gt;3&lt;/sup&gt;)</th>
<th>PTV&lt;sub&gt;KG&lt;/sub&gt; (cm&lt;sup&gt;3&lt;/sup&gt;)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>130</td>
<td>134</td>
<td>131</td>
<td>113</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>193</td>
<td>187</td>
<td>200</td>
<td>167</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>248</td>
<td>263</td>
<td>247</td>
<td>208</td>
<td>N/A</td>
</tr>
<tr>
<td>4</td>
<td>220</td>
<td>236</td>
<td>227</td>
<td>196</td>
<td>197</td>
</tr>
<tr>
<td>5</td>
<td>468</td>
<td>500</td>
<td>492</td>
<td>422</td>
<td>409</td>
</tr>
<tr>
<td>6</td>
<td>369</td>
<td>393</td>
<td>390</td>
<td>390</td>
<td>N/A</td>
</tr>
<tr>
<td>7</td>
<td>231</td>
<td>251</td>
<td>222</td>
<td>182</td>
<td>205</td>
</tr>
<tr>
<td>8</td>
<td>305</td>
<td>331</td>
<td>317</td>
<td>284</td>
<td>308</td>
</tr>
<tr>
<td>9</td>
<td>131</td>
<td>155</td>
<td>134</td>
<td>120</td>
<td>107</td>
</tr>
<tr>
<td>10</td>
<td>150</td>
<td>166</td>
<td>166</td>
<td>139</td>
<td>157</td>
</tr>
<tr>
<td>11</td>
<td>164</td>
<td>162</td>
<td>156</td>
<td>125</td>
<td>126</td>
</tr>
</tbody>
</table>

Average 237 ± 107 253 ± 114 244 ± 114 213 ± 107 N/A

IV.10.c. Plan comparison

The resulting treatment plans were compared quantitatively by the use of dose statistics, DVHs and CIs as described in section 9 of this chapter. Plans were created based on their respective planning target volumes as described in section 5. All plans were evaluated by comparing the coverage of the planning target volume of reference Physician 1 (PTV<sub>1</sub>).
IV.10.d. Dose statistics

Table 9 shows the average dose statistics for the different treatment plans as percent of boost treatment dose (see Table 6) and part of the planning volume not receiving at least 95% of the boost dose. All values are displayed in percent volume of PTV. Note that the mean average dose is larger than the prescribed dose (95% of boost dose) for all treatment plans indicating good target coverage. Paired t-tests with Bonferroni correction indicate that the kNN treatment plans have a significantly lower mean dose than plans of Physician 1.

Table 9.
Average dosimetric values for PTV$_1$ from the different treatment plans

<table>
<thead>
<tr>
<th>Dosimetric Quantity</th>
<th>Physician 1 (%)</th>
<th>Physician 2 (%)</th>
<th>Physician 3 (%)</th>
<th>KNN (%)</th>
<th>KG (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average dose</td>
<td>99.8 ± 2.1</td>
<td>99.8 ± 1.1</td>
<td>99.8 ± 1.2</td>
<td>99.5 ± 1.4</td>
<td>99.3 ± 0.6</td>
</tr>
<tr>
<td>Maximum dose</td>
<td>103.9 ± 2.1</td>
<td>104.2 ± 2.2</td>
<td>104.0 ± 2.1</td>
<td>104.2 ± 2.1</td>
<td>104.2 ± 2.1</td>
</tr>
<tr>
<td>Minimum dose</td>
<td>91.4 ± 3.1</td>
<td>87.2 ± 10</td>
<td>90.6 ± 3.3</td>
<td>75.0 ± 16.6*</td>
<td>79.3 ± 11.9</td>
</tr>
<tr>
<td>Volume below 95% isodose</td>
<td>1.0 ± 1.4</td>
<td>1.3 ± 1.9</td>
<td>1.1 ± 1.3</td>
<td>3.5 ± 2.4*</td>
<td>4.5 ± 3.4</td>
</tr>
</tbody>
</table>

(*) statistically significant as compared with Physician 1
The dose maximum and minimum values were used to evaluate the homogeneity of the dose distributions from the automated segmentation plans and the physician plans. Note similar results for all plans regarding maximum dose in Table 9. Results from minimum dose seem to indicate that kNN and KG lead to less homogeneous plans. Since the minimum dose could be misleading in interpreting dose uniformity in the target volume, an additional calculation of the percent volume that receives less than the prescription dose (95%) was determined. It can be seen that even though the dose is not as homogeneous when using the plan obtained from the KG and kNN volumes, nevertheless, the amount of underdose volume is insignificant for clinical purposes (<5% on average). In practice dose variation within the target is usually held to no more than 10% of the prescribed dose.

IV.10.e. Cumulative dose volume histogram

A comparison of the amount of normal brain tissue that would be treated with the different physician volumes and computer segmentation was calculated and compared with the plan from Physician 1. Table 10 shows the average dose to normal brain resulting from boost plans as percentage of total prescription dose (95% of total boost dose (Table 6)). Normal brain was calculated by subtracting the PTV$_1$ volume from the total brain volume (Normal brain volume = Total brain volume – PTV$_1$). Note the slightly less dosage of normal brain by using the plans from the kNN and KG segmentation. Paired t-tests showed that there is no statistical difference in normal tissue dose between the plans from the different physician or segmentation GTVs as compared with physician 1.

110
Table 10.
Average dose to normal brain irradiation for boost plans

<table>
<thead>
<tr>
<th>Patient #</th>
<th>Physician1 (%)</th>
<th>Physician2 (%)</th>
<th>Physician3 (%)</th>
<th>kNN (%)</th>
<th>KG (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>29.4</td>
<td>28.9</td>
<td>28.9</td>
<td>26.7</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>32.2</td>
<td>30.0</td>
<td>32.2</td>
<td>28.9</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>46.0</td>
<td>48.0</td>
<td>46.0</td>
<td>43.0</td>
<td>N/A</td>
</tr>
<tr>
<td>4</td>
<td>36.0</td>
<td>39.0</td>
<td>41.5</td>
<td>39.5</td>
<td>36.5</td>
</tr>
<tr>
<td>5</td>
<td>60.0</td>
<td>65.0</td>
<td>67.0</td>
<td>65.5</td>
<td>55.5</td>
</tr>
<tr>
<td>6</td>
<td>43.0</td>
<td>41.5</td>
<td>44.5</td>
<td>43.0</td>
<td>N/A</td>
</tr>
<tr>
<td>7</td>
<td>40.6</td>
<td>42.8</td>
<td>39.4</td>
<td>35.6</td>
<td>38.9</td>
</tr>
<tr>
<td>8</td>
<td>45.0</td>
<td>47.5</td>
<td>46.5</td>
<td>43.5</td>
<td>47.5</td>
</tr>
<tr>
<td>9</td>
<td>34.5</td>
<td>39.0</td>
<td>35.0</td>
<td>33.0</td>
<td>27.0</td>
</tr>
<tr>
<td>10</td>
<td>36.7</td>
<td>39.4</td>
<td>39.4</td>
<td>36.1</td>
<td>37.8</td>
</tr>
<tr>
<td>11</td>
<td>41.0</td>
<td>41.0</td>
<td>39.5</td>
<td>35.5</td>
<td>35.5</td>
</tr>
</tbody>
</table>

Average 40.4±8.4  42.0±9.7  41.8±10.0  39.1±10.4  39.8±9.2

Figure 37 shows a graphical comparison of the DVHs for patient 11. Notice that the KG and kNN treatment plans give slightly less dosage to normal brain tissue and maintain proper coverage of PTV\textsubscript{1}; i.e. dose $\geq 95\%$. Similar results were encountered for the other patients.
Figure 37.
Comparisons of DVHs for patient 11 for PTV₁ and normal brain tissue

Comparisons of DVHs for patient 11 showing PTV₁ target volume (superior set of lines) and normal brain tissue (inferior set of lines) for (a) Physician 1 (solid line), Physician 2 (dashed line) and Physician 3 (doted line) treatment plans; and (b) Physician 1 (solid line), KG (dashed line), kNN (dotted line).
IV.10.f. Conformity index

The conformity index of each of the treatment plans based on their corresponding PTVs was calculated and denoted as $\text{CI}_1$, $\text{CI}_2$, and $\text{CI}_3$ for the three physicians, $\text{CI}_{\text{kNN}}$ for the plan generated by the kNN segmentation method contours, and $\text{CI}_{\text{KG}}$ for the plan generated by the knowledge guided system contours. This provided information on the degree of difficulty to design a plan that closely conforms to each outline. Table 11 shows the conformity indices for all treatment plans.

<table>
<thead>
<tr>
<th>Patient #</th>
<th>Physician 1 $\text{CI}_1$</th>
<th>Physician 2 $\text{CI}_2$</th>
<th>Physician 3 $\text{CI}_3$</th>
<th>kNN $\text{CI}_{\text{kNN}}$</th>
<th>KG $\text{CI}_{\text{KG}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.34</td>
<td>1.28</td>
<td>1.31</td>
<td>1.32</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>1.52</td>
<td>1.51</td>
<td>1.50</td>
<td>1.54</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>1.41</td>
<td>1.43</td>
<td>1.40</td>
<td>1.45</td>
<td>N/A</td>
</tr>
<tr>
<td>4</td>
<td>1.35</td>
<td>1.37</td>
<td>1.36</td>
<td>1.38</td>
<td>1.41</td>
</tr>
<tr>
<td>5</td>
<td>1.57</td>
<td>1.53</td>
<td>1.53</td>
<td>1.62</td>
<td>1.59</td>
</tr>
<tr>
<td>6</td>
<td>1.41</td>
<td>1.34</td>
<td>1.40</td>
<td>1.35</td>
<td>N/A</td>
</tr>
<tr>
<td>7</td>
<td>1.48</td>
<td>1.45</td>
<td>1.48</td>
<td>1.56</td>
<td>1.51</td>
</tr>
<tr>
<td>8</td>
<td>1.42</td>
<td>1.36</td>
<td>1.42</td>
<td>1.46</td>
<td>1.47</td>
</tr>
<tr>
<td>9</td>
<td>2.08</td>
<td>1.97</td>
<td>2.07</td>
<td>2.10</td>
<td>2.28</td>
</tr>
<tr>
<td>10</td>
<td>1.35</td>
<td>1.36</td>
<td>1.36</td>
<td>1.40</td>
<td>1.38</td>
</tr>
<tr>
<td>11</td>
<td>1.50</td>
<td>1.50</td>
<td>1.50</td>
<td>1.59</td>
<td>1.53</td>
</tr>
</tbody>
</table>

Average $1.49 \pm 0.21$  $1.46 \pm 0.19$  $1.48 \pm 0.21$  $1.52 \pm 0.22$  N/A
Note that all values displayed in Table 11, except for patient 9, are within the acceptable range as defined in Figure 28. Patient 9 resulted in conformity index value considered a minor deviation but acceptable for treatment planning. A comparison with the data from table 6 shows that a better conformity index, that is, values closer to one, results for treatment plans with a larger number of fields. Statistical analysis showed that all plans for each patient were created to the same level of conformity.

Table 12. Values of cross-conformity indices

<table>
<thead>
<tr>
<th>Patient #</th>
<th>Physician 1 CCI&lt;sub&gt;1&lt;/sub&gt;-&lt;sub&gt;1&lt;/sub&gt;</th>
<th>Physician 2 CCI&lt;sub&gt;2&lt;/sub&gt;-&lt;sub&gt;1&lt;/sub&gt;</th>
<th>Physician 3 CCI&lt;sub&gt;3&lt;/sub&gt;-&lt;sub&gt;1&lt;/sub&gt;</th>
<th>kNN CCI&lt;sub&gt;kNN-1&lt;/sub&gt;</th>
<th>KG CCI&lt;sub&gt;KG-1&lt;/sub&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.42</td>
<td>1.43</td>
<td>1.42</td>
<td>1.34</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>1.52</td>
<td>1.49</td>
<td>1.55</td>
<td>1.39</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>1.43</td>
<td>1.53</td>
<td>1.41</td>
<td>1.31</td>
<td>N/A</td>
</tr>
<tr>
<td>4</td>
<td>1.42</td>
<td>1.55</td>
<td>1.48</td>
<td>1.39</td>
<td>1.49</td>
</tr>
<tr>
<td>5</td>
<td>1.57</td>
<td>1.64</td>
<td>1.61</td>
<td>1.49</td>
<td>1.52</td>
</tr>
<tr>
<td>6</td>
<td>1.53</td>
<td>1.60</td>
<td>1.59</td>
<td>1.63</td>
<td>N/A</td>
</tr>
<tr>
<td>7</td>
<td>1.49</td>
<td>1.58</td>
<td>1.45</td>
<td>1.30</td>
<td>1.38</td>
</tr>
<tr>
<td>8</td>
<td>1.43</td>
<td>1.48</td>
<td>1.48</td>
<td>1.39</td>
<td>1.52</td>
</tr>
<tr>
<td>9</td>
<td>2.08</td>
<td>2.34</td>
<td>2.11</td>
<td>1.94</td>
<td>2.30</td>
</tr>
<tr>
<td>10</td>
<td>1.37</td>
<td>1.51</td>
<td>1.51</td>
<td>1.41</td>
<td>1.51</td>
</tr>
<tr>
<td>11</td>
<td>1.51</td>
<td>1.48</td>
<td>1.46</td>
<td>1.32</td>
<td>1.28</td>
</tr>
<tr>
<td>Average</td>
<td>1.52 ± 0.19</td>
<td>1.60 ± 0.25</td>
<td>1.55 ± 0.20</td>
<td>1.45 ± 0.19</td>
<td>N/A</td>
</tr>
</tbody>
</table>
IV.10.g. Cross-conformity indices

Table 12 shows the values for cross-conformity indices (CCIs) as defined by Eq. 6 above. These values were based on the comparison of the 95% isodose coverage for the reference PTV and a correction factor to account for location of the dose volume. Notice how the CCIs with correction for Physician 1 are very close to original CI calculated for physician one (Table 11). This confirms that the correction factor is applied correctly and does not alter the actual definition of CI when the reference isodose volume covers the evaluated target volume. The correction factor only penalizes the CI for cases where the isodose volume does not cover the target volume.

Most cases in Table 12 show equal conformity. Paired t-tests with Bonferroni correction for multiple comparisons on the data in Table 12 showed only a statistical difference between $\text{CCI}_{1,1}$ and $\text{CCI}_{2,1}$.

IV.10.h. Post-treatment evaluation

The post-treatment MR images were used to compare the resulting volumes that were irradiated with those that would be irradiated under the different treatment plans. Figure 38 shows the dose coverage of the different tumor volumes identified in the post-MR for patient 11. Figure 38(b) is an example of the excellent agreement found in location between the area identified as tumor by physician 1 and the area identified as tumor in post-MR by the neuro-radiologist. Similar results were obtained for the other patients confirming the selection of Physician 1 as the expert radiation oncologist for purpose of this study. Figures 38 shows how the plan and PTV resulting from physician
1 and the automated segmentation methods provide coverage to the area identified as tumor and necrosis that represent the actual treated region.

Figure 38.
Post-MRI and isodose coverage for tumor and necrosis for patient 11

Figure (a) shows volume of tumor (outer volume in red) and necrosis (inner volumes in green) identified in axial post-MR image for patient 11. Figures (b)-(d) show the same volumes together with average GTV (next outline) and PTV (outer outline) for physician 1 (a), kNN (b) and KG (d). The isodose colorwash is overlaid in the figures to display doses from 20% (outside starting in shades of blue) to 100% (inner in shades of red) resulting from the plans generated by each PTV.
Figure 39.
Post-MRI and isodose coverage for tumor, necrosis and damaged tissue

Figure (a) shows tumor (two outer contours in red), necrosis (contours inside tumor displayed in green) and damaged normal (tissue surrounding tumors displayed in blue) for patient 9 on post MRI axial image. Figures b, c, d, show the dose distribution resulting from plans based on volume from Physician 1, kNN and KG respectively. Dose colorwash display shows dose distribution from 20% (outside shades in blue) to 100% (inner shades in red).
Figure 39 shows an example of volumes identified in the post-MR and the coverage resulting from the different treatment plans. Notice that even though kNN and KG have a slightly smaller treatment area, there is still coverage of the area that was in fact treated, i.e. tumor, necrosis and damaged normal tissue.

Cumulative dose histograms were calculated for the areas identified as (a) tumor, which may represent tissue identified as GTV in the actual treatment plan; (b) necrosis, which may represent tumor tissue that was targeted by the treatment; and (c) damaged normal tissue, which may be correlated with inaccuracies in the used tumor definition or additional area treated due to the PTV margin expansion. All data is summarized in tables 13, 14 and 15 in percentages of boost prescription dose. Values shown are in percentages of boost prescription dose. The volume (column 2) is only the volume within the three slices contoured by the expert neuroradiologist. Notice that due to non-tumor enhancement of Patients 1 and 2, only necrosis and damaged normal was identified in the post-MR images.

It was found that the coverage of these areas by the original patient treatment plans was on average of $98.2\% \pm 8.0\%$. This shows that all areas identified in the post-MR were treated by the treatment plan. Small under-dosage was the result of additional blocking for patients with PTV close to adjacent critical organs. Analysis of the coverage from the plans resulting from the segmentation methods shows an average of $99.1\% \pm 2.8\%$ and $98\% \pm 6.4\%$ for kNN and KG, respectively. Similar values were found for the physicians, indicating that both the physicians and the segmentation methods would have treated all tissues as identified in the post-MR. Therefore, the use of an alternate plan
resulting from the automated methods would have yielded irradiation areas that are in agreement with the areas identified from the actual patient treatment.

Table 13.
Data evaluation for post-MRI tumor volume

<table>
<thead>
<tr>
<th>Patient</th>
<th>Volume (cm³)</th>
<th>% Dose Phys.1</th>
<th>% Dose Phys.2</th>
<th>% Dose Phys.3</th>
<th>% Dose kNN</th>
<th>% Dose KG</th>
<th>% Dose TXPlan</th>
</tr>
</thead>
<tbody>
<tr>
<td>1*</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>2*</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>3</td>
<td>20.9</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>98.0</td>
<td>N/A</td>
<td>101.0</td>
</tr>
<tr>
<td>6</td>
<td>6.6</td>
<td>99.0</td>
<td>99.0</td>
<td>99.0</td>
<td>98.5</td>
<td>N/A</td>
<td>99.0</td>
</tr>
<tr>
<td>8</td>
<td>16.7</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>103.0</td>
</tr>
<tr>
<td>9</td>
<td>16.1</td>
<td>103.0</td>
<td>102.5</td>
<td>103.0</td>
<td>103.5</td>
<td>100.5</td>
<td>104.0</td>
</tr>
<tr>
<td>10</td>
<td>22.2</td>
<td>95.0</td>
<td>96.1</td>
<td>96.1</td>
<td>96.1</td>
<td>96.1</td>
<td>86.1</td>
</tr>
<tr>
<td>11</td>
<td>11.5</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>103.0</td>
</tr>
<tr>
<td>Ave.</td>
<td>15.7</td>
<td>99.3</td>
<td>99.4</td>
<td>99.4</td>
<td>99.4</td>
<td>N/A</td>
<td>99.4</td>
</tr>
<tr>
<td>± 5.8</td>
<td>± 2.6</td>
<td>± 2.1</td>
<td>± 2.3</td>
<td>± 2.5</td>
<td>± 6.7</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(*) Patients 1 & 2 did not show tumor on post-treatment MRI.
Table 14.
Data evaluation for post-MRI necrosis volume

<table>
<thead>
<tr>
<th>Patient</th>
<th>Volume (cm³)</th>
<th>% Dose Phys.1</th>
<th>% Dose Phys.2</th>
<th>% Dose Phys.3</th>
<th>% Dose kNN</th>
<th>% Dose KG</th>
<th>% Dose TXPlan</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.8</td>
<td>99.4</td>
<td>99.4</td>
<td>99.4</td>
<td>99.4</td>
<td>N/A</td>
<td>101.1</td>
</tr>
<tr>
<td>2</td>
<td>2.0</td>
<td>98.9</td>
<td>99.4</td>
<td>99.4</td>
<td>99.4</td>
<td>N/A</td>
<td>98.9</td>
</tr>
<tr>
<td>3</td>
<td>9.9</td>
<td>98.0</td>
<td>98.0</td>
<td>98.0</td>
<td>98.0</td>
<td>N/A</td>
<td>101.0</td>
</tr>
<tr>
<td>6</td>
<td>1.3</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>98.5</td>
<td>N/A</td>
<td>98.5</td>
</tr>
<tr>
<td>8</td>
<td>11.5</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>99.5</td>
<td>103.0</td>
</tr>
<tr>
<td>9</td>
<td>2.6</td>
<td>103.5</td>
<td>103.0</td>
<td>103.5</td>
<td>104.0</td>
<td>98.5</td>
<td>104.0</td>
</tr>
<tr>
<td>10</td>
<td>7.9</td>
<td>97.8</td>
<td>98.3</td>
<td>97.8</td>
<td>98.9</td>
<td>99.4</td>
<td>86.1</td>
</tr>
<tr>
<td>11</td>
<td>3.7</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>103.0</td>
</tr>
<tr>
<td>Ave.</td>
<td>5.1</td>
<td>99.5</td>
<td>99.6</td>
<td>99.6</td>
<td>99.8</td>
<td>N/A</td>
<td>99.5 ± 5.7</td>
</tr>
</tbody>
</table>

Table 14 and 15 show the results of coverage for necrosis and normal tissue volume. These were tissues that were also irradiated by the treatment. Notice similar coverage with the plans resulting from the segmentation methods compared to physician 1 and the actual treatment plan (last column). Table 15 shows that the KG treatment plan for Patient 9 delivered a lower dose to the damaged normal tissue. The tumor volume for this patient included a small enhancing area with cystic center in the posterior section of
the brain. This was not identified by the KG method as is evidenced in the 3D reconstruction in Figure 32. This effect was discussed in the previous chapter.

Table 15.
Data evaluation for post-MRI damaged normal tissue volume

<table>
<thead>
<tr>
<th>Patient</th>
<th>Volume (cm³)</th>
<th>% Dose Phys.1</th>
<th>% Dose Phys.2</th>
<th>% Dose Phys.3</th>
<th>% Dose kNN</th>
<th>% Dose KG</th>
<th>%Dose TXPlan</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6.5</td>
<td>99.4</td>
<td>99.4</td>
<td>99.4</td>
<td>99.4</td>
<td>N/A</td>
<td>101.1</td>
</tr>
<tr>
<td>2</td>
<td>26.3</td>
<td>98.9</td>
<td>99.4</td>
<td>99.4</td>
<td>99.4</td>
<td>N/A</td>
<td>98.9</td>
</tr>
<tr>
<td>3</td>
<td>28.7</td>
<td>97.0</td>
<td>97.0</td>
<td>96.5</td>
<td>96.5</td>
<td>N/A</td>
<td>100.0</td>
</tr>
<tr>
<td>6</td>
<td>12.7</td>
<td>102.5</td>
<td>102.5</td>
<td>102.5</td>
<td>102.5</td>
<td>N/A</td>
<td>102.0</td>
</tr>
<tr>
<td>8</td>
<td>78.2</td>
<td>98.0</td>
<td>98.0</td>
<td>98.0</td>
<td>96.5</td>
<td>97.0</td>
<td>91.5</td>
</tr>
<tr>
<td>9</td>
<td>30.6</td>
<td>101.0</td>
<td>100.5</td>
<td>101.0</td>
<td>101.0</td>
<td>91.5</td>
<td>101.5</td>
</tr>
<tr>
<td>10</td>
<td>22.3</td>
<td>88.9</td>
<td>92.2</td>
<td>91.7</td>
<td>89.4</td>
<td>93.3</td>
<td>68.3</td>
</tr>
<tr>
<td>11</td>
<td>11.5</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>103.0</td>
</tr>
</tbody>
</table>

Ave. 27.1 ± 22.4 98.2 ± 4.1 98.6 ± 3.1 98.6 ± 3.3 98.1 ± 4.1 N/A ± 11.7
IV.11. Discussion

In Chapter II it was shown that the in-plane contour extraction process allowed reasonable gross tumor volume outlines, but left the computer segmentation GTVs lacking in the slices immediately above and below the volume. Using this insight, the results from the computer segmentation methods were modified vertically to compensate for their under-segmentation by mimicking the process used by the physicians to contour the most superior and inferior tumor volume outlines. This expert knowledge could be automatically applied to the segmentation method routine.

A planning target volume was determined using the automatic margin expansion technique. It was found that it provided volumes comparable to those of the reference physician (PTV1). This was the result of both the modification of the vertical extension of the resulting automated GTV volume and the averaging effect of the volume expansion technique used for the creation of the planning target volume. The rolling-ball algorithm averaged the differences found in original contouring of the GTV, which were mostly at the edges.

The resulting treatment plans from both physician and computer segmentation PTVs were evaluated and compared to the plan resulting from physician 1 PTV. Only the cone down treatment plan was evaluated since the treatment volume generated for this plan is based solely on the GTV. In this study, there was no evaluation of the capabilities of the segmentation methods to identify edema. Further studies are required to identify the accuracy of edema outlining and the consequences for the initial planning target volume. Nevertheless, the use of the GTV volume is appropriate since the GTV
receives the greatest radiation dose and is the most important volume in radiation therapy planning. Studies addressing the issue of dose escalation concentrate in delivering higher dose of radiation solely to the gross tumor volume. No efforts are taken to cover areas of edema for these escalation protocols\(^{(3,7)}\).

The treatment plans designed from operator defined tumor volumes with treatment plans developed using computer segmentations were compared based on dose statistics, dose volume histograms, and conformity indices. All comparisons were made with respect of the coverage of the reference physician volume, PTV1. The dose statistics provided a measure of the differences between physicians and automatic segmentations in terms of delivered dose. The results indicated a good target coverage. A slight under-dosage was found from the kNN and KG treatment plans in less than 5% of the total PTV volume indicating no significant possible clinical effects. An analysis of how much normal brain tissue would be irradiated based on the dose volume histogram showed no significant difference between the different treatment plans.

All treatment plans were created without modification of beam aperture to eliminate human bias intervention. Nevertheless, conformity indices showed adequate results for all plans. A comparison of the data showed a better conformity for treatment plans with a higher number of fields, as shown by previous research\(^{(48)}\).

A novel modified definition of a conformity index, referenced herein as cross-conformity indices, was introduced to calculate the conformity of dose coverage to the reference target volume resulting from a plan generated from a computer segmentation or physician contour. The traditional CI comparisons are used to evaluate different
treatment plan modalities that were developed based on the same target\(^{(47)}\). For this case, the CCIs were used to compare plans created on other targets and superimposing their coverage on a reference target. The same level of conformity was found for the dose distributions resulting from the different treatment plans. This analysis showed that the plans generated from the computer segmentations were acceptable to treat the outline generated by the reference physician and could have been used in a clinical situation.

A novel evaluation mechanism based on follow-up imaging was developed to validate the results of the hypothetical treatment plans based on the outcome of the actual clinical patient treatment. The method presented utilizes existing imaging modalities for the patients under consideration, but the evaluation scheme could be applied to any imaging modality. The results showed that the segmentation plans provide similar dose coverage to areas being treated by the original treatment plans.

It can be seen that the automated systems evaluated in this study produced treatment plans consistent with the reference radiation oncologist. Additionally, accuracy of the dose coverage was validated by comparing the hypothetical treated areas with the actual patient treated areas. This was the result of the modification of the vertical extension of the resulting automated GTV volume and the averaging effect of the volume expansion technique used for the creation of the planning target volume. This mechanism shows that automated computer segmentation methods could be a viable tool in radiation therapy treatment planning.
CHAPTER V

CONCLUSIONS

Despite advances in diagnosis and treatment modalities, the prognosis for patients with brain tumors remains poor. Malignant gliomas are the most frequent primary brain tumors encountered in the adult age group and represent a major cause of mortality in neurological practice. Surgery followed by radiation therapy remains the primary therapeutic approach for most brain tumors. This type of tumor infiltrates normal brain tissues quite deeply, making complete surgical resection virtually impossible. Therefore, radiation therapy remains the most effective postoperative treatment for these patients.

Radiation therapy attempts to minimize radiation dose to normal brain tissue while delivering the highest possible dose to the target volume. Recent studies indicate that under-irradiation is a major contributor to the relatively poor success rate in treating glioma patients. The current efficacy of radiotherapy is limited since these tumors are highly resistant to radiation. This has motivated the exploration of escalating the dose delivered to the tumor bed, where the majority of treatment failures occur. The proper detection of tumor volume is a pre-requisite for such dose escalation. However, these tumors are infiltrative in nature and the definition of tumor margins is extremely difficult.

The need for improvements was identified when evaluating the current practice of tumor volume definition by radiation oncologists. Our findings indicate that the current technique of manual tumor outlining is time-intensive and highly subjective resulting in a
large variability between different physicians, even among outlines produced by the same physician when faced with the same data at different points in time. The use of automated computer segmentation algorithms as a tool to improve the efficiency, accuracy, and reproducibility of tumor definition in radiation treatment planning presents one possible solution to this problem.

This study investigated the application of two state-of-the-art automated tumor segmentation methods as tools for tumor volume definition for radiation therapy, kNN as a representative of the operator assisted semi-automated method and KG as a candidate for the fully automated method. Although not yet perfected enough for stand-alone use in a clinical setting in their current incarnations, automated MRI segmentation systems hold the promise of integrating data from multiple sequences to define tumor volumes for brain radiotherapy with a greater degree of precision and in a more consistent manner than outlining by radiation oncologists.

In the course of this study, various algorithms were developed to integrate the results from the automated computer segmentations techniques with radiation therapy treatment planning. These included the extraction of volume of interest, the manipulation of image data and segmentation results to match the treatment planning coordinate system, the extraction of contour points, and the formatting of the contour files to include the results of the segmentations in the treatment planning software.

An evaluation of the resulting automated contours showed that the kNN method performed better than the KG, probably due to its user input for initial selection of training data. The completely automated KG method was limited to enhancing tumors
and gliomas which exhibited clear enhancing edges. There is need for more work on the KG method to make it fully compatible for use in radiation therapy, including modifications to permit contouring of partially enhancing tumors, resection cavities and non-enhancing tumors.

The study assessed the viability of the resulting automated contours with the contours from radiation oncologists by introducing a novel, probabilistic measurement of accuracy. Starting with the assumption that the true target volume is found through the consensus of expert radiation oncologists, the results demonstrated that the kNN and KG methods under-segment the tumor volume compared with the physicians but are within the variability of the contouring performed by experienced radiation oncologists based on the same data. Additionally, it was found that the contouring produced by both of the automated segmentation systems became increasingly less accurate as the contouring moved from the central regions of a tumor towards its edges. The largest discrepancy was found at the vertical edges of the GTV. These results led to the proposed modifications to the results obtained from the kNN and KG routines that made them significantly more useful for creating radiation therapy treatment plans through an improvement in tumor delineation near the vertical axes to compensate for their consistent under-segmentation in these regions.

The next step in the study was to compare the gross tumor volumes (with vertical enhancements) produced by the automated systems with those produced from identical data by experienced radiation oncologists in a hypothetical clinical setting. This analysis compared the cone down radiation treatment plans designed from the physician defined
tumor volumes with similar plans developed from the modified segmentation results produced by the automated system. The comparison showed that the treatment plans based on both sets of tumor volume definitions resulted in dose coverage to the reference planning target volume within the same level of conformity as presented in Chapter IV. Only the cone down treatment plan was evaluated since the treatment volume generated for this type of plan is based solely on the gross tumor volume. Nevertheless, the use of the GTV volume is appropriate since the GTV receives the greatest radiation dose and is the most important volume in radiation therapy dose escalation studies.

A validation mechanism was designed to determine the accuracy of the results from the segmentation treatment plans with the original patient treatment plans. This evaluation scheme was based on existing follow-up MRI imaging but could be applied to any imaging modality. The results showed that the plans produced using the images generated by the automated systems provide similar dose coverage to the target areas as those produced using the images provided by the radiation oncologists.

Based on the results of this study it can be seen that the current level of sophistication of the automated systems is insufficient for them to replace the contouring generated by radiation oncologists; however, it is also apparent that with the continued evolution of these systems, they are likely to reach a level of expertise in contouring which will eventually be close to that of experienced radiation oncologists in accuracy and exceed it in consistency. Even in their current incarnations, the automated systems produced treatment plans in this study of comparable efficacy to those produced by the physicians. Thus combining the information obtained from the segmentation methods
could help radiation oncologists (especially those with limited experience) identify the target volume with greater accuracy and produce three-dimensional tumor volume models which would be more consistent and reproducible than those currently produced by different physicians or at different facilities. Automatic tumor outlining has the potential to speed the contouring process in radiation treatment planning and aid in multi-center trials since it would prevent physician- and center- bias that can affect trial outcomes.

This study was the first attempt at identifying and designing the steps needed to incorporate automatic brain tumor delineation in radiation therapy and the evaluation mechanisms to identify their success in treatment planning. Based on the foregoing, improvements are clearly needed to each of the elements of the proposed automatic segmentation systems to reach the goal of fully automatic GTV determination for radiation therapy treatment planning. This study has identified certain areas for improvement and made suggestions to address some of these areas. However, this study has also shown, through quantitative and qualitative analysis, that the long-term goal of fully automatic tumor delineation for glioma radiation treatment planning is not just theoretically possible, but is likely to be achieved within the next decade as others with more specific expertise in the relevant fields will build on this research in the future.
REFERENCES


Appendix A: Intra- and inter-operator variability

The intra-operator variability was calculated as the ratio of the average disagreement, that is, the size of each volume minus the intersection of the three volumes, divided by the average size of the three volumes. This calculation is represented in the following figure:

![Figure A1](image)

**Figure A1**
Representation of intra-operator variability calculation

In the above figure, $V_{i1}$, $V_{i2}$, $V_{i3}$ indicate tumor volume delineated by the radiation oncologist $i$ three times and shaded area $V_{i}(\text{int})$ represents the intersection of all 3 volumes. This results in the following formula to calculate the intra-operator variability:

$$\text{COV}_{i}\text{ intra} = \frac{1}{3} \sum_{j=1}^{3} (V_{i} - V_{i(\text{int})}) \times 100\%$$

**Formula [A.1]**

The inter-operator variability was calculated using the nine sets of outlines for each of the eleven patients and then calculating the disagreement from the outline prepared by each physician for each patient with the corresponding outline prepared by each of the other two physicians for that same patient. Figure A2 shows the disagreement of one volume of one physician $i$ ($V_{i1}$) with one volume of another physician $j$ ($V_{j1}$).
The final intra-operator variability is the average of the comparison of all volumes of one physician with all the volumes of the other physician as shown in the following formula:

\[
COV_{i\text{in}}^{\text{er}} = \frac{1}{9} \sum_{i=1}^{3} \sum_{n=1}^{3} \frac{V_{in} - V_{jn}}{V_{in} + V_{jn}} \times 100\% 
\]

This calculation is done for each of the three physicians resulting in the average variability between the three contours for each patient prepared by one physician with the other six sets of contours prepared by the other two physicians for that same patient.
Appendix B: Calculation of accuracy

Accuracy was calculated by assuming that the probability that a region is part of the definition of gross tumor volume is reflected by the number of times that region is included in any of the nine outline volumes produced by the three radiation oncologists. Every pixel in the image is labeled with an integer value (0 to 9) corresponding to the number of physician contours in which it was included. This pixel label provides the weight for measuring accuracy. Figure B1 shows a single volume of either physician ($V_{ij}$) or segmentation ($V_k$) compared with three (out of the nine) physician volumes. An area of higher pixel label weight is represented by level of gray in the figure, i.e. the area where the single volume being evaluated intersects more physician contours:

![Figure B1](representation_of_accuracy_calculation.png)

Final accuracy or true-positive is then expressed as the ratio of the total sum of pixel weights that was included by the physician or segmentation volume ($V_{ij}$ or $V_k$, showed as shaded areas in previous figure) to the total sum of pixel weights of the nine
Appendix B  (Continued)

volumes produced by the physicians (represented by all area enclosed by the three volumes \( V_{i1}, V_{i2} \) and \( V_{i3} \) in Figure B1):

\[
\text{Accuracy}_{(V_{i}\ or \ V_{k})} = \frac{\sum_{\text{image pixels}} (V_{ij} \ or \ V_{kj}) \cdot \frac{1}{9} \sum_{i=1}^{3} \sum_{j=1}^{3} V_{ij}}{\sum_{\text{image pixels}} \frac{1}{9} \sum_{i=1}^{3} \sum_{j=1}^{3} V_{ij}} \times 100\% \quad \text{Equation [B.1]}
\]

Similarly, excluded accuracy or false-positive is expressed as:

\[
\text{Exc. Accuracy}_{(V_{i}\ or \ V_{k})} = \frac{\sum_{\text{image pixels}} (V_{ij} \ or \ V_{kj}) \cdot (1 - \frac{1}{9} \sum_{i=1}^{3} \sum_{j=1}^{3} V_{ij})}{(1 - \sum_{\text{image pixels}} \frac{1}{9} \sum_{i=1}^{3} \sum_{j=1}^{3} V_{ij})} \times 100\% \quad \text{Equation [B.2]}
\]
Appendix C: IDL software programs

IDL software programs used throughout this dissertation can be found by enclosed CD-rom or by the enclosed file link.
ABOUT THE AUTHOR

Gloria Patrika Mazzara received a B.S. in physics from the University of Costa Rica in 1989, a M.S. in Physics from Florida State University in 1993, and a M.S. in Medical Health Physics from the University of Florida in 1995. She started working as a radiation therapy medical physicist in 1994 and entered the Ph.D. program at the University of South Florida in 1997 while continuing to work full-time in the field of radiation oncology.

While working on her Ph.D., Ms. Mazzara was certified in Therapeutic Radiological Physics in 1998, acted as an officer of the Florida Chapter of the American Association of Physicists in Medicine from 1998 to 2001, and presented numerous lectures on radiation therapy. She is currently working with Varian Medical Systems where she is involved in the education and implementation of treatment planning systems and new radiation therapy techniques. Her current interests include the application of imaging technology to the field of radiation therapy.