STOCHASTIC MODELS FOR EVOLUTION OF TUMOR GEOMETRY FOR CERVICAL CANCER DURING RADIATION THERAPY

by

Yifang Liu

A thesis submitted in conformity with the requirements for the degree of Master of Applied Science
 Graduate Department of Mechanical and Industrial Engineering
 University of Toronto

© Copyright 2013 by Yifang Liu
Abstract

Stochastic Models For Evolution Of Tumor Geometry for Cervical Cancer During Radiation Therapy

Yifang Liu
Master of Applied Science
Graduate Department of Mechanical and Industrial Engineering
University of Toronto
2013

Adaptive radiation therapy re-optimizes treatment plans based on updated tumor geometries from magnetic resonance imaging scans. However, the imaging process is costly in labor and equipment. In this study, we develop a mathematical model that describes tumor evolution based on a Markov assumption. We then extend the model to predict tumor evolution with any level of information from a new patient: weekly MRI scans are used to estimate transition probabilities when available, otherwise historical MRI scans are used. In the latter case, patients in the historical data are clustered into two groups, and the model relates the new patient’s behavior to the existing two groups. The models are evaluated with 33 cervical cancer patients from Princess Margaret Cancer Centre. The result indicates that our models outperform the constant volume model, which replicates the current clinical practice.
Acknowledgements

I want to express my deep gratitude towards Dr. Chi-Guhn Lee and Dr. Timothy Chan for their excellent supervision over the course of my masters program. Their sharp intelligence and prudent guidance have been a great source of motivation and inspiration for me. In addition, I would like to thank Dr. Mohammad Islam and Dr. Young-Bin Cho in their support in obtaining data for cervical cancer at Princess Margaret Cancer Centre, and their constant feedback to this project from a clinical perspective. This largely enhances the practicality of this research. Because of the effective collaboration with these four professors, I am able to finish two papers during the program. I want to thank all of them for their time and patience in reviewing the papers, and special thank to Tim for writing with me in person, where I learn the most in academic writing.

For support far beyond the ordinary, I deeply appreciate the frequent comments from both labs that I belong to on my papers and conference talks. I also want to thank Dr. Birsen Donmez and Dr. Jeffery Rosenthal for their help in statistics related problems.

This program is funded by the Department of Mechanical and Industrial Engineering, the Ontario Consortium for Adaptive Interventions in Radiation Oncology (OCAIRO) via the Ontario Research Fund (ORF), and the MITACS Accelerate Internship Program. I would like to thank all sponsors for their vision and their kind support.

Last but not least, I would like to thank my parents Fang and Songhao for their lifelong support that helped me to get where I am and eventually made this work possible, and my boyfriend Andrew for his love and support in good times and bad.
Contents

1 Introduction .................................................. 1
   1.1 Adaptive Radiation Therapy for Cervical Cancer ......... 1
   1.2 Tumor Motion and Deformation ........................... 3
   1.3 Mathematical Models for Tumor Evolution During Radiation Therapy ............................................. 4
       1.3.1 Tumor evolution ........................................... 4
       1.3.2 Stochastic behavior of tumor evolution ................. 4
       1.3.3 Discrete cell population models ........................ 5
   1.4 Tumor Classification ........................................ 5
   1.5 Research Objectives ....................................... 6

2 A General Model Setup ....................................... 8
   2.1 Patient Data ............................................... 8
   2.2 Data Processing .......................................... 11
   2.3 Spatial system ............................................. 11
       2.3.1 Voxel based spatial system ............................ 11
       2.3.2 Bi-voxel based spatial system ......................... 12
   2.4 Tumor Alignment ......................................... 14
       2.4.1 The COM approach: minimize the distance between centroids ........................................... 14
       2.4.2 The surface approach: minimize the distance between tumor surfaces ............................ 14
       2.4.3 Summary ................................................ 18
   2.5 The Markov Assumption .................................... 18

3 Stochastic Models for Tumor Evolution with Individual Information .................................................. 21
   3.1 Methods .................................................... 21
       3.1.1 Markov model I: prediction with individual information .............................................. 21
       3.1.2 An isomorphic shrinkage model: modeling voxel layers on the tumor surface ............... 23
       3.1.3 A constant volume model ................................ 23
   3.2 Performance Metrics ....................................... 23
   3.3 Results and discussion .................................... 25
       3.3.1 Performance of the Markov model I .................. 25
       3.3.2 Comparing Markov model I, the isomorphic shrinkage model, and the constant volume model ........................................... 29
       3.3.3 A statistical comparison between the Markov and isomorphic shrinkage models .......................... 31
Chapter 1

Introduction

This chapter presents a comprehensive literature review. Section 1.1 provides a general background about cervical cancer, radiation treatment, and a recently established concept of adaptive radiation treatment. Section 1.2 introduces the detection of tumor motion and deformation, which calls for a dynamic treatment plan over the course of treatment. We incorporate the adaptation of “margin” into our model in chapter 3 and 4. Since this study focuses on the prediction of tumor shrinkage and growth, we reduce the effect from tumor motion using tumor alignment in chapter 2. Section 1.3 reviews mathematical models that have been used to describe tumor evolution with and without radiation therapy in the literature, and our approach falls within the class of discrete cell models. More details about the model setup can be found in chapter 2. Section 1.4 provides an early incentive for tumor classification in the literature and prepares the tumor classification in chapter 4. Section 1.5 concludes this chapter with long-term and short-term goals of the current study.

1.1 Adaptive Radiation Therapy for Cervical Cancer

According to statistics from Canadian Cancer Society [1] and American Cancer Society [2], in the year 2013 throughout North America, about 13,790 new cases of invasive cervical cancer will be diagnosed. Among them 32.66% women will die from cervical cancer. Treatment for cervical cancers includes surgery, radiation therapy, chemotherapy, and complementary therapies. Radiation therapy has been the cornerstone in the treatment of this disease and now shows much potential for progress [3].
Radiation therapy irradiates and kills cancer cells by damaging their DNAs. The objective of radiation therapy is to irradiate the target with sufficient dose while sparing the organs-at-risk (OARs) simultaneously. Due to microscopic spread of cancer cells to normal tissues around the gross tumor, clinical target volume (CTV) is much larger than the gross tumour volume (GTV). In cervical cancer, CTV consists of GTV, upper vagina, lower uterus within 2cm from GTV, parametrium tissue and involved lymph nodes. Figure 1.1 shows one slice of a CTV including its three major components.

High precision radiation dose delivery has been one of the main objectives for enhancing tumor control and minimizing radiation-induced toxicity in the management of gynaecological malignancies. Chan et al. reported on the use of magnetic resonance imaging (MRI) to characterize intra-and inter-fractional motion of the cervix and surrounding structures for the explicit purpose of developing intensity-modulated radiation therapy (IMRT) approaches. These studies were repeated on a weekly basis on 20 patients with biopsy-proven cervical cancer undergoing radical radiation therapy, and represent one of the most comprehensive libraries of radiation therapy induced changes in diseased and normal anatomy in the literature. Of greater interest is the potential to dynamically adapt the treatment to these changing structures for the purpose of reducing the volume of normal tissue irradiated and guaranteeing coverage of the target with sufficient dose.

Standard treatment for cervical cancer is chemo-radiation therapy with an external radiation beam with dose prescription of 1.8Gy or 2Gy per fraction for 25 fractions to the whole pelvis (± node), followed by brachytherapy of 40Gy to the residual target, or IMRT if brachytherapy is not feasible. The quality of an external beam radiation therapy (EBRT) treatment for cervical cancer is heavily influenced by the change in the volume, shape and location of the target over the course of treatment. In the case of cervical cancer, the target volume can be reduced to a quarter of its original volume in five weeks of treatment, as also shown in Figure 1.2 with the subject group in this study. Thus, the effectiveness of a treatment assuming a static target can be rapidly degraded, particularly in terms of sparing OARs. Lim et al. proposed a radiobiological model for evaluating tumor regression in terms of volume change.

![Figure 1.2: Tumor volume change before and after the treatment](image-url)
using weekly MRI scans for over 27 patients with cervical cancer, and then correlated the results with pretreatment measurements of tumor hypoxia. The relationship between radiobiological determinants and tumor regression was examined and estimated using simulations. The authors concluded that early prediction of tumor regression can improve disease control and reduce toxicity.

In the recent decade, the concept of adaptive treatment planning in cervical cancer has shown substantial promise [6]. Adaptive radiation therapy (ART) is a feedback control strategy to manage patient specific variations during the course of radiation treatment. Key components of ART include dose assessment, variation evaluation, modification decisions and treatment modification [9]. Most clinical ART applications have been limited to target positional corrections and much effort is still required in 4D uncertainty modeling in order to “close the loop.” Our work in this paper aims to develop a mathematical model that can “predict” the evolution of the geometry (primarily shrinkage) of a cervical tumor undergoing radiation therapy. Such a model may support components of adaptive radiation therapy like variation evaluation by supplementing cheaper imaging modalities and reducing the need for expensive imaging activities and modification decisions by forecasting patient specific variations.

1.2 Tumor Motion and Deformation

Tumor motion, including rotational and translational movement, has been detected in clinical cases. In 2005, Court et al. proposed a CT-guided online ART algorithm [10], which compared online CT images with planning CT images and took into account local shifts and longitudinal extensions, in addition to global shifts. The developed algorithm was compared with the commonly used couch-shift method during two fractions for two prostate cancer patients: improvement was found in the superior portion of the prostate and the seminal vesicles when the in-slice shape change was small. In 2008, Chan et al. [5] studied daily MRI scans of 20 cervical cancer patients, constructed points of interest to consider rotational and translational movement, and investigated the inter- and intra-fractional tumor and organ motion. They observed that the motion is greatest at the fundus of the uterus, less at the isthmus and least at the cervical os. In addition, inter-fractional motion is more conspicuous than intra-fractional motion. Accordingly, they suggested that daily imaging be used before each fraction to ensure precise ART. Based on the observation from Chan et al., James et al. [6] implemented a weekly ART for 33 cervical cancer patients and investigated the dosimetric effect as a result of the interplay of the tumor deformation and the organ motion. The strategy achieved a reduction in treatment-induced toxicities. However, the geometric deformation algorithm has not been widely adopted in cervical cancer and hence requires further study.

In order to account for tumor motion and setup variation, a margin (of 7mm used by the Princess Margaret Cancer Centre) is added to the CTV to form the planning target volume (PTV), which defines the region where dose is delivered to. In 2006, Court et al. [11] developed a technique that considered global and local rotations as well as shifts. For head and neck cancer, a 5mm margin was applied to the CTV, and the overall dose was reduced while the percentage volume receiving the prescribed dose was maintained, i.e., OAR sparing was enhanced. Recently, Lim et al. [8] showed that whole pelvis IMRT with an adequate margin was superior to a four field conformal plans, and the quality of the whole pelvis IMRT could be improved by adaptive on-line re-planning. Previous research suggested that the planning margin should consider stochastic behavior of the target motion which may be specific to each patient [8]. Development of an anisotropic planning target margin and its adaptation over the course of treatment is highly desirable. The current study adopts the concept of margin and explores the model performance
over a range of margin size.

In addition to tumor and OAR movement, Lim et al. [4] considered tumor volume regression and deformation in cervical cancer. An algorithm was introduced to resolve the geometric discrepancy in a tumor over time, and the accumulated dose was calculated to illustrate the dosimetric effect of tumor regression. A retrospective study of 20 cervical cancer patients suggested that smaller PTV margins could be applied to reduce dose on normal tissues. In the same year, Wu et al. [12] proposed a new geometric descriptor, the overlap volume histogram (OVH), to capture the geometric relationship between the CTV and the OARs. A new criterion involving the relationship between the OVH and the dose volume histogram (DVH) is established, which forms the basis of a quality control method for IMRT with prior data obtained from a database. The effectiveness of the approach is demonstrated through a retrospective study on head and neck cancer. The replanning results for parotids glands of over 13 patients shows that it might be possible to reduce the dose under equal constraints in treatment plans, which suggests that the proposed quality control method can guide physicians towards a more efficient treatment plan. One of the methods of the current work adopts the OVH approach and investigates the geometric relationships between the tumor and the OARs (see section 2.4.2).

1.3 Mathematical Models for Tumor Evolution During Radiation Therapy

1.3.1 Tumor evolution

As experimental studies on radiation therapy continued, many researchers became interested in modeling the tumor evolution during radiation therapy. Early models in the 1900s measured tumor growth in terms of tumor volume and tumor size (i.e., the number of cells in the tumor). Looney et al. [13, 14] performed a least square fitting with a non-linear function to the change of tumor volume over time, and then classified tumors with respect to their response to radiation treatment. Later, studies on the influence of the structural biology of tumor cells on the determination of the mechanical characteristics of cancerous mass growth began to prevail. Sachs et al. [15] investigated tumor growth kinetics with linear quadratic (LQ) models, and took into account the interplay between tumor and neovascularization during cancer growth and cancer therapy. McAneney et al. [16] and O’Rourke et al. [17] further investigated various mechanisms of solid tumor growth within the LQ model during radiotherapy. Other than modeling change in tumor volume or size over time, we focus on modeling evolution in tumor geometry or shape in this paper.

1.3.2 Stochastic behavior of tumor evolution

In the 2000s, the stochastic behavior of tumor growth called general attention. While Escudero [18] reported two stochastic partial differential equations to model the physical properties of cell proliferation and surface diffusion, Enderling et al. [19] studied cancer cell growth under radiation treatment and for the first time introduced a random motility term to the invasion model of Anderson et al. [20]. Recently, the study from Navin et al. [21] demonstrated that tumor evolution can be inferred by examining the sequence of multiple cells from the same cancer, which implies that the change of state of such population cells can be modeled as a stochastic process. We investigate the evolution of tumor geometry in terms of the changes of state of cells in this paper. We start with a simple assumption that such evolution follows
a Markov process and test the Markov model using real patient data from Princess Margaret Cancer Centre.

1.3.3 Discrete cell population models

There are two primary classes of techniques used to model tumor shape change under radiation therapy: continuum models and discrete cell models. Continuum models describe the natural growth and shrinkage due to radiation using differential equations [15, 16, 17, 19], thus modeling changes in continuous time. These models focus primarily on the volume of the tumor or the number of cells comprising the tumor. On the other hand, discrete cell models describe the evolution of voxels between different discrete states in discrete time [22, 23, 24].

With the huge advances in biotechnology, large amounts of data on phenomena occurring on a single cell scale became available. Researchers started to model single-cell-scale phenomena and then upscaled to obtain information about the large scale of tumor growth. Discrete cell models consider the state of each population of cells inside the tumor to be characterized by the velocity, the internal biological state and the interaction with the local biochemical environment. Since the system of equations for interactions between cells governed by statistical mechanics or Newton’s laws of motion is complicated and extremely computationally expensive, it is more common to replace physical laws of cell motion with cell movement rules.

Different discrete cell models incorporate different levels of complexity into cell movement, proliferation and extinction laws. Düehling et al. [22] proposed one of the first models of this kind. They conducted a series of three-dimensional simulations to model the stochastic invasion process, and determined the influence of radiotherapy on tumors. Though simulations of a dividing tumor cell have been validated by in vitro experiments, exact parameters chosen for cell-cycle phase duration are difficult to obtain. Qi et al. [23] proposed a model that was better parametrized with a set of minimal cellular automata rules. However, the measurement of kinetic parameters still proved difficult. Unlike these two models using a regular lattice, Kansal et al. [24] recently proposed a three-dimensional model, which considered a random fixed lattice and Voronoi tessellation. They pointed out the importance of considering the impact of cell mutations in addition to cell divisions.

The current study describes the evolution of voxels between different discrete states in discrete time, and the data used to support this study is provided in a static space. Therefore the discrete cell models with a static grid system are more appropriate for studying changes in the actual geometry or shape of the tumor over time.

1.4 Tumor Classification

Tumor volume changes, which reflect the speed that cells are cleared inside a tumor, vary according to patients and weeks. Figure 1.3 shows various tumor volume change ratios for all patients all weeks: some tumors shrink with a high rate (to around half of its size a week ago, i.e. \( \approx 0.5 \)), some shrink with a low rate (\( \approx 0.8 \)), and some grow (\( > 1 \)). Lim et al. [8] introduced the concept of cell clearance time, which defined the time until clearance of irreparably damaged cells from the tumor. It can range from 5 to 25 days. Tumors with a large cell clearance time bear with a large volume ratio. This motivates our work in this paper to classify tumors into two groups and relate the tumor evolution of an incoming patient to that of the two existing groups.
Research has been done to classify tumors into either radioresponsive or radioresistant, according to which the treatment outcome of the radiation therapy is predicted. Most studies focus on molecular classification [25, 26, 26], and our work classify tumors from a macro perspective. The degree of tumor shrinkage is commonly used to represent tumor radioresponsiveness [27, 28, 29]. Therefore, our work also studies the classification by tumor shrinkage rate and explores the importance of obtaining the tumor volume information.

1.5 Research Objectives

The long-term goal of this research is to develop mathematical models to predict the CTV evolution in cervical cancer, analyze the dosimetric effect for different treatment plans and eventually optimize the treatment plan. Figure 1.4 provides an example of dose delivery region with different models. The shaded region represents the tumor shape. The red dotted line represent the dose delivery region generated by a mathematical model that can predict tumor evolution. And the black dotted line shows the region determined the constant volume model that replicates the current clinical practice, respectively.

In this study, our goal is to develop mathematical models to describe the change in a cervical tumor’s shape and size over time. Our approach falls within the class of discrete cell models, as we model the evolution of voxels in discrete time as a stochastic process. We classify static voxels as being in one of two states, tumor or non-tumor, at every epoch. Then we develop six Markov models, which use probabilities to govern the process of state transitions of individual voxels. An isomorphic shrinkage model takes a more coarse view of the tumor and models volume change by analyzing layers of voxels instead of individual voxels. We derive parameter values for our models from sequential MRI scans from Princess Margaret Cancer Centre in Toronto, Canada. Our developed models are then used to simulate the change in geometry of a cervical tumor over time. These simulations are tested against the actual images. Lastly, we quantify the trade-off between tumor coverage and OAR sparing of the developed
Figure 1.4: An example of dose delivery region with different models. The shaded region is the actual tumor. The region bounded by the black and red dotted line represent the prediction from the constant volume model and the Markov model.

models and a constant volume model.
Chapter 2

A General Model Setup

This chapter starts with a detailed description of the data used in the study, which was obtained from the Princess Margaret Cancer Centre. Section 2.2 presents the data inclusion/exclusion criteria. Then section 2.3 introduces a voxel based spatial system and a bi-voxel based spatial system, which enhances the former. As briefly introduced in chapter 1, the change of state is due to tumor motion and/or tumor shrinkage, and since this study focuses on the prediction of tumor shrinkage, we reduce the effect from tumor motion with tumor alignment, i.e., for each patient we align the tumor at \( t \) to the tumor at \( t - 1 \). Two objectives are proposed for tumor alignment in section 2.4 to minimize the distance between tumor centroid (center of mass), and to minimize the distance between tumor surfaces. The two objectives reach the same optimal solution for convex shapes, in particular for the GTV. We use tumor centroid alignment for this study, because it is more computationally efficient. Finally, section 2.5 introduces geometric configurations and formulates the transition probability that will be estimated in later chapters.

2.1 Patient Data

We obtained MRI scans for 33 cervical cancer patients treated at Princess Margaret Cancer Centre who were enrolled in a research ethics board approved study [8]. All patients received one MRI and one computed tomography (CT) scan for planning purposes, and either four or five additional weekly MRI scans during radiation therapy. The GTVs were contoured by a radiologist and reviewed by a radiation oncologist. Standard treatment for cervical cancer patients is chemo-radiation therapy with an external radiation beam with dose prescription of 1.8Gy or 2Gy per fraction for 25 fractions to the whole pelvis (± node) followed by brachytherapy of 40Gy to the residual target, or intensity-modulated radiation therapy if brachytherapy is not feasible.

Table 2.1 and Table 2.2 show the number of voxels in the tumor or on the tumor surface, respectively, at each week for all patient if it is available (otherwise a “-” is displayed). For our analysis, we excluded four patients (shaded red in Table 2.1) who had small tumors (fewer than 280 voxels) in the planning image. We also excluded four of the weekly MRI scans (shaded gray in Table 2.1) from the remaining 29 patients where the tumor shrinkage was substantial, spanning at least two layers of voxels. E.g., for patient indexed 12, the total number of voxels in the tumor at \( t = 4 \) is 284, and the number of voxels on the tumor surface is 177 (see Table 2.2). From \( t = 4 \) to \( t = 5 \) the actual number of voxel reduces by 192 (284-92) which is more than possible. In total, there were 53 pairs of consecutive weekly scans, which we
Table 2.1: Number of voxels in the tumor. The shaded lines are the patients that are eliminated because there are less than 280 voxels at $t = 2$. The shaded pairs of cells are the transitions (a pair of weeks) that are eliminated because the tumor changes are beyond one bi-voxel layer.

<table>
<thead>
<tr>
<th>Patient Index</th>
<th>$t = 2$</th>
<th>$t = 3$</th>
<th>$t = 4$</th>
<th>$t = 5$</th>
<th>$t = 6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>881</td>
<td>976</td>
<td>700</td>
<td>-</td>
<td>547</td>
</tr>
<tr>
<td>2</td>
<td>701</td>
<td>-</td>
<td>410</td>
<td>257</td>
<td>127</td>
</tr>
<tr>
<td>3</td>
<td>245</td>
<td>164</td>
<td>209</td>
<td>-</td>
<td>176</td>
</tr>
<tr>
<td>4</td>
<td>1487</td>
<td>887</td>
<td>532</td>
<td>402</td>
<td>421</td>
</tr>
<tr>
<td>5</td>
<td>705</td>
<td>433</td>
<td>421</td>
<td>332</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>841</td>
<td>837</td>
<td>715</td>
<td>657</td>
<td>513</td>
</tr>
<tr>
<td>7</td>
<td>1008</td>
<td>897</td>
<td>682</td>
<td>633</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>3660</td>
<td>2548</td>
<td>1493</td>
<td>1080</td>
<td>810</td>
</tr>
<tr>
<td>9</td>
<td>195</td>
<td>163</td>
<td>142</td>
<td>85</td>
<td>100</td>
</tr>
<tr>
<td>10</td>
<td>474</td>
<td>471</td>
<td>352</td>
<td>312</td>
<td>253</td>
</tr>
<tr>
<td>11</td>
<td>1767</td>
<td>1301</td>
<td>1003</td>
<td>675</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>317</td>
<td>245</td>
<td>284</td>
<td>92</td>
<td>87</td>
</tr>
<tr>
<td>13</td>
<td>1121</td>
<td>687</td>
<td>507</td>
<td>272</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>2796</td>
<td>2328</td>
<td>1960</td>
<td>1538</td>
<td>1540</td>
</tr>
<tr>
<td>15</td>
<td>463</td>
<td>329</td>
<td>342</td>
<td>216</td>
<td>187</td>
</tr>
<tr>
<td>16</td>
<td>830</td>
<td>-</td>
<td>435</td>
<td>351</td>
<td>244</td>
</tr>
<tr>
<td>17</td>
<td>1407</td>
<td>417</td>
<td>201</td>
<td>127</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>3322</td>
<td>-</td>
<td>1507</td>
<td>1242</td>
<td>1050</td>
</tr>
<tr>
<td>19</td>
<td>1446</td>
<td>1374</td>
<td>1142</td>
<td>824</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>562</td>
<td>501</td>
<td>284</td>
<td>218</td>
<td>120</td>
</tr>
<tr>
<td>21</td>
<td>339</td>
<td>-</td>
<td>187</td>
<td>152</td>
<td>122</td>
</tr>
<tr>
<td>22</td>
<td>328</td>
<td>178</td>
<td>-</td>
<td>131</td>
<td>73</td>
</tr>
<tr>
<td>23</td>
<td>823</td>
<td>734</td>
<td>773</td>
<td>715</td>
<td>732</td>
</tr>
<tr>
<td>24</td>
<td>581</td>
<td>-</td>
<td>509</td>
<td>433</td>
<td>340</td>
</tr>
<tr>
<td>25</td>
<td>-</td>
<td>1814</td>
<td>1044</td>
<td>628</td>
<td>386</td>
</tr>
<tr>
<td>26</td>
<td>450</td>
<td>362</td>
<td>196</td>
<td>163</td>
<td>144</td>
</tr>
<tr>
<td>27</td>
<td>1826</td>
<td>-</td>
<td>955</td>
<td>757</td>
<td>635</td>
</tr>
<tr>
<td>28</td>
<td>401</td>
<td>286</td>
<td>211</td>
<td>144</td>
<td>48</td>
</tr>
<tr>
<td>29</td>
<td>504</td>
<td>389</td>
<td>211</td>
<td>98</td>
<td>115</td>
</tr>
<tr>
<td>30</td>
<td>80</td>
<td>67</td>
<td>44</td>
<td>50</td>
<td>42</td>
</tr>
<tr>
<td>31</td>
<td>1105</td>
<td>760</td>
<td>467</td>
<td>365</td>
<td>-</td>
</tr>
<tr>
<td>32</td>
<td>225</td>
<td>144</td>
<td>107</td>
<td>55</td>
<td>-</td>
</tr>
<tr>
<td>33</td>
<td>439</td>
<td>393</td>
<td>222</td>
<td>-</td>
<td>125</td>
</tr>
</tbody>
</table>
Table 2.2: Number of voxels on the tumor surface. The shaded cells show the number of voxels that can be eliminated during the transitions shaded in Table 2.1.

<table>
<thead>
<tr>
<th>Patient Index</th>
<th>$t = 2$</th>
<th>$t = 3$</th>
<th>$t = 4$</th>
<th>$t = 5$</th>
<th>$t = 6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>395</td>
<td>440</td>
<td>347</td>
<td>-</td>
<td>304</td>
</tr>
<tr>
<td>2</td>
<td>336</td>
<td>-</td>
<td>229</td>
<td>161</td>
<td>92</td>
</tr>
<tr>
<td>3</td>
<td>166</td>
<td>116</td>
<td>143</td>
<td>-</td>
<td>128</td>
</tr>
<tr>
<td>4</td>
<td>595</td>
<td>421</td>
<td>305</td>
<td>254</td>
<td>249</td>
</tr>
<tr>
<td>5</td>
<td>322</td>
<td>244</td>
<td>235</td>
<td>190</td>
<td>-</td>
</tr>
<tr>
<td>6</td>
<td>438</td>
<td>410</td>
<td>335</td>
<td>340</td>
<td>284</td>
</tr>
<tr>
<td>7</td>
<td>470</td>
<td>400</td>
<td>331</td>
<td>324</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>1065</td>
<td>810</td>
<td>596</td>
<td>516</td>
<td>441</td>
</tr>
<tr>
<td>9</td>
<td>124</td>
<td>106</td>
<td>94</td>
<td>58</td>
<td>77</td>
</tr>
<tr>
<td>10</td>
<td>248</td>
<td>229</td>
<td>196</td>
<td>179</td>
<td>147</td>
</tr>
<tr>
<td>11</td>
<td>653</td>
<td>526</td>
<td>440</td>
<td>354</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>202</td>
<td>153</td>
<td>177</td>
<td>75</td>
<td>71</td>
</tr>
<tr>
<td>13</td>
<td>524</td>
<td>340</td>
<td>250</td>
<td>171</td>
<td>-</td>
</tr>
<tr>
<td>14</td>
<td>872</td>
<td>769</td>
<td>694</td>
<td>581</td>
<td>567</td>
</tr>
<tr>
<td>15</td>
<td>257</td>
<td>202</td>
<td>214</td>
<td>147</td>
<td>131</td>
</tr>
<tr>
<td>16</td>
<td>380</td>
<td>-</td>
<td>229</td>
<td>199</td>
<td>155</td>
</tr>
<tr>
<td>17</td>
<td>550</td>
<td>240</td>
<td>134</td>
<td>87</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>1112</td>
<td>-</td>
<td>649</td>
<td>580</td>
<td>508</td>
</tr>
<tr>
<td>19</td>
<td>588</td>
<td>563</td>
<td>468</td>
<td>398</td>
<td>-</td>
</tr>
<tr>
<td>20</td>
<td>293</td>
<td>261</td>
<td>177</td>
<td>139</td>
<td>87</td>
</tr>
<tr>
<td>21</td>
<td>204</td>
<td>-</td>
<td>144</td>
<td>117</td>
<td>96</td>
</tr>
<tr>
<td>22</td>
<td>191</td>
<td>123</td>
<td>-</td>
<td>99</td>
<td>63</td>
</tr>
<tr>
<td>23</td>
<td>366</td>
<td>344</td>
<td>365</td>
<td>343</td>
<td>327</td>
</tr>
<tr>
<td>24</td>
<td>331</td>
<td>-</td>
<td>279</td>
<td>249</td>
<td>210</td>
</tr>
<tr>
<td>25</td>
<td>-</td>
<td>640</td>
<td>433</td>
<td>288</td>
<td>217</td>
</tr>
<tr>
<td>26</td>
<td>242</td>
<td>214</td>
<td>132</td>
<td>116</td>
<td>107</td>
</tr>
<tr>
<td>27</td>
<td>636</td>
<td>-</td>
<td>405</td>
<td>337</td>
<td>322</td>
</tr>
<tr>
<td>28</td>
<td>236</td>
<td>166</td>
<td>137</td>
<td>110</td>
<td>42</td>
</tr>
<tr>
<td>29</td>
<td>248</td>
<td>204</td>
<td>111</td>
<td>76</td>
<td>84</td>
</tr>
<tr>
<td>30</td>
<td>64</td>
<td>55</td>
<td>41</td>
<td>44</td>
<td>39</td>
</tr>
<tr>
<td>31</td>
<td>489</td>
<td>357</td>
<td>255</td>
<td>212</td>
<td>-</td>
</tr>
<tr>
<td>32</td>
<td>144</td>
<td>111</td>
<td>87</td>
<td>45</td>
<td>-</td>
</tr>
<tr>
<td>33</td>
<td>281</td>
<td>293</td>
<td>183</td>
<td>-</td>
<td>100</td>
</tr>
</tbody>
</table>
refer to as “transitions” (e.g., if we have scans from week 2, 4, and 5 for a patient, we have one transition: 4 → 5). The time interval between the planning image and the start of treatment (i.e., between the first and second MRI scan) varied across patients, and therefore the planning scans (at t = 1) were not included in the study.

2.2 Data Processing

Based on the voxel size of the available reconstructed image, a $0.4 \times 0.4 \times 0.4 \text{ cm}^3$ spatial voxel grid was used for each image. A spatial voxel was classified as “tumor” if the proportion of the voxel that was covered by the tumor, as determined by the treatment planning software, was larger than a tumor classification threshold $\rho$. Otherwise, it was classified as “non-tumor”. An illustrative figure for one slice of the tumor is shown in Figure 2.1. The dotted outline shows the tumor contour and the shaded cells represent tumor voxels in our spatial grid. We plot the number of tumor voxels with $\rho$ ranging from 0 to 1 with an increment of 0.1 against the actual tumor volume. The best fit is obtained at $\rho = 0.4$. However, since physicians put more emphasize on tumor coverage compared to OAR sparing, we choose $\rho = 0$ for this study, i.e., as long as a voxel is covered (either completely or partially) by the actual tumor, it is classified as tumor. At the same time, the choice also benefit the study by including a larger number of voxels for statistical analysis.

Figure 2.1: An example slice from a 3D tumor. The dotted outline shows the tumor contour and the shaded cells represent tumor voxels in our spatial voxel grid. The voxels bounded by the solid lines comprise the bi-voxel layer on the tumor surface.

2.3 Spatial system

2.3.1 Voxel based spatial system

Formally, we define the state of voxel $i$ at time $t$ to be $S_t(i)$, which can take either a value of 1 (tumor state) or 0 (non-tumor state). Rather than modeling the state change of each voxel in the tumor, we focus on state changes of voxels on the tumor surface. Such cells are killed more effectively with radiation [30, 31] and tumor shrinkage can be represented as surface voxels changing from a tumor to non-tumor
state. Formally, the state of a voxel is,

\[ S_t(i) = \begin{cases} 
1, & \text{if the voxel } i \text{ is in tumor state at time } t \\
0, & \text{if the voxel } i \text{ is in non-tumor state at time } t 
\end{cases} \]  

(2.1)

State changes of a voxel \( i \) may be influenced by the states of the neighboring voxels. In other words, these changes may be influenced by whether the voxels adjacent to \( i \) in the sagittal, coronal, and transverse directions are tumor or non-tumor. Thus, we define distinct geometric configurations around a voxel, based on the states of the neighboring voxels. Each voxel has six neighbors in the sagittal, coronal, and transverse directions. There were a total of 41,441 voxels in all MRI scans analyzed. The frequency of occurrence for each geometric configuration is shown in Table 2.3 along with a visual representation of the configuration. A solid dot represents a voxel in a tumor state and a clear dot represents a voxel in a non-tumor state. Configuration \( g_8 \) is the most prevalent and is intuitively what we would expect to see – tumor on one side (five solid dots on the left) and the normal tissues on the other (one clear dot on the right).

Table 2.3: Frequency of occurrence of key geometric configurations

<table>
<thead>
<tr>
<th>Index</th>
<th>Frequency</th>
<th>Geometry</th>
<th>Index</th>
<th>Frequency</th>
<th>Geometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>( g_1 )</td>
<td>33</td>
<td><img src="image" alt="Configuration" /></td>
<td>( g_5 )</td>
<td>5037</td>
<td><img src="image" alt="Configuration" /></td>
</tr>
<tr>
<td>( g_2 )</td>
<td>376</td>
<td><img src="image" alt="Configuration" /></td>
<td>( g_6 )</td>
<td>10,793</td>
<td><img src="image" alt="Configuration" /></td>
</tr>
<tr>
<td>( g_3 )</td>
<td>0</td>
<td><img src="image" alt="Configuration" /></td>
<td>( g_7 )</td>
<td>0</td>
<td><img src="image" alt="Configuration" /></td>
</tr>
<tr>
<td>( g_4 )</td>
<td>139</td>
<td><img src="image" alt="Configuration" /></td>
<td>( g_8 )</td>
<td>25,063</td>
<td><img src="image" alt="Configuration" /></td>
</tr>
</tbody>
</table>

2.3.2 Bi-voxel based spatial system

To account for both shrinkage and growth on the tumor surface, we define a new quantity, the bi-voxel, which is composed of two adjacent voxels. For example, a bi-voxel overlapping the surface of the tumor will include one voxel on the surface of the tumor and one adjacent non-tumor voxel. An example bi-voxel layer on the tumor surface is shown in Figure 2.1 as the region between the two solid lines.

Formally, we define the state of bi-voxel \( (i, j) \) at time \( t \) to be \( B_t(i, j) := S_t(i) + S_t(j) \). In particular,

\[ B_t(i, j) = \begin{cases} 
0, & \text{if } S_t(i) = 0, S_t(j) = 0, \\
1, & \text{if } S_t(i) = 1, S_t(j) = 0 \text{ or } S_t(i) = 0, S_t(j) = 1, \\
2, & \text{if } S_t(i) = 1, S_t(j) = 1.
\end{cases} \]  

(2.2)

On the tumor surface, all bi-voxels are in state 1. Modeling tumor shrinkage or growth can now be
accomplished by modeling a change in the state of the appropriate bi-voxels from state 1 to either state 0 or 2, respectively. Unless otherwise stated, future references to “bi-voxel” refer to a bi-voxel on the tumor surface.

Similar to a voxel, state changes of a bi-voxel \((i,j)\) may be influenced by the states of the neighboring voxels. That is, these changes may be influenced by whether the voxels adjacent to either \(i\) or \(j\) in the sagittal, coronal, and transverse directions are tumor or non-tumor. Each voxel has six neighbors so a bi-voxel has 10 neighboring voxels. Each of the 10 neighboring voxels can be in one of the two states (tumor or non-tumor). Therefore, there are \(2^{10}\) possible geometric configurations surrounding each bi-voxel, some of which would be highly unlikely to be realized in any realistic tumor.

There were a total of 41,441 bi-voxels in all MRI scans analyzed. After combining isomorphic configurations and eliminating any geometric configurations that have a frequency of occurrence of less than 1% (i.e., less than 423 bi-voxels with such a configuration), 14 geometric configurations (out of \(2^{10}\)) remained. The frequency of occurrence for each geometric configuration is shown in Table 2.4 along with a visual representation of the configuration. A solid dot represents a voxel in a tumor state and a clear dot represents a voxel in a non-tumor state. The first configuration is the most prevalent and is intuitively what we would expect to see – a complete separation between the tumor (six solid dots on the left) and the non-tumor voxels (six clear dots on the right).

Table 2.4: Frequency of occurrence of key geometric configurations

<table>
<thead>
<tr>
<th>Index</th>
<th>Frequency</th>
<th>Geometry</th>
<th>Index</th>
<th>Frequency</th>
<th>Geometry</th>
</tr>
</thead>
<tbody>
<tr>
<td>(g_1)</td>
<td>10,697</td>
<td>![Image of configuration]</td>
<td>(g_8)</td>
<td>1,390</td>
<td>![Image of configuration]</td>
</tr>
<tr>
<td>(g_2)</td>
<td>5,918</td>
<td>![Image of configuration]</td>
<td>(g_9)</td>
<td>1,142</td>
<td>![Image of configuration]</td>
</tr>
<tr>
<td>(g_3)</td>
<td>5,245</td>
<td>![Image of configuration]</td>
<td>(g_{10})</td>
<td>1,041</td>
<td>![Image of configuration]</td>
</tr>
<tr>
<td>(g_4)</td>
<td>4,486</td>
<td>![Image of configuration]</td>
<td>(g_{11})</td>
<td>1,041</td>
<td>![Image of configuration]</td>
</tr>
<tr>
<td>(g_5)</td>
<td>2,525</td>
<td>![Image of configuration]</td>
<td>(g_{12})</td>
<td>974</td>
<td>![Image of configuration]</td>
</tr>
<tr>
<td>(g_6)</td>
<td>2,368</td>
<td>![Image of configuration]</td>
<td>(g_{13})</td>
<td>902</td>
<td>![Image of configuration]</td>
</tr>
<tr>
<td>(g_7)</td>
<td>1,446</td>
<td>![Image of configuration]</td>
<td>(g_{14})</td>
<td>850</td>
<td>![Image of configuration]</td>
</tr>
</tbody>
</table>

The comparison among transition probability estimates from different tumor alignment in section 2.4
considers geometric configurations based on voxel spatial system for simplicity. All Markov models for tumor evolution in chapter 3 and 4 are developed using bi-voxel spatial system.

2.4 Tumor Alignment

There are two main explanations for changes of state of voxels: target motion and tumor shrinkage or growth. This study assumes that target motion can be detected with inexpensive online imaging scans, and it focuses on the prediction of tumor shrinkage or growth. In order to reduce the effect from the target motion, for each patient we align the MRI scan at $t$ to the scan at $t - 1$. If the scan at $t - 1$ is not available, then we align it to the scan at $t - 2$ and so on. Two objectives are used to align the tumors of two consecutive MRI scans: to minimize the distance between centroids, and to minimize the distance between tumor surfaces.

2.4.1 The COM approach: minimize the distance between centroids

Formally, we write down the objective as follows:

$$
\begin{align*}
\text{minimize} & \quad \| \bar{c}^t - \bar{c}^{t-1} \| \\
\text{subject to} & \quad \bar{c}^t = R \circ T(c^t)
\end{align*}
$$

(2.3)

where $c^t$ is the centroid (center of mass) of the GTV at time $t$, and the rotation matrices is $R = R_{x,y,z} = R_x \circ R_y \circ R_z$ where

$$
R_z = \begin{bmatrix}
\cos \theta(z) & \sin \theta(z) & 0 & 0 \\
-\sin \theta(z) & \cos \theta(z) & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}, \quad R_x = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \cos \theta(x) & \sin \theta(x) & 0 \\
0 & -\sin \theta(x) & \cos \theta(x) & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}, \quad R_y = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & \cos \theta(y) & \sin \theta(y) \\
0 & 0 & -\sin \theta(y) & \cos \theta(y)
\end{bmatrix}
$$

(2.4)

and $\theta(x)$ and $\theta(y)$ are the angles of rotation around $x$ and $y$ axes, respectively; and translation matrix $T_{x,y,z}$ is:

$$
T_{x,y,z} = \begin{bmatrix}
1 & 0 & 0 & d(x) \\
0 & 1 & 0 & d(y) \\
0 & 0 & 1 & d(z) \\
0 & 0 & 0 & 1
\end{bmatrix},
$$

(2.5)

where $d=[d(x), d(y), d(z)]$ is the translation vector.

The COM approach, along with soft-tissue matching and bone matching (current clinical standard), is adopted in clinical study. This study adopts COM approach, because of its simplicity and computational efficiency.

2.4.2 The surface approach: minimize the distance between tumor surfaces

The second objective is to minimize the distance between two tumor surfaces, where surface is defined as the outermost layer of the tumor. Note that all voxels on the surface are in tumor state. Formally, we
write down the model as follows:

\[
\begin{align*}
\text{minimize} \quad & y(R,T) = \sum_{p \in S_t, q \in S_{t-1}} w_p^t \| \hat{L}_p - L_q \| \\
\text{subject to} \quad & q = \arg \min_{q \in S_{t-1}} \{ \min_{p \in S_t} \| \hat{L}_p - L_q \| \}
\end{align*}
\]

where \( L_p \) is the location of voxel \( p \), \( w_p^t \) is the weight assigned to voxel \( p \), \( S_t \) is the tumor surface at time \( t \), and the rotation matrix \( R \) and translation matrix \( T \) are defined in Equation 2.4 and Equation 2.5, respectively.

The steps to obtain the optimal values for both \( R \) and \( T \) with numerical approach are as follows:

1. Compute \( y \) with \( R_x, R_y, R_z \) and \( T_x, T_y, T_z \) being identity matrix, i.e. \( y_0 = y(I_4, I_4) \).
2. While keeping \( R_x, R_y, R_z \) identity matrix, compute \( T_x, T_y, T_z \) with the translation vector \( d + \Delta d \), where \( \Delta d \) takes all six vectors: \( [0 0 \pm \Delta d], [0 \pm \Delta d 0], \) and \( [\pm \Delta d 0 0] \). Since the study is based on a \( 0.4 \times 0.4 \times 0.4 \) cm\(^3\) voxel spatial system, we adopt an increment unit \( \Delta d = 0.4 \) cm. Then compute \( y_1 = y(I_4, T) \).
3. If \( y_1 < y_0 \), then go back to Step 1, else continue with Step 4 with \( y^* \) at \( T^* \), i.e., the sum of all \( T \) taken.
4. While keeping \( T = T^* \), compute \( R_x \circ R_y \circ R_z \circ T_{x,y,z}^*(L_p) \) with each \( R \) over \( \theta \) ranging from \( 0 \) to \( 360^\circ \) with an increment of \( 2^\circ \). Then compute \( y_2 = y(R,T) \).

Alter. 4. Write \( L(X_p^t) = [x_1, x_2, x_3] \) with polar coordinate such that

\[
\begin{align*}
x_1 &= r \cdot \sin(\theta_1) \cdot \cos(\theta_2) \\
x_2 &= r \cdot \sin(\theta_1) \cdot \sin(\theta_2) \\
x_3 &= r \cdot \cos(\theta_1)
\end{align*}
\]

Replace \( \theta_1 \) with \( \theta_1 + \Delta \theta_1 \) and \( \theta_2 \) with \( \theta_2 + \Delta \theta_2 \) for all \( \Delta \theta_1 \) between \( 0 \) to \( 180^\circ \) with an increment of \( 2^\circ \), and for all \( \Delta \theta_2 \) between \( 0 \) to \( 360^\circ \) with an increment of \( 2^\circ \); and then compare the old and the new \( y \) similarly. The alternative approach is more computationally efficient, since it is equivalent to Step 4 and requires fewer iterations.

5. If \( y_2 < y^* \), then go back to Step 3, else an optimal \( y^{**} \) is obtained at \( T^* \) (the optimal \( T \)) and \( R^* \) (the optimal \( R \)).

The second model is able to account for the influence from OAR movement by adding a weight function \( w_p^t \) in front of the smallest distance from one voxel on one tumor surface to another, weighted according to the proximity of OARs. The following subsections present four approaches to compute the weight \( w_p^t \).

**Weighting approach I: uniform weight**

The uniform weight assumes \( w_p^t = 1 \) for all \( p \in S_t \). Then \( S_{t-1} \) and \( S_t \) are aligned simply with respect to the smallest distance between the two surfaces.
Weighting approach II: distance to centroid of an OAR

The weight is formulated so that the closer the tumor to the OAR, the greater the influence of the cancer cell on the OAR, and the larger value $w_{tp}^t$ takes. We assume that such influence follows an exponential decay with respect to the distance between the tumor and the OAR, and therefore we have the following formulation for $w_{tp}^t$:

$$ w_{tp}^t = \sum_{r \in OAR} \exp(-\alpha_1 d^t(p, r)), \quad (2.8) $$

$$ w_{tp}^t = \exp(\sum_{r \in OAR} -\alpha_2 d^t(p, r)), \quad (2.9) $$

where $d^t(p, r)$ is the distance between voxel $p$ in the tumor and the centroid of the $r^{th}$ OAR, $\alpha_1$ and $\alpha_2$ are tunable parameters.

Weighting approach III: sensitivity to the movement of OAR

The weight is formulated so that the more an OAR moves towards the tumor (i.e., the closer an voxel on the tumor surface to the moving OAR), the greater is the influence, and the larger value $w_{tp}^t$ takes. This formulation is based on the approach II, and therefore it also takes into account the influence of the distance.

$$ w_{tp}^t = \sum_{r \in OAR} \exp(-d^t(p, r) \cdot \max(0, \frac{\alpha_3}{\Delta d^t(p, r)})), \quad (2.10) $$

where $\Delta d^t(p, r) = d^t(p, r) - d^{t-1}(p, r)$, and $\alpha_3$ is a tunable parameter.

Weighting approach IV: sensitivity to the rate of movement of an OAR

The overlap volume histogram (OVH) describes the spatial configuration of an OAR relative to the tumor. It defines the overlap volume of an OAR and the tumor as the tumor shrinks or grows while the surrounding OAR remains static. Figure 2.2 shows the OVH curves for different OARs, which indicates the different configurations that OARs possess. A relative smooth OVH curve indicates a relatively convex shape with a smooth surface of an OAR. The shape of the uterus (Figure 2.2a) is more convex than the shape of the sigmoid (Figure 2.2b) for the patient under study. The weight in this context relates to the slope at $x = 0$ (i.e., the point where the tumor is at its original size). The larger the slope, the greater is the influence and hence the larger value $w_{tp}^t$ obtains. Similar to approach II, we have two sets of formulation for $w_{tp}^t$ regarding the location of the exponential function:

$$ w_{tp}^t = \sum_{r \in OAR} \exp(-\alpha_4 d^t(p, r)) \cdot \varphi_p^t, \quad (2.11) $$

$$ w_{tp}^t = \exp(\sum_{r \in OAR} -\alpha_5 d^t(p, r) \cdot \varphi_p^t), \quad (2.12) $$

where $\varphi_p^t = 1 - \exp(-\alpha_4 \cdot o_p \cdot s_p)$ denotes the influence on voxel $p$ at $t$, $o_p$ is the original size of the tumor, $s_p$ is the slope of OVH at the original size of the tumor, $\alpha_4$ and $\alpha_5$ are tunable parameters.
Figure 2.2: OVH with OARs for a patient. The x-axis shows the increment from the tumor, and the y-axis shows the corresponding overlap volumes between the tumor and the OAR of interest.
2.4.3 Summary

The results of tumor GTV alignment with the two objectives are similar. Since the COM approach is more computationally efficient, we will use COM approach for tumor alignment as a preliminary step to prediction of tumor evolution. CTV, as a combination of three relatively convex shapes, is less convex. So there exist some discrepancy in tumor CTV alignment. Our long term goal of this research is to predict tumor CTV evolution. Therefore the result for CTV alignment is presented in this section as a reference for future study.

As a result of the COM approach, there are 3 out of 33 transitions where the rotation degree is greater than 15° in one axis. And the surface approach with weight 1 results in 4 out of 33 transitions where the rotation degree is greater than 15° in one axis. The tumor shape of the exceptional transitions are more symmetric, i.e., the shapes do not change significantly during rotation. In addition, the extreme rotations occur during early weeks (e.g., $t = 1$ to $t = 2$, or $t = 2$ to $t = 3$), which are the base of the prediction in the main model and hence require no tumor alignment. Furthermore, the average rotation for other transitions is less than 3° for both cases. Therefore for this study we ignore the rotation effect and consider only the translation effect.

We compute $y$ in Equation 2.6 for 33 patients’ CTV from MRI scans with all derived weights except for the initial scan before treatment ($t = 1$), since the initial scan is discarded for the study. Neither do we perform tumor alignment on the first scan during treatment ($t = 2$), since it is the first scan within our study. The different weighting methods may affect transition probability estimation in the main model, which determines the accuracy of the prediction. Figure 2.3 illustrates the difference in transition probability estimates using different weighting approaches. In each subfigure, the $x$ axis represents the 8 possible geometric configurations based on the voxel spatial system that was discussed in section 2.3.1. The colored bars in each column represent different weeks. And the red line represents the change of mean of the weekly values. The $y$ axis shows the transition probability estimate that a voxel (in state 1) on the tumor surface at $t$ will change to a non-tumor state at $t + 1$. The figure indicates that $g_2$ and $g_7$, though possible theoretically, do not exist in the real case. For geometric configurations that most frequently occur (e.g., $g_5$, $g_6$, $g_8$), the probability estimates are similar among different weighting methods. The values vary though for the least frequent geometries. However, since the predicted tumor shape is, to a large extent, determined by the transition probability estimates, and since the transition probability estimates are dominated by the most frequent geometries, the prediction as a result of different weighting approaches are similar.

2.5 The Markov Assumption

In a Markov model, we assume that the evolution of tumor geometry follows a Markov process. A Markov process is a stochastic process where the future depends only on the current state of the process and not on any additional history [32]. In the present context, the Markov assumption implies that only the current geometry of the tumor (and not previous geometries) influences the evolution of the tumor shape in the future. To support this assumption, we assume that the dose delivered is uniform over the tumor each week and affects all voxels in the same way.

In a bi-voxel spatial system, state changes of the voxels can be influenced by multiple factors including motion of the tumor and cell death due to radiation. We focus on changes induced by radiation and therefore we reduce the effect of motion via the tumor alignment process that was discussed in section 2.4.
Figure 2.3: Transition probability estimates for each geometric configuration over weeks. The x axis shows the 8 possible geometric configurations (see Table 2.3) and y axis is the probability that a voxel is in non-tumor state during the next week. The various colors represent 5 weeks.
We assume that the state of voxel $i$ at time $t$ is dependent only on the state of voxel $i$ and of its neighboring voxels at time $t-1$. In other words, we assume a Markov process for the state transitions according to the following equation:

$$P(B_t(i,j) = s_t \mid B_{t-1}(i,j) = s_{t-1}, G_{t-1} = g), \quad (2.13)$$

where $s_t$ and $s_{t-1}$ take values in $\{0, 1, 2\}$, $G_{t-1}$ is the state of the geometric configuration of the neighboring voxels at time $t-1$, and $g \in \{g_1, \ldots, g_{14}\}$ (cf. Table 2.4).

Since we focus on bi-voxels on the tumor surface, we are only interested in the transition probabilities in equation (2.13) where $s_{t-1} = 1$. We define a matrix $\Pi_t$ of dimension $14 \times 3$, which contains the transition probabilities for all 14 possible configurations and three possible state changes. The $(l,k)$-th element of the matrix is

$$P(B_t(i,j) = k \mid B_{t-1}(i,j) = 1, G_{t-1} = g_l), \ \forall k = 0, 1, 2, \ l = 1, \ldots, 14. \quad (2.14)$$

The next chapter estimates the transition probabilities, which governs the Markov process and hence determines the accuracy of the prediction.
Chapter 3

Stochastic Models for Tumor Evolution with Individual Information

This chapter aims at verifying that a Markov model performs well, especially compared to the constant volume model that replicates the current clinical approach. For each patient, we calculate the transition probability $\Pi_t$ from $t-1$ to $t$ using the individual MRI scans at $t-1$ and $t$, and use $\Pi_t$ to predict the tumor shape at $t+1$ based on the actual tumor shape at $t$. The transition probability estimation is discussed in section 3.1. In the same section, we develop an isomorphic shrinkage model that does not require complicated parameter estimation, and formally define the constant volume model. Before presenting the results, section 3.2 lists the metrics to evaluate the model performance. We choose to present sensitivity and specificity as they indicate the tumor coverage and OAR sparing, respectively. In order to compare the trade-off curves consisting of the two metrics, we define an acceptable region and the Hausdorff distance. Finally, we conclude that the Markov model performs better than the constant volume model in terms of the trade-off between sensitivity and specificity, and the isomorphic shrinkage approach is a good alternative to the Markov model when high-resolution imaging scans are not available.

3.1 Methods

3.1.1 Markov model I: prediction with individual information

The Markov model I analyzes the changes of voxels on the tumor surface from $t-1$ to $t$, and based on the tumor shape at $t$ the model predicts the tumor shape at $t+1$. We calculate the transition probabilities from $t-1$ to $t$ using maximum likelihood estimation (MLE): if there are $n$ bi-voxels with geometry $g_t$ on the tumor surface at time $t-1$, (note that the $n$ bi-voxels are in state 1,) and $n_0$, $n_1$ and $n_2$ ($n_0 + n_1 + n_2 = n$) are found to convert to state 0, 1 and 2 respectively at time $t$, then the transition probability conditional on geometry $g_t$ is

$$P(B_t(i, j) = k \mid B_{t-1}(i, j) = 1, G_{t-1} = g_t) = \frac{n_k}{n}, \quad (3.1)$$
Chapter 3. Stochastic Models for Tumor Evolution with Individual Information

for \( k = 0, 1, 2 \) and \( g_t = 0, 1, \ldots, 14 \).

To prove the statement, we assume that the probabilities that given \( g_t \), a bi-voxel change to \( k \) is \( p_k \), for \( k = 0, 1, 2 \). Then it follows that

\[
p_0 + p_1 + p_2 = 1 \quad (3.2)
\]

The log likelihood that \( p_{1,0} = p_0, p_{1,1} = p_1, p_{1,2} = p_2 \) is the following:

\[
LL(p_{1,0} = p_0, p_{1,1} = p_1, p_{1,2} = p_2) = \log(p_0^{x_0} \cdot p_1^{x_1} \cdot p_2^{x_2}) \quad (3.3)
\]

To maximize the log likelihood, we differentiate \( LL \) with respect to \( p_k \), for \( k = 0, 1, 2 \). Below illustrates the differentiation with respect to \( p_1 \).

\[
\frac{\partial LL}{\partial p_1} = \frac{x_1 \cdot p_1^{x_1-1} \cdot p_2^{x_2} \cdot p_0^{x_0} - x_0 \cdot p_1^{x_1} \cdot p_2^{x_2} \cdot p_0^{x_0-1}}{p_1^{x_1} \cdot p_2^{x_2} \cdot p_0^{x_0}} \quad (3.4)
\]

Set Equation (3.4) = 0, we obtain

\[
\frac{p_1}{p_3} = \frac{x_1}{x_0}
\]

Similarly, by differentiating Equation (3.3) with respect to \( p_2 \) we obtain

\[
\frac{p_2}{p_0} = \frac{x_2}{x_0}
\]

From the property that \( p_1 + p_2 + p_3 = 1 \),

\[
p_1 = \frac{x_1}{x}, \quad p_2 = \frac{x_2}{x}, \quad p_0 = \frac{x_0}{x} \quad (3.5)
\]

Therefore the point estimate of the transition probability is the frequency count of the transitions for a given geometric configuration.

We estimate \( \Pi_t \) for each week \( t \) and each patient. For those patients who have few (less than 20) or no bi-voxels of configuration \( g_t \) in week \( t - 1 \), we derive the transition probability from the entire population of patients who have configuration \( g_t \) in week \( t - 1 \). The value 20 was chosen as a lower limit to ensure granularity in the probability numbers of at most 0.05 (\( 20 = 1/0.05 \)). Therefore more than 20 voxels associated with each of the 14 geometric configurations is required on the tumor surface in order for the tumor to be included in the study (recall the definition of “small” in our inclusion/exclusion criteria and note that 280 = 14 \times 20).

For each patient, we estimate \( \Pi_t \) from MRI scans at time \( t - 1 \) and \( t \). Then, based on the MRI scan at time \( t \), which tells us the states of the bi-voxels at time \( t \), we predict the state of the bi-voxels at time \( t + 1 \). At time \( t \), for each configuration \( g_t \), we simulate the state change of the associated bi-voxels 500 times according to the estimated transition probabilities. Then we average the simulated images to generate a probability map, where each voxel has a value corresponding to the probability that it is in a tumor state at time \( t + 1 \). Finally, we convert the voxel-specific probability back to a binary “tumor”/“non-tumor” classification by using a probability threshold \( \beta \), above which the voxel is classified as “tumor”. Lastly, we add a margin of size \( \alpha \) around the predicted tumor shape where \( \alpha \) ranges from 0 to 0.4 cm.
3.1.2 An isomorphic shrinkage model: modeling voxel layers on the tumor surface

For the second approach, we consider a simple alternative to the Markov model that focuses on the volume of the entire tumor rather than modeling individual voxels. The isomorphic shrinkage approach assumes that the proportional change in tumor volume from week $t - 1$ to $t$ is the same for week $t$ to $t + 1$. For each patient, the tumor volume for week $t - 1$ and $t$ in terms of the number of voxels in the tumor, $\text{Vol}(t - 1)$ and $\text{Vol}(t)$, are obtained from the MRI scans. To generate the volume in week $t + 1$, we first calculate the number of layers of voxels, in increments of 0.1 cm, to eliminate from the volume at week $t - 1$ so that the new volume matches the volume at week $t$. Formally, we write down the volume after eliminating $z$ layers of size 0.1 cm to be $\text{Vol}_z(t - 1)$, and $z$ is obtained when Equation 3.6 is satisfied.

$$\text{Vol}_z(t - 1) < \text{Vol}(t) < \text{Vol}_{z-1}(t - 1)$$ (3.6)

Then, the number of layers we remove is proportional to the radius of an equivalent spherical volume, i.e. the nearest integer to $z \times (\sqrt[3]{\text{Vol}(t + 1)/\text{Vol}(t)})$, from the volume at week $t$ to generate the volume at week $t + 1$. Like the Markov model, after the proportional shrinkage operation we add a margin of size $\alpha$ around the predicted tumor shape where $\alpha$ ranges from 0 to 0.4 cm.

3.1.3 A constant volume model

In this approach, we assume the GTV remains constant for the duration of the treatment. This replicates the most common practice currently employed in clinics. In this approach, as well as the two previous approaches, we add a margin of size $\alpha$ around the predicted tumor shape.

3.2 Performance Metrics

In the computational results, we compare the predicted tumor geometries generated by the three models against the actual geometries using sensitivity and specificity. In this context, sensitivity indicates the level of tumor coverage provided, and specificity relates to OAR sparing. Let TP (true positive) be the number of voxels that we correctly predict to be tumor, TN (true negative) be the number of voxels that we correctly predict to be non-tumor, FP (false positive) be the number of voxels that we incorrectly predict to be tumor, and FN (false negative) be the number of voxels that we incorrectly predict to be non-tumor. Table 3.1 summarizes the above definitions.

<table>
<thead>
<tr>
<th>Test outcome</th>
<th>actual class (observation)</th>
</tr>
</thead>
<tbody>
<tr>
<td>predicted class (expectation)</td>
<td>T, N</td>
</tr>
<tr>
<td></td>
<td>T, TP, FP</td>
</tr>
<tr>
<td></td>
<td>F, FN, TN</td>
</tr>
</tbody>
</table>

Formally, sensitivity and specificity are defined as

$$\text{sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$,
$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}.$$ (3.7)
Chapter 3. Stochastic Models for Tumor Evolution with Individual Information

An F score is commonly used in statistics to evaluate the classification problem with a single number, which is defined as

\[ F = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}, \]  

where precision and recall are defined as

\[ \text{precision} = \frac{TP}{TP + FP}, \quad \text{recall} = \frac{TP}{TP + FN}. \]  

Though F-score summarizes the performance into a single number, it assumes that FP and FN are equally important for the classification. In clinical sense, it assumes that the influence of mis-prediction over the tumor voxels and the non-tumor voxels are the same, which conflicts with the clinical practice that the tumor coverage is more important than the OAR sparing. Therefore, we can compare a weighted measure of the two metrics presented above, and the weight shall be chosen by the clinicians.

In addition, we compute two conformity indexes (CIs): the volume ratio (VR) between the simulated and the actual tumors, and the volume overlap ratio (VOR) or the overlap ratio (OR) as defined by the intersection of the simulated and the actual tumors divided by their union. Formally, we define the two metrics as

\[ \text{VR} = \frac{TP + FP}{FN + TP}, \]  

\[ \text{VOR} = \frac{TP}{FN + TP + FP}. \]

In the result section in both chapter 3 and 4, we choose sensitivity and specificity to present in the result section since they are sufficient to reflect the goodness of the model in terms of the two objectives in clinical study: the tumor coverage and the OAR sparing.

Qualitatively, an “acceptability” threshold of 0.93 was chosen for both sensitivity and specificity. These metrics have a similar interpretation as the conformity index. Our choice of 0.93 is derived from the conformity index criteria from Hazard et al. Note that the exact value chosen is somewhat arbitrary, but the primary observations we make in the next section hold regardless of the threshold. Quantitatively, we compute the Hausdorff distance in order to compare the trade-off curves generated by different models. The directional Hausdorff distance \( \delta_H(a, b) \) between two curves \( a \) and \( b \) is defined as follows:

\[ \delta_H(a, b) = \max_{p \in a} \min_{q \in b} d(p, q), \]  

where \( d(p, q) \) is the Euclidean distance between point \( p \in a \) and \( q \in b \). Note that the optimal solution to the max-min problem to obtain the directional Hausdorff distance is affected by the length of the two curves of interest, e.g., an additional point in curve \( a \) may increase \( \delta_H(a, b) \), and therefore \( \delta_H(a, b) \) is sensitive to the choice of \( a \). In order to eliminate the sensitivity, we derive a new metric \( \Delta_H(a, b) \):

\[ \Delta_H(a, b) = \min(\delta_H(a, b), \delta_H(b, a)). \]

We introduce a benchmark curve in later section and compute the distance between all derived models and the benchmark curve, the smaller the value \( \Delta_H(a, b) \), the better the performance of the model under the clinical objectives carried by benchmark curve.
3.3 Results and discussion

3.3.1 Performance of the Markov model I

Table 3.2 shows the sensitivity and specificity values for all transitions over a range of margin size \( \alpha \) and probability threshold \( \beta \). All figures in the following sections are produced using data from this table.

The Markov model has two tunable parameters, i.e., the margin size \( \alpha \) and the probability threshold \( \beta \). Figure 3.1 and Figure 3.2 show the influence to the mean sensitivity and specificity from the probability threshold \( \beta \) and the margin size \( \alpha \), respectively. Figure 3.1a and Figure 3.1b indicate that as \( \beta \) increases, there are fewer voxels on the tumor surface classified as being in tumor state, and therefore sensitivity decreases while specificity increases for the Markov model. The constant volume model and the isomorphic shrinkage model do not have the parameter \( \beta \) and they produce binary tumor shapes. The figures also show that the Markov model takes values between the other two models: it is more flexible than the constant model and yet more conservative than the isomorphic shrinkage model.

Figure 3.2a and Figure 3.2b show the influence from the margin size \( \alpha \). Note that we add margin to the constant volume model only for fair comparison among models, and the one that replicates the current clinical practice is the constant volume model at \( \alpha = 0 \). As \( \alpha \) increases, we add a larger number of voxel layers above the predicted tumor shape, which results in a larger predicted tumor shape that increases the sensitivity while decreasing the specificity. Since \( \alpha \) is also applied to the volume model and the isomorphic shrinkage model, they all share a similar trend when tuning the parameter. Note that while the mean specificity of the three models share a parallel decrease, the increase of mean sensitivity tends to converge given a large margin size. It is justifiable that with a large margin size, the three models are less distinguished from each other because the margin depreciates the accurate prediction from the Markov model. The trend implies that the Markov model performs best among all, because with the same mean sensitivity it can still maintain a large specificity.

While the Markov model seems to perform well against the isomorphic shrinkage and constant volume models, there were instances when it did not predict tumor geometry change well. Out of the 53 transitions (i.e., a prediction of tumor geometry change in consecutive weeks for a particular patient) considered, there were 50 transitions where some part of the sensitivity-specificity trade-off curve overlapped the acceptable region for some value of \( \alpha \) and \( \beta \). In the remaining three cases, no value of \( \alpha \) or \( \beta \) resulted in performance within the acceptable region. For those three transitions, we examined several factors, including tumor size at time of prediction \( t \), tumor shape at time of prediction \( t \), and the consistency of volume change ratios during \( t - 1 \) to \( t \) and \( t \) to \( t + 1 \), to try to understand the poor performance of the Markov model. Figure 3.3 and Figure 3.5 visualize the correlation between the factors under examination and the performance of the corresponding transitions. We did principal component analysis (PCA) on the possible factors and we determined that the convexity of the tumor has a major influence on how well the Markov model predicted geometry change: tumor shapes that are highly non-convex are much harder to predict. We computed a convexity measure \( \text{[35]} \) for the tumor at the start of each transition: the higher the convexity measure the more convex the shape is. Two examples of tumors with different degrees of convexity are shown in Figure 3.4.

Figure 3.5 shows that tumors with the smallest convexity measures are hard to predict from the perspective of sensitivity. Specificity prediction does not seem to be affected by tumor convexity, which is expected as specificity relates to OAR sparing. Note that there are examples in Figure 3.5 where high convexity measures may still generate relatively low sensitivity values (in the range 0.94). Further study is needed to understand the factors that challenge tumor geometry prediction using our Markov model.
Table 3.2: Sensitivity and specificity values for all patients and available images from weeks 4-6 for a range of $\beta$ values.

<table>
<thead>
<tr>
<th>Pat.</th>
<th>Week</th>
<th>$\beta = 0$</th>
<th>$\beta = 0.1$</th>
<th>$\beta = 0.2$</th>
<th>$\beta = 0.3$</th>
<th>$\beta = 0.4$</th>
<th>$\beta = 0.5$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Sens Spec</td>
<td>Sens Spec</td>
<td>Sens Spec</td>
<td>Sens Spec</td>
<td>Sens Spec</td>
<td>Sens Spec</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>0.99 0.93</td>
<td>0.99 0.93</td>
<td>0.99 0.93</td>
<td>0.98 0.94</td>
<td>0.94 0.96</td>
<td>0.93 0.96</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>1.00 0.97</td>
<td>1.00 0.97</td>
<td>1.00 0.97</td>
<td>1.00 0.97</td>
<td>1.00 0.97</td>
<td>1.00 0.97</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>1.00 0.94</td>
<td>1.00 0.96</td>
<td>0.98 0.97</td>
<td>0.97 0.98</td>
<td>0.98 0.96</td>
<td>0.96 0.98</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>1.00 0.97</td>
<td>0.97 0.98</td>
<td>0.96 0.98</td>
<td>0.96 0.98</td>
<td>0.96 0.98</td>
<td>0.96 0.98</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>0.91 0.98</td>
<td>0.88 0.99</td>
<td>0.88 0.99</td>
<td>0.88 0.99</td>
<td>0.89 0.99</td>
<td>0.88 0.99</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>0.98 0.97</td>
<td>0.97 0.98</td>
<td>0.97 0.98</td>
<td>0.97 0.98</td>
<td>0.97 0.98</td>
<td>0.97 0.98</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>1.00 0.96</td>
<td>1.00 0.96</td>
<td>1.00 0.96</td>
<td>1.00 0.96</td>
<td>1.00 0.96</td>
<td>1.00 0.97</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>1.00 0.94</td>
<td>0.99 0.95</td>
<td>0.93 0.96</td>
<td>0.93 0.96</td>
<td>0.93 0.96</td>
<td>0.93 0.97</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>0.97 0.95</td>
<td>0.94 0.95</td>
<td>0.92 0.95</td>
<td>0.91 0.96</td>
<td>0.91 0.96</td>
<td>0.91 0.96</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>1.00 0.95</td>
<td>1.00 0.95</td>
<td>1.00 0.96</td>
<td>0.99 0.96</td>
<td>0.99 0.96</td>
<td>0.99 0.96</td>
</tr>
<tr>
<td>11</td>
<td>5</td>
<td>0.94 0.98</td>
<td>0.92 0.98</td>
<td>0.91 0.98</td>
<td>0.91 0.98</td>
<td>0.91 0.98</td>
<td>0.91 0.98</td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>1.00 0.94</td>
<td>0.99 0.96</td>
<td>0.96 0.97</td>
<td>0.94 0.98</td>
<td>0.93 0.98</td>
<td>0.93 0.98</td>
</tr>
<tr>
<td>13</td>
<td>5</td>
<td>1.00 0.97</td>
<td>0.99 0.98</td>
<td>0.97 0.99</td>
<td>0.97 0.96</td>
<td>0.99 0.99</td>
<td>0.99 0.98</td>
</tr>
<tr>
<td>14</td>
<td>6</td>
<td>0.99 0.97</td>
<td>0.98 0.99</td>
<td>0.97 0.99</td>
<td>0.95 0.99</td>
<td>0.95 0.99</td>
<td>0.95 0.98</td>
</tr>
<tr>
<td>15</td>
<td>4</td>
<td>1.00 0.95</td>
<td>1.00 0.95</td>
<td>1.00 0.95</td>
<td>1.00 0.95</td>
<td>1.00 0.95</td>
<td>1.00 0.95</td>
</tr>
<tr>
<td>16</td>
<td>5</td>
<td>0.97 0.97</td>
<td>0.93 0.99</td>
<td>0.90 1.00</td>
<td>0.90 1.00</td>
<td>0.90 1.00</td>
<td>0.90 1.00</td>
</tr>
<tr>
<td>17</td>
<td>6</td>
<td>0.96 0.98</td>
<td>0.93 0.98</td>
<td>0.93 0.98</td>
<td>0.93 0.98</td>
<td>0.93 0.98</td>
<td>0.93 0.98</td>
</tr>
<tr>
<td>18</td>
<td>5</td>
<td>0.99 0.96</td>
<td>0.93 0.96</td>
<td>0.93 0.97</td>
<td>0.93 0.97</td>
<td>0.93 0.97</td>
<td>0.93 0.97</td>
</tr>
<tr>
<td>19</td>
<td>6</td>
<td>0.99 0.98</td>
<td>0.99 0.98</td>
<td>0.98 0.98</td>
<td>0.98 0.98</td>
<td>0.98 0.98</td>
<td>0.98 0.98</td>
</tr>
<tr>
<td>20</td>
<td>4</td>
<td>1.00 0.97</td>
<td>1.00 0.97</td>
<td>0.99 0.97</td>
<td>0.99 0.97</td>
<td>0.99 0.97</td>
<td>0.99 0.97</td>
</tr>
<tr>
<td>21</td>
<td>5</td>
<td>0.99 0.99</td>
<td>0.99 0.99</td>
<td>0.98 0.99</td>
<td>0.88 1.00</td>
<td>0.88 1.00</td>
<td>0.88 1.00</td>
</tr>
<tr>
<td>22</td>
<td>6</td>
<td>0.93 0.97</td>
<td>0.90 0.97</td>
<td>0.90 0.98</td>
<td>0.88 0.98</td>
<td>0.88 0.98</td>
<td>0.88 0.98</td>
</tr>
<tr>
<td>23</td>
<td>4</td>
<td>1.00 0.94</td>
<td>1.00 0.94</td>
<td>0.99 0.96</td>
<td>0.98 0.96</td>
<td>0.96 0.96</td>
<td>0.96 0.96</td>
</tr>
<tr>
<td>24</td>
<td>5</td>
<td>0.99 0.95</td>
<td>0.98 0.96</td>
<td>0.95 0.96</td>
<td>0.95 0.96</td>
<td>0.95 0.96</td>
<td>0.95 0.96</td>
</tr>
<tr>
<td>25</td>
<td>6</td>
<td>1.00 0.94</td>
<td>1.00 0.94</td>
<td>1.00 0.95</td>
<td>1.00 0.95</td>
<td>1.00 0.95</td>
<td>1.00 0.95</td>
</tr>
<tr>
<td>26</td>
<td>4</td>
<td>0.98 0.97</td>
<td>0.97 0.97</td>
<td>0.97 0.97</td>
<td>0.97 0.97</td>
<td>0.97 0.97</td>
<td>0.97 0.97</td>
</tr>
<tr>
<td>27</td>
<td>5</td>
<td>1.00 0.97</td>
<td>1.00 0.97</td>
<td>1.00 0.97</td>
<td>1.00 0.97</td>
<td>1.00 0.97</td>
<td>1.00 0.97</td>
</tr>
<tr>
<td>28</td>
<td>4</td>
<td>1.00 0.95</td>
<td>0.99 0.96</td>
<td>0.98 0.97</td>
<td>0.97 0.98</td>
<td>0.97 0.98</td>
<td>0.97 0.98</td>
</tr>
<tr>
<td>29</td>
<td>5</td>
<td>0.97 0.97</td>
<td>0.96 0.98</td>
<td>0.96 0.98</td>
<td>0.96 0.98</td>
<td>0.96 0.98</td>
<td>0.96 0.98</td>
</tr>
<tr>
<td>30</td>
<td>4</td>
<td>0.99 0.95</td>
<td>0.97 0.95</td>
<td>0.97 0.95</td>
<td>0.96 0.96</td>
<td>0.88 0.97</td>
<td>0.88 0.97</td>
</tr>
</tbody>
</table>
Figure 3.1: Influence to the mean sensitivity and specificity from the probability threshold $\beta$, holding the margin size $\alpha = 0$. 
Figure 3.2: Influence to the mean sensitivity and specificity from the margin size $\alpha$, holding the probability threshold $\beta = 0.3$. 

(a) Mean sensitivity

(b) Mean specificity
3.3.2 Comparing Markov model I, the isomorphic shrinkage model, and the constant volume model

Figure 3.6 compares the sensitivity and specificity values achieved by the three models across a range of parameter settings. Each curve of solid dots is associated with a different $\alpha$ value: increasing $\alpha$ corresponds to curves moving towards the right. Within a curve, the dots correspond to different $\beta$ values: points generally move down and to the right with increasing $\beta$. The diamonds represent the isomorphic shrinkage model, one per $\alpha$ moving to the right with increasing $\alpha$. The circle represents the constant volume model. Figure 3.6a plots the mean sensitivity and specificity values, averaged over all transitions of all patients. Figure 3.6b plots the combination of 10th percentile of the sensitivity and specificity distributions, derived from all transitions of all patients.

First, a clear and expected trade-off is evident across all models: increasing sensitivity is generally accompanied by decreasing specificity and vice versa. For the Markov model, given a fixed $\alpha$, sensitivity decreases and specificity increases as $\beta$ increases, as expected. Figure 3.6a shows that the isomorphic shrinkage model is comparable with the Markov model when mean performance is considered. However, it underperforms the Markov model in the lower part of the sensitivity and specificity distributions, as shown in Figure 3.6b. Figure 3.6b shows that the Markov model dominates both the isomorphic shrinkage model and the constant volume model, and that the Markov model has the most “overlap”
with the acceptable region. In other words, there are parameter settings for the Markov model in which
the results demonstrate both sensitivity and specificity values above 0.93. On the other hand, the constant
volume model does not enter the acceptable region. Furthermore, both subfigures indicate that the best
performance in sensitivity and specificity is attained by combinations of small margin parameter ($\alpha$)
with a small probability threshold ($\beta$) or a large margin parameter with a large probability threshold.

We believe the Markov model offers several advantages over the other models. The Markov model
takes the most granular approach to modeling geometry, considering each voxel (or bi-voxel) separately.
This, along with the additional adjustable parameter $\beta$, provides the Markov model with the capability to
achieve generally dominant performance in terms of the sensitivity-specificity trade-off. We imagine that
sensitivity outweighs specificity in clinical situations. However, given a particular acceptable threshold for
sensitivity, it appears that it is possible for the Markov model to meet the same level of sensitivity but with
increased specificity, compared with the other models. Note that if the definition of “acceptable” changes,
the Markov model provides flexibility through the $\beta$ parameter to try and meet the new acceptability
criteria. Note that the strength of Markov model over the other two models is most evident in Figure 3.6b,
which depicts performance at the 10th percentile (i.e., 90% of transitions), which is more clinically relevant
than mean performance.

Figure 3.7 compares the weekly performance of the models. Generally, the performance of the Markov
model improves over the course of treatment. On the other hand, the performance of the constant model,
particularly specificity, degrades as $t$ increases. Notice that the isomorphic shrinkage model prefers a
larger margin to maintain acceptable sensitivity. Regardless of the value of the chosen margin parameter,
the Markov model generates acceptable results by adjusting the parameter $\beta$ accordingly (larger $\beta$ for
larger $\alpha$ and vice versa).

As shown in Figure 3.7, the performance of the Markov model generally improves as the treatment
progresses. Though tumor cells may be killed right after the start of treatment, it may take a few weeks
for complete cell clearance, only after which volume reduction is observed [8]. The degradation in the
performance of the constant volume model as $t$ increases is intuitive: specificity relates to OAR sparing
and modeling a shrinking tumor using a constant volume will result in increased OAR dose.
3.3.3 A statistical comparison between the Markov and isomorphic shrinkage models

We performed a Wilcoxon signed-rank test against the hypothesis that the Markov model is significantly different from the isomorphic shrinkage model in terms of sensitivity and specificity. The p-values are shown in Table 3.3a (sensitivity) and 3.3b (specificity), for different combinations of $\alpha$ and $\beta$. Shaded cells indicate a p-value of less than 0.05 (i.e., significant at the 95% level). Generally, the two models perform differently (Markov outperforms isomorphic shrinkage) when both $\alpha$ and $\beta$ are small.

In the case that high quality scans are not available to train the Markov model, the isomorphic shrinkage model may be a viable alternative. While the isomorphic shrinkage model is generally dominated by the Markov model, recall that performance was similar for larger $\alpha$ and $\beta$ values (cf. Tables 3.3a and 3.3b). Increasing $\beta$ will result in more voxels in the bi-voxel layer of interest being classified as non-tumor and increasing $\alpha$ will put less emphasis on the voxel by voxel classification of the Markov model, eventually eliminating the difference between the Markov and isomorphic shrinkage models.
Figure 3.6: Performance of the three models in terms of sensitivity and specificity, where $\alpha$ is the margin size and $\beta$ is the probability threshold.
Figure 3.7: Weekly performance of the three models with and without a margin, where $t$ is the week at prediction and $\beta$ is the probability threshold.
Table 3.3: p-values of Wilcoxon signed-rank test on sensitivity. Null hypothesis: no difference between the Markov and isomorphic shrinkage models in sensitivity. Alternative hypothesis: difference between the two models in sensitivity. Shaded cells correspond cases where the null hypothesis is rejected ($p < 0.05$).
Chapter 4

Stochastic Models for Tumor Evolution with Population Information

Based on the conclusion from chapter 3 that the Markov model performs better than the constant volume model, we extend the Markov model from a theoretical level to a practical level: chapter 4 assumes that information of the patient under prediction is not completely known, and that weekly MRI scans are available for all patients in the historical data. Therefore we estimate transition probability for the patient under prediction using the transition probabilities in the historical data. We classify tumors in the past into two groups: one representing the “fast-shrinking” group, and the other representing the “slow-shrinking/growing” group. Then we claim that each incoming patient behaves in a way as the combination of these two groups, i.e., the transition probability $\Pi_t$ for the incoming patient is a weighted average of $\Pi_{1,t}$ and $\Pi_{2,t}$ from the two groups. A few Markov models are developed in this chapter to estimate the transition probabilities $\Pi_{1,t}$ and $\Pi_{2,t}$, and the weight $\Theta_t$ associated with each group. While section Finally we compare the 6 Markov models (cf. Table 4.1) developed in this study and conclude the importance of an extra piece of information and of the good use of the information that are available in prediction of tumor geometry.

4.1 Extention of Markov Model I

In chapter 3 we develop a Markov model (denoted as Markov model I thereafter) that incorporates the transition probabilities obtained from individual MRI scans into prediction of the tumor evolution. This chapter considers several clinical scenarios when partial (but not all) information is observed for the patient under prediction. We assume that a database consisting of weekly MRI scans of patients in the past is available, and we will make use of this population information to estimate the transition probabilities of an incoming patient (i.e., a new patient under prediction). According to Figure 3.3, the rate at which a tumor shrinks is different among transitions. This motivates us to classify patients in the historical data set into groups using k-means clustering [36]. For simplicity, we start with two groups: group 1 consisting of tumors that shrink with a large rate, and group 2 consisting of tumors that shrink with either a small rate or a negative rate (grow). Note that the clustered group is assigned to 1 and
2 based on the mean value of the first column of the transition probability matrix, i.e., the transitions corresponding to the bi-voxel change from 1 to 0. In Figure 4.1, the $x$ axis is the mean value of the transition probabilities from 1 to 0 of 14 geometric configurations for a patient, and the $y$ axis is the volume change ratio $\frac{\text{vol}(t)}{\text{vol}(t+1)}$. The linear association indicates that the larger the mean value, the faster the tumor shrinks. Then we estimate the transition probability for each group. For an incoming patient, the transition probability is a weighted average of that of the two groups. Unlike chapter 3 where we simulate the tumor shape at $t + 1$ based on $t$, this chapter starts at the given MRI scan at $t = 2$, and the simulation for $t + 1$ is always based on the simulated tumor shape at $t$.

Formally, the steps to predict the evolution of tumor geometry for an incoming patient are provided below. The Markov models that will be discussed follow the same steps and step 3 to 5 will be elaborated.

1. Divide the 29 patients in the subject group into a training set and a testing set. The testing set consists of one patient, who represents the new incoming patient. And the training set consists of the rest 28 patients and replicates the database in reality. By assumption, weekly MRI scans are known for the training set but not for the testing set. Our goal is to write transition probability for the patient in the testing set at week $t$ as

$$\Pi_t = \Pi_{1,t} \cdot \theta_{1,t} + \Pi_{2,t} \cdot \theta_{2,t},$$

(4.1)
where $\Pi_{1,t}$ and $\Pi_{2,t}$ are the transition probabilities for the two groups clustered with the training set; $\Theta_t = [\theta_{1,t}, \theta_{2,t}]$ is the weight associated with the two groups and $\theta_{1,t} + \theta_{2,t} = 1$.

2. Estimate the transition probabilities for the patients in the training set using frequency count (see Equation 3.1). For patient $m$ in the training set at time $t$, the number of bi-voxels converts from state 1 to state 0, 1 or 2 is summarized in $Y_{m,t}$ below:

$$Y_{m,t} = \begin{bmatrix} y_0 | g_1 & y_1 | g_1 & y_2 | g_1 \\ \vdots & \vdots & \vdots \\ y_0 | g_k & y_1 | g_k & y_2 | g_k \\ \vdots & \vdots & \vdots \\ y_0 | g_{14} & y_1 | g_{14} & y_2 | g_{14} \end{bmatrix}_{m,t}$$  \hspace{1cm} (4.2)

where $y_k$ for $k \in \{0, 1, 2\}$ is the number of bi-voxels converting from state 1 to $k$, and $g_i \in g_1, \ldots, g_{14}$ is the given geometric configuration.

We estimate the transition probability $\Pi_{m,t}$ for patient $m$ at $t$ using frequency count, where

$$\Pi_{m,t} = \begin{bmatrix} p_0 | g_1 & p_1 | g_1 & p_2 | g_1 \\ \vdots & \vdots & \vdots \\ p_0 | g_k & p_1 | g_k & p_2 | g_k \\ \vdots & \vdots & \vdots \\ p_0 | g_{14} & p_1 | g_{14} & p_2 | g_{14} \end{bmatrix}_{m,t}$$  \hspace{1cm} (4.3)

where $p_0$, $p_1$ and $p_2$ are the transition probabilities that a bi-voxel converts from 1 to 0, 1 and 2, respectively.

Sometimes we are interested in the transition probabilities given a particular geometric configuration, and we denote such transition probability vector $V_{l,m,t}$ given geometry $g_l$ as below:

$$V_{l,m,t} = [p_0 | g_l \ p_1 | g_l \ p_2 | g_l]_{m,t},$$  \hspace{1cm} (4.4)

and the corresponding transition count vector $U_{l,m,t}$ is written as:

$$U_{l,m,t} = [y_0 | g_l \ y_1 | g_l \ y_2 | g_l]_{m,t}$$  \hspace{1cm} (4.5)

We also introduce the matrix operator $F$ of frequency count so that $\Pi_{m,t} = F \cdot Y_{m,t}$, where

$$F : \{p_k | g_l\}_{p \in \Pi_{m,t}} = \left\{ \frac{y_k | g_l}{\sum_{k \in \{0,1,2\}} y_k | g_l} \right\}_{y \in Y_{m,t}}, \forall g_l \in \{g_1, \ldots, g_{14}\}, k \in \{0, 1, 2\}.$$  \hspace{1cm} (4.6)

3. Cluster tumors in the training set into two groups $X_{1,t}$ and $X_{2,t}$ using k-means clustering. Since k-means clustering yields a local optimum in each iteration, we perform the clustering for 100 times to search for a global optimum.

4. Estimate the transition probability $\Pi_{1,t}$ and $\Pi_{2,t}$ for $X_{1,t}$ and $X_{2,t}$, respectively.
5. Estimate the weight of the assignment $\Theta_t = [\theta_{1,t}, \theta_{2,t}]$ to $X_{1,t}$ and $X_{2,t}$, respectively, for each patient in the testing set at $t$.

6. Obtain $\Pi_t$ for the incoming patient with Equation 4.1 and simulate the tumor shape at $t + 1$ using $\Pi_t$ based on the simulated tumor shape at $t$.

7. Obtain the transition probability map and the predicted tumor shape with the same approach described in section 3.1.1.

Along with the Markov model in chapter 3, we develop six Markov models and Table 4.1 summarizes the methodology for different steps within each model. The following subsections elaborate on the steps for each Markov model. In this section, we have a global assumption that weekly MRI scans are available for patients in the training set, while each model has different level of information from the testing set.

<table>
<thead>
<tr>
<th>Model</th>
<th>Data Source</th>
<th>Clustering</th>
<th>Transition Probability</th>
<th>Weight estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>MRI scans at t-1 and t</td>
<td>-</td>
<td>Frequency count</td>
<td>-</td>
</tr>
<tr>
<td>II</td>
<td>MRI scan at t=2</td>
<td>-</td>
<td>Weighted average for all patients geometry by geometry</td>
<td>Markov chain assumption on group association, with the first weight generated by voxel count and geometry map</td>
</tr>
<tr>
<td>III</td>
<td>MRI scans at t=2 in the historical data for all t</td>
<td>-</td>
<td>Linear association between the volume change ratio and the weight</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td>MRI scan at t=2, tumor volume at t-1 and t for all t</td>
<td>-</td>
<td>Weighted average of transition probability matrix</td>
<td>Generate two s-curves according to the means of the clusters</td>
</tr>
<tr>
<td>V</td>
<td>MRI scan at t=2</td>
<td>on transition probability vector for all patient-geometry pairs</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>VI</td>
<td>MRI scan at t=2, tumor volume at t-1 and t for all t</td>
<td>on transition probability vector geometry by geometry</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: Comparison of the models

### 4.2 The Markov Models with The First MRI Scan

#### 4.2.1 Markov model II: transition probabilities from MRI scans

In this section, we assume that the first MRI scan during the treatment ($t = 2$) is available for the patient in the testing set. We follow the steps listed in Section 4.1 and steps 3-5 are elaborated as follows:

3. At $t$ we perform k-means clustering on the $14 \times 28$ (number of geometries $\times$ number of patients in the training set) transition probability vectors $v_{1,m,t}$ (see Equation 4.4 for $m \in$ training set for all $g_t \in \{g_1, \ldots, g_{14}\}$). Figure 4.2 shows the two groups $X_{1,t}$ and $X_{2,t}$ as a result of clustering, denoted as red and blue circles. Each vertex of the simplex denotes a 100% transition probability for a bi-voxel to convert from 1 to 0, 1 or 2. The red circles represent $X_{1,t}$ (the “fast-shrinking” group) because the red region contains the probability that a bi-voxel converts from 1 to 0 being 100%. The darkness of the circle indicates the number of voxels having specific transition probability vector. E.g., the
4. We compute the transition probability $\Pi_{1,t}$ and $\Pi_{2,t}$ row by row as the weighted average of the transition probability vectors in $X_{1,t}$ and $X_{2,t}$, respectively. We write $\Pi_{i,t}$ for $i \in \{1, 2\}$ in the following form:

$$\Pi_{i,t} = \begin{bmatrix}
\hat{V}_{i,1,t} \\
\vdots \\
\hat{V}_{i,l,t} \\
\vdots \\
\hat{V}_{i,14,t}
\end{bmatrix} \quad \forall i \in \{1, 2\},$$

(4.7)
and the corresponding transition count matrix can be written as follows:

\[
Z_{i,t} = \begin{bmatrix}
\hat{U}_{i,1.t} \\
\vdots \\
\hat{U}_{i,l.t} \\
\vdots \\
\hat{U}_{i,14,t}
\end{bmatrix} \quad \forall i \in \{1, 2\}. \tag{4.8}
\]

For each geometric configuration \( g_i \in \{g_1, \ldots, g_{14}\} \),

\[
\hat{U}_{i,l.t} = \sum_{m \in X_i} U_{l,m,t} \quad \forall i \in \{1, 2\}. \tag{4.9}
\]

Then \( \Pi_{i,t} \) is obtained by \( F \cdot Z_{i,t} \) for \( i \in \{1, 2\} \). If all transitions associated with geometry \( g_i \) are clustered in \( X_i \), i.e., geometry \( g_i \) represents either a “fast” geometry or a “slow” geometry, the transition probability vector \( \hat{U}_{i,l,t} \) is estimated to be the mean of the vectors in such group, i.e.,

\[
\hat{U}_{i,l,t} = \sum_{k \in \{g_1, \ldots, g_{14}\}} \sum_{m \in X_i} U_{k,m,t} \quad \forall i = \{1, 2\}. \tag{4.10}
\]

5. At each \( t \), we construct a geometric configuration map. We count the number of bi-voxels \( r_{l,1} \) and \( r_{l,2} \) associated with \( g_i \) that are clustered in \( X_1 \) and \( X_2 \), respectively. For each \( g_i \), the probability of being associated with \( X_1 \) or \( X_2 \) are \( \gamma_{l,1,t} \) and \( \gamma_{l,2,t} \), where

\[
\gamma_{l,1,t} = \frac{r_{l,1}}{r_{l,1} + r_{l,2}}, \quad \gamma_{l,2,t} = 1 - \gamma_{l,1,t}. \tag{4.11}
\]

The map is written in matrix form as follows:

\[
\Gamma_t = \begin{bmatrix}
\gamma_{1,1,t} & \gamma_{1,2,t} \\
\vdots & \vdots \\
\gamma_{l,1,t} & \gamma_{l,2,t} \\
\vdots & \vdots \\
\gamma_{14,1,t} & \gamma_{14,2,t}
\end{bmatrix}. \tag{4.12}
\]

For each patient in the testing set, count the number of voxels associated with each geometry, and then normalize this vector to be

\[
\Omega_t = \begin{bmatrix}
\omega_{1,t} \\
\vdots \\
\omega_{l,t} \\
\vdots \\
\omega_{14,t}
\end{bmatrix}. \tag{4.13}
\]

The weight \( \Theta_t \) for the patient in the testing set is obtained from

\[
\Theta_t = \Omega_t^T \cdot \Gamma_t. \tag{4.14}
\]
Since we assume that the only available MRI scan for the patient in the testing set is the scan at $t = 2$, where we obtain $\Theta_2$. In order to estimate $\Theta_t$ for $t > 2$, we assume that the type of the patient (either 1 or 2) changes according to a (non-stationary) Markov chain, as shown below:

We can use Equation 4.15 to obtain $\Theta_t$ iteratively:

$$\Theta_t = \Theta_{t-1} \cdot \Psi_{t-1}$$

(4.15)

where

$$\Psi_t = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}_t.$$

Statistics indicate that the Markov chain is non-stationary, and therefore we estimate $\Psi_t$ for each $t$. We apply least square approach to the following linear system with $\tilde{\Theta}_{m,t}$ being the weight assigned to patient $m$ in the training set, which can be obtained with the same approach as we estimate $\Theta_t$ for the patient in the testing set.

$$\tilde{\Theta}_{m,t} = \tilde{\Theta}_{m,t-1} \cdot \Psi_{t-1}$$

(4.16)

### 4.2.2 Markov Model III: additional tumor volume from MRI scans

Model III shares the same assumption and same data source with Model II. In addition to the transition probabilities, we obtain the tumor volume from MRI scans in the training set. We implement this extra piece of information to step 3: the k-means clustering is performed on the following vector with an entry representing the volume change ratio:

$$\begin{bmatrix} p_{10} | g_t, p_{11} | g_t, p_{12} | g_t, \frac{\text{vol}_{t-1}}{\text{vol}_t} \end{bmatrix} \forall g_t \in \{g_1, ..., g_{14}\}$$

(4.17)

Step 4 and 5 are identical to Markov model II.

### 4.3 The Markov Models with Additional Volume Information

In addition to the first MRI scan during radiation treatment ($t = 2$), we assume the tumor volume (but not the MRI scans) at $t - 1$ and $t$ are known in the testing set at time $t$. Then we develop Markov models IV, V and VI that can take advantage of this extra piece of information to predict the tumor shape at time $t + 1$.

#### 4.3.1 Markov model IV: k-means clustering on patient-geometry pairs

Markov model IV shares the same steps 1 to 4 with Markov model II. To estimate the weight $\Theta_t$, since we obtain an extra piece of information, we analyze the relationship between $\Theta_t$ and the volume change ratio $R_t = \frac{\text{vol}(t-1)}{\text{vol}(t)}$. In step 5, model II and III iteratively obtain the weight $\Theta_t$ by assuming a Markov
chain on change of types for a patient, while this model uses a polynomial to obtain the \( \Theta_t \) given the volume change ratio. Figure 4.3 illustrates a linear relationship between \( \Theta_t \) on the y axis and \( R_t \) on the x axis. We may fit the data points with other well-known functions and the performance may improve.

Figure 4.3: Linear association between the tumor volume ratio \( R_t \) (on the x axis) and the group association \( \theta_{1,t} \) (on the y axis). Note that the weight \( \Theta_t = [\theta_{1,t}, \theta_{2,t}] \).

4.3.2 Markov model V: k-means clustering on patients for each geometry

Model IV clusters tumors with transition probability vectors mixing patients and geometries together, this model clusters tumors for all patients geometry by geometry. Formally we write steps 3 to 5 as follows:

3. For geometry \( g_l \), at time \( t \), we perform k-means clustering on the 28 (number of patients in the training set) transition probability vectors \( U_{l,m,t} \). As a result, we obtain group 1 \( X_{1,g_l} \) and group 2 \( X_{2,g_l} \) for each geometric configuration \( g_l \). The clustering process is similar to step 3 in model II and III: while model II yields results on one simplex (see Figure 4.2), this model yields results on 14 simplexes corresponding to 14 geometric configuration.

4. Since we perform clustering analysis for each geometry, we estimate the transition probability \( \Pi_{1,t} \) and \( \Pi_{2,t} \) row by row, i.e., we estimate the transition probability vector \( \hat{V}_{i,t} \) for \( i \in \{1, 2\}, l \in \)
{1, ..., 14}. Formally, we write down the estimation as follows:

$$\hat{V}_{i, l, t} = F \cdot \hat{U}_{i, l, t} \forall i \in \{1, 2\}, l \in \{1, ..., 14\},$$

(4.18)

where

$$\hat{U}_{i, l, t} = \sum_{m \in X_{i, l}} U_{l, m, t} \forall i \in \{1, 2\}, l \in \{1, ..., 14\}.$$  

(4.19)

We perform k-means clustering geometry by geometry for 14 geometries and Figure 4.4 shows the transition probability distribution of the two resulting clusters, with red representing $X_{1, l}$ and blue $X_{2, l}$ for the 4 most frequent geometric configurations. Each row represents a geometry $g_l$ and the three columns represent that a bi-voxel converts from 1 to 0, 1, and 2, respectively. The red bars indicate transition probability vectors associated with group 1.

![Figure 4.4: The distribution of transition probability vectors for 4 characterized geometries for all patients during one week. Each row represents a geometric configuration, and the three columns represent transition of a bi-voxel from state 1 to 0, 1 and 2, respectively. The red bars indicate transition probability vectors associated with group 1.](image-url)

Note that the two clusters are distinguished for the transitions 1 → 0 and 1 → 1. However, the probability tends to converge to 0 for both clusters for the transition 1 → 2, i.e., tumor growth are rare for all geometries. The distribution takes into account the frequency of occurrence of a transition probability, i.e., $\frac{1}{10}$ and $\frac{10}{1000}$ share the same probability 0.1 but the latter contributes 10 times that of the former.
5. Now that we have 14 sets of clusters $X_{1,l}$ and $X_{2,l}$ for $l \in \{1, \ldots, 14\}$, we need to integrate the group association with each geometry into a single weight $s_{m,t}$, indicating the likelihood that patient $m$ belongs to $X_1$ at time $t$ (i.e., the likelihood that patient $m$ belongs to $X_2$ at time $t$ is $1 - s_{m,t}$). Formally, we give the following definition:

$$s_{m,t} = \frac{\sum_{i=1}^{q} 1_{i \in X_1} f_m(g_l)}{\sum_{i=1}^{q} f_m(g_l)}, \quad (4.20)$$

where $f(g_l)$ is the number of voxels associated with geometric configuration $g_l$ for patient $m$. In this context $i \in 1, \ldots, q$ means that we only consider the $q$ most frequent geometries. Figure 4.5 shows the computation of the likelihood $s_{m,t}$ for patient $m$ at $t$ with $q = 4$. Each row represents a geometry and each column represents a patient. The red and blue dots represent patient-geometry pairs that belong to $X_1$ and $X_2$, respectively. The likelihood $s_{m,t}$ is calculated approximately as the

Figure 4.5: Group association for all patients during one week with $q = 4$ characteristic geometries. Each row represents a geometric configuration, and each column represents a patient. A red dot shows that from time $t$ to $t+1$, the patient belongs to $X_1$ for a particular geometric configuration. It is ideal that a patient consistently belongs to the same group over all the geometric configuration. The likelihood $s_{m,t}$ that the patient is associated with $X_1$ (red) conceptually estimates the number of red dots each column has.
number of red dots divided by 4 (total number of dots) in each column. However, the weight of the 4 dots differs according to the number of voxels associated with them in the formulation. Table 4.2 shows the likelihood $$s_{m,t}$$ that patient $$m$$ belongs to $$X_1$$ from $$t$$ to $$t+1$$ with $$q = 4$$. Like in

Table 4.2: Likelihood $$s_{m,t}$$ that the patient belongs to $$X_1$$

<table>
<thead>
<tr>
<th></th>
<th>$$t = 3$$</th>
<th>$$t = 4$$</th>
<th>$$t = 5$$</th>
<th>$$t = 6$$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.00</td>
<td>0.58</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>0.66</td>
<td>0.92</td>
</tr>
<tr>
<td>3</td>
<td>1.00</td>
<td>1.00</td>
<td>0.22</td>
<td>0.49</td>
</tr>
<tr>
<td>4</td>
<td>0.39</td>
<td>0.16</td>
<td>0.39</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>0.00</td>
<td>0.41</td>
<td>0.58</td>
<td>1.00</td>
</tr>
<tr>
<td>6</td>
<td>0.14</td>
<td>0.83</td>
<td>0.00</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>0.66</td>
</tr>
<tr>
<td>8</td>
<td>0.00</td>
<td>0.69</td>
<td>0.25</td>
<td>0.78</td>
</tr>
<tr>
<td>9</td>
<td>1.00</td>
<td>0.69</td>
<td>0.92</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>0.17</td>
<td>-</td>
<td>-</td>
<td>0.16</td>
</tr>
<tr>
<td>11</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>-</td>
</tr>
<tr>
<td>12</td>
<td>0.34</td>
<td>0.83</td>
<td>0.86</td>
<td>0.14</td>
</tr>
<tr>
<td>13</td>
<td>0.67</td>
<td>0.00</td>
<td>0.86</td>
<td>0.47</td>
</tr>
<tr>
<td>14</td>
<td>-</td>
<td>-</td>
<td>0.16</td>
<td>0.66</td>
</tr>
<tr>
<td>15</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>16</td>
<td>-</td>
<td>-</td>
<td>0.84</td>
<td>0.83</td>
</tr>
<tr>
<td>17</td>
<td>0.08</td>
<td>0.08</td>
<td>0.78</td>
<td>-</td>
</tr>
<tr>
<td>18</td>
<td>0.00</td>
<td>1.00</td>
<td>0.08</td>
<td>0.22</td>
</tr>
<tr>
<td>19</td>
<td>-</td>
<td>-</td>
<td>0.00</td>
<td>0.65</td>
</tr>
<tr>
<td>20</td>
<td>1.00</td>
<td>-</td>
<td>-</td>
<td>0.51</td>
</tr>
<tr>
<td>21</td>
<td>0.00</td>
<td>0.00</td>
<td>0.08</td>
<td>0.68</td>
</tr>
<tr>
<td>22</td>
<td>-</td>
<td>-</td>
<td>0.08</td>
<td>0.51</td>
</tr>
<tr>
<td>23</td>
<td>-</td>
<td>1.00</td>
<td>1.00</td>
<td>0.81</td>
</tr>
<tr>
<td>24</td>
<td>0.34</td>
<td>0.92</td>
<td>0.16</td>
<td>0.14</td>
</tr>
<tr>
<td>25</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>26</td>
<td>0.69</td>
<td>0.75</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>27</td>
<td>0.35</td>
<td>1.00</td>
<td>1.00</td>
<td>0.22</td>
</tr>
<tr>
<td>28</td>
<td>1.00</td>
<td>1.00</td>
<td>0.78</td>
<td>-</td>
</tr>
</tbody>
</table>

model IV, we want to relate $$s_{m,t}$$ with the volume change ratio. We improve the model fitting by using $$V_t$$ rather than $$R_t$$, which involves both the change in tumor volume from $$t-1$$ to $$t$$ and the actual tumor volume at $$t$$:

$$V_t = \frac{\text{vol}(t-1)}{\text{vol}(t)} \cdot \log[\text{vol}(t-1)]$$ (4.21)

where $$\text{vol}(t)$$ is the tumor volume at time $$t$$.

We define the $$q$$ ($$q \leq 14$$) most frequent geometric configurations to be the characteristic geometries, which means that the likelihood $$s_{m,t}$$ integrated from these $$q$$ geometries are as good as if integrated from all 14 geometries. In section 2.3.2, a 1% cut-off is used and any geometric configuration that has less than 414.41 (41,441×1%) bi-voxels is not considered in the study, and the resulting 14 geometries contribute to 96.58% of the 41,441 bi-voxels. A 5% cut-off results in 6 geometries that contribute to 75.41%. A 8% cut-off results in 4 geometries that contribute to 63.63%. We compare the regression between $$V_t$$ (on the y axis) and the likelihood $$s_{m,t}$$ of being in $$X_1$$ (on the x axis) with $$q = 4$$, $$q = 6$$ and $$q = 14$$ in Figure 4.6a, Figure 4.6b and Figure 4.6c respectively. Inside each
Figure 4.6: Regression between group association $s_{m,t}$ and volume change ratio $V_t$ with different number $q$ of characteristic geometries.

(a) $q = 4$

(b) $q = 6$

(c) $q = 14$
subfigure the subgraphs show the regression from time $t - 1$ to $t$. Given $V_t$ of an incoming patient, $s_{m,t}$ can be obtained from the linear relationships.

The regression line obtained are similar for $q = 4$ and $q = 6$. Since the $s_{m,t}$ values are estimated as an average weighted according to the occurrence of a geometry (see Equation 4.20), meaning that the geometry with fewer occurrence contributes less in terms of clustering. However, there is a large difference in the obtained regression line between $q = 14$ and the other cases, implying that 4 or 6 geometries are not sufficient to characterize the clustering process. Therefore we use 14 geometries in the study and the result is presented in the next section.

#### 4.3.3 Markov model VI: biclustering on patients

Unlike other models performing k-means clustering on vectors consisting of transition probabilities, this model performs a biclustering [37] on the volume ratio $R_t = \frac{\text{Vol}(t)}{\text{Vol}(t-1)}$. Steps 2 to 5 are elaborated as follows:

3. We perform biclustering using the volume ratios $R_{m,t}$ for the patient $m$ in the training set. Then we obtain two clusters $X_1$ and $X_2$ with mean $x_1$ and $x_2$, respectively:

$$x_i = \frac{\sum_{m \in X_i} R_{m,t}}{\sum_{m} 1_{m \in X_i}}, \forall i \in \{1, 2\}.$$  \hspace{1cm} (4.22)

4. We take a weighted average of the transition probability from patients in $X_1$ and $X_2$, respectively:

$$\Pi_{i,t} = F \cdot W_{i,t}, \forall i \in \{1, 2\},$$  \hspace{1cm} (4.23)

where

$$W_{i,t} = \sum_{m \in X_i} Y_{m,t}, \forall i \in \{1, 2\},$$  \hspace{1cm} (4.24)

and $Y_{m,t}$ is the transition count (see Equation 4.2) for patient $m$ in the training set.

5. Given the volume ratio $x^*$ of the patient in the testing set, we obtain the weight $\Theta_t$ using the following equation [38]:

$$\theta_{1,t} = \begin{cases} 
\max(1 - \lambda, \lambda), & \text{if } |x^* - x_1| < |x^* - x_2|, \\
\min(1 - \lambda, \lambda), & \text{otherwise}. \end{cases}$$

where

$$\lambda = \frac{1}{1 + e^{-\alpha_6 |x^* - q|}}, \quad q = \frac{x_1 + x_2}{2},$$

$\alpha_6$ is a tunable parameter for the equation. It is obtained by assuming that the mean $x_1$ of the volume ratios in $X_1$ locates at the $z^{th}$ percentile of the distribution of all volume ratios (in both $X_1$ and $X_2$). It is optimized through enumeration of $z$ value ranging from 0.05 to 0.40 with an increment of 0.05. Figure 4.7 plots the Markov curves with different choice of $z$. We compare the choices quantitatively by computing the directional Hausdorff distance (see Equation 3.12) between the Markov curves generated by Model V with different $z$ values and a benchmark curve close to the ideal point (1,1), shown as the red line in Figure 4.7. The smaller the distance, the better the performance. Table 4.3 shows that the best performance is obtained at $z = 0.25$, i.e., the
Table 4.3: Hausdorff distance between Markov curves over a range of \( z \) values and a standard curve. The smallest distance (red curve) indicates the best performance.

<table>
<thead>
<tr>
<th>( x^{th} ) Percentile of all data</th>
<th>0.40</th>
<th>0.35</th>
<th>0.30</th>
<th>0.25</th>
<th>0.20</th>
<th>0.15</th>
<th>0.10</th>
<th>0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.1868</td>
<td>0.1888</td>
<td>0.1872</td>
<td>0.1786</td>
<td>0.1872</td>
<td>0.1874</td>
<td>0.1885</td>
<td>0.1885</td>
</tr>
</tbody>
</table>

model performs best when we assume the mean tumor change ratio is at 25\% percentile of the data distribution. Unless otherwise stated, the Markov curve for Model V presented in the following section uses the parameter \( z=0.25 \).

Figure 4.7: Markov curves generated for Markov model V over a range of \( z \) values.

4.4 Results and discussion

Figure 4.8 compares the sensitivity and specificity among the curves generated by the six developed Markov models and the constant volume model over a range of \( \alpha \) in terms of mean values (see Figure 4.8a) and a combination of 10\% percentile sensitivity and specificity (see Figure 4.8b). The deviation between the performance of the models becomes more apparent when we look at the 10\% percentile (i.e., the worst case performance). It indicates that certain Markov model performs better than others with the most unpredictable patients. Similar to the observations from last chapter: there is an apparent trade-off between sensitivity and specificity. The Markov model performs slightly better than the constant volume model with respect to the mean. However, the deviation becomes obvious when we investigate the 10\% percentile values. Table 4.4 and Table 4.5 show the mean value curves and a quantitative comparison at \( \alpha = 0 \) and \( \alpha = 0.3 \), respectively. The tables start with an illustrative figure consisting of the Markov curves and the constant model. For each transition, we compute the Hausdorff distance between the Markov curve generated from the individual patient and a benchmark curve. Part a computes the mean...
Figure 4.8: Comparison of the models

(a) Mean performance

(b) 10th percentile performance
of the Hausdorff distance of all transitions. Part b ranks the models according to the mean distance presented in a. The following sections compare the different Markov models.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|}
\hline
\textbf{Markov Model} & \textbf{I} & \textbf{II} & \textbf{III} & \textbf{IV} & \textbf{V} \\
\hline
\textbf{\(\Delta H\)} & 0.0383 & 0.0237 & 0.0245 & 0.0239 & 0.0607 \\
\hline
\end{tabular}
\caption{Comparison between the mean sensitivity and specificity of the Markov models at \(\alpha = 0.4\).}
\end{table}

\subsection{4.4.1 Choice of a benchmark curve}

An acceptable region was defined in section 3.2 to evaluate the performance and the boundary of the region is subject to change according to the clinical requirement. In this chapter we compute Hausdorff distance to further quantify the model performance. The smaller the distance between a Markov curve and a benchmark curve, the better the performance. We introduce a few benchmark curves so that
Table 4.5: Comparison between the mean sensitivity and specificity of the Markov models at $\alpha = 0.3$. 

<table>
<thead>
<tr>
<th>$\Delta H$</th>
<th>Benchmark Curves</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Markov Model I</td>
<td></td>
<td>0.0198</td>
<td>0.0732</td>
<td>0.0172</td>
<td>0.0190</td>
<td>0.0174</td>
</tr>
<tr>
<td>Markov Model II</td>
<td></td>
<td>0.0280</td>
<td>0.0710</td>
<td>0.0498</td>
<td>0.0443</td>
<td>0.0542</td>
</tr>
<tr>
<td>Markov Model III</td>
<td></td>
<td>0.0270</td>
<td>0.0702</td>
<td>0.0452</td>
<td>0.0406</td>
<td>0.0494</td>
</tr>
<tr>
<td>Markov Model IV</td>
<td></td>
<td>0.0324</td>
<td>0.0830</td>
<td>0.0488</td>
<td>0.0439</td>
<td>0.0526</td>
</tr>
<tr>
<td>Markov Model V</td>
<td></td>
<td>0.0251</td>
<td>0.0483</td>
<td>0.0177</td>
<td>0.0214</td>
<td>0.0179</td>
</tr>
<tr>
<td>Markov Model VI</td>
<td></td>
<td>0.0232</td>
<td>0.0687</td>
<td>0.0407</td>
<td>0.0366</td>
<td>0.0442</td>
</tr>
</tbody>
</table>

(a) Hausdorff distance

(b) Rank
both the sensitivity and specificity range from 0.93 to 1 (within the acceptable region that we defined in section 3.2), and it is determined by the clinicians in real practice. The figure in Table 4.4 shows five benchmark curves as examples, and each curve puts emphasize on a different perspective. Benchmark curve A is a straight horizontal line with sensitivity equal to 1 over a range of specificity from 0.93 to 1. Benchmark curve B is a straight vertical line with specificity equal to 1 over a range of sensitivity from 0.93 to 1. While A shares the same emphasis with the current clinical practice on a complete tumor coverage, B puts emphasis completely on OAR sparing. Benchmark curves C and D take both objectives into account. The angle between the benchmark curve and x axis indicates the degree of conservativeness (the more emphasis on tumor coverage, the more conservative is a treatment plan). The larger the degree, the less conservative the treatment plan will be. We can either include or not include a change of degree of conservativeness in a benchmark curve. The point (x, y) where the change occurs means that if the OAR sparing can reach as much as x, then we can sacrifice more sensitivity than before in exchange for more OAR sparing.

The performance among models largely depends on the choice of the benchmark curve. Table 4.4b shows the rank of the mean value Markov curves at $\alpha = 0$. Markov model II performs best with benchmark curve A, while Markov model VI performs best with benchmark curve B. The rank is identical with benchmark curve C and D, with model I ranking the first and V being the worst. Benchmark curve E has the largest span among the benchmark curves, i.e., it adds more constraints to the max-min problem to find the Hausdorff distance. Therefore a completely different rank is shown with curve E. The choice of the benchmark curve needs to be discussed with clinicians in real practice.

4.4.2 Importance of individual information: model I vs. others

We examine the performance of the developed Markov models over time $t$ (see Figure 4.9). The spread among curves of Markov model I is larger than that of the other models. This implies that Markov model I improves over time (see Figure 4.9a) while the others perform stably. Since model I uses the actual tumor shape at $t$ as the base of prediction at $t + 1$, it incorporates the most updated information of the tumor, and therefore the goodness of performance at each $t$ is independent. The performance becomes slightly worse towards the end of treatment for other models. Unlike the weekly updated actual tumor shape we obtain in model I, since the prediction at $t + 1$ is based on the predicted tumor shape at $t$, the deviation from the actual tumor shape at $t$ will influence the goodness of prediction throughout the treatment.

We also compare model I and others in terms of sensitivity and specificity with some margin. Table 4.5a shows the Hausdorff distance between the benchmark curves and the mean value curves generated from Markov model with a margin $\alpha = 0.3$. Accordingly, Table 4.5b shows a more consistent rank among different benchmark curves: Markov model I ranks the first unless we put a complete emphasis on specificity (benchmark curve B). We infer that the most updated and accurate information leads to a less conservative treatment plan. When we add some margin, the prediction becomes more conservative and exceeds other models in terms of sensitivity.

4.4.3 Importance of volume information: model I vs. V/VI

Unlike Markov model I that uses incoming patient’s MRI scans at $t - 1$ and $t$ to predict the tumor shape at $t + 1$, model V and VI use the tumor volume at $t - 1$ and $t$ of an incoming patient and takes advantage of the known weekly MRI scans in the historical data. However, they are comparable in Table 4.4b and
Figure 4.9: Sensitivity vs. specificity over $t = 4$ to $t = 6$
Table 4.5a This implies that the two levels of information mentioned above are comparable in terms of contribution to the goodness of tumor prediction. Note that the cost for MRI scans from the past is considered sunk cost, and therefore it is not included in the cost for developing a treatment plan for an incoming patient. Since we can potentially obtain volume information in a less expensive way than obtaining MRI scans, the cost of information from model VI or V is smaller than model I. Therefore, the former outperforms the latter since it is more cost-effective. In addition, the curves generated by Markov model V and VI span over a larger range than model I, and its tunable probability threshold $\beta$ allows it to obtain more combinations of sensitivity and specificity than model I. Therefore it is more adaptive to the clinical requirement. Furthermore, when we add some margin to our prediction, model V maintains a good balance between achieving both large sensitivity and large specificity: it ranks the first with OAR sparing as our target (benchmark curve B), and ranks the second or the third when the objective shifts towards specificity.

### 4.4.4 Information utilization: model II vs. III

Model II and III both use only the population data. Unlike the discussion in the previous subsection, there is no additional cost of the information. The mere difference between the two approaches is in the amount of information they obtain: model III incorporates one more piece of information, i.e., volume change ratio, into the prediction. Table 4.4a shows that there is at most 0.1% difference between these two models, but the better model varies according to the choice of a benchmark curve. When we put more emphasis on sensitivity (e.g., benchmark curve A and E), Markov model II exceeds III. However, when we put more emphasis on specificity (e.g., benchmark curve B, C and D), Markov model III outperforms II. This implies that because of the additional piece of information, Markov model III becomes less conservative in terms of tumor prediction. Therefore, we need to decide which information to include in the prediction according to a clinical objective.

### 4.4.5 Clustering method

In the study, we classify tumors in the training set into two groups with three different methods. Each patient $m$ in the training set has a transition probability matrix $\Pi_{m,t}$ (see Equation 2.13) consisting of 14 rows $\hat{U}_{m,l,t}$ (see Equation 4.4) corresponding to 14 geometric configurations $g_l \in \{g_1, ..., g_{14}\}$. Markov model II, III and IV adopt the first method and perform the clustering on $\hat{U}_{m,l,t}$ for all patient-geometry pairs; Markov model VI uses the second method and performs the clustering on $\hat{U}_{m,l,t}$ for all patients geometry by geometry; Markov model VI uses the third method with a single indicator (i.e., change in tumor volume) for the clustering.

For an incoming patient, the first method estimates the weight $\theta_{1,t}$ associated with $X_1$ to be closer to 0.5, and therefore the behavior of the incoming patient is close to the mean behavior of the patients in the historical data. This method diminishes the difference among newly observed extreme cases. However, it is justifiable that without any updated information, the approach assigns each new patient similar transition probabilities with adjustment in small degree. It can produce a very conservative treatment plan that more likely to cover the tumor and some OARs around it.

On the other hand, the second method pushes the weight towards the end value 0 or 1. Compared to the first method, it allows a larger range for the transition probabilities of an incoming patient. The first method clusters the transition probability vectors for all geometries all at once, classifying certain geometries as “fast geometries” while some as “slow geometries”, i.e., for fast geometries, the values that
a new vector can take concentrate around certain probabilities. However, the second method performs clustering for each geometry so that each has its own “fast probability” and “slow probability”, and therefore a new vector has a larger range of values to take. In other words, the method pushes the weight $\Theta_t$ towards 0 or 1 and the resulting transition probability is closer to two extremes of the two existing groups.

Furthermore, the third method uses Equation 2.6 to generate a weight for an incoming patient, given the tumor volume ratio. Since the new volume ratio may fall out of the existing range, the formulation drives the weight $\Theta_t$ even more towards the two extremes than the second method. Therefore the incoming patient most of the time belongs to a group (either $X_1$ or $X_2$), rather than standing in the middle of the two groups and sharing the same property with both groups. The effect of tumor clustering in step 3 in section 4.1 is more remarkable in model VI than other models with clustering process.

### 4.4.6 Improvement on poorly behaved patients

Previous subsections justify model performance by comparing models in terms of the mean sensitivity and specificity. Table 4.6 compares model performance for the least 10% well behaved patients. With an equal amount of improvement in the mean performance, we can either improve enormously on the poorly behaved and rarely on the well behaved, or improve moderately on both the poorly and well behaved patients. In clinics, however, the former improvement is preferred. Without a margin (i.e., $\alpha = 0$), model VI performs steadily well with all benchmark curves: it ranks first with both curve A and C. Though model I ranks high when we compare the mean values, it becomes less competitive considering the improvement on the poorly behaved patients, with model I ranking within last three with three benchmark curves. Markov model VI uses tumor volume at $t − 1$ and $t$ from the incoming patient, and the weekly MRI scans from the historical data. It implies that a considerable amount of information from the poorly behaved patients can be helpful in terms of prediction, but more information may become a distractive factor, as model I sometimes ranks next to last. It is justifiable that a poorly behaved patient does not have a consistent behavior over time, and therefore information from $t − 1$ to $t$ for this patient may not be better than information from $t$ to $t + 1$ from patients in the historical data.

When we consider a margin added to the prediction as shown in Table 4.7, the model performance is similar to that with the mean values, since adding a margin diminishes the difference among models. This can also be seen from the deviation among the curves with (see Table 4.4 and Table 4.5) or without (see Table 4.5 and Table 4.7) a margin.

### 4.4.7 The tunable parameters

Instead of computing the mean of Hausdorff distance for a fixed margin $\alpha$ (as in Table 4.4), we calculate the Hausdorff distance between a benchmark curve and a set of Markov curves (corresponding to different margin size $\alpha$) generated by one Markov model. Unlike the Hausdorff distance in the previous tables reflecting the performance of the model with a single tunable parameters $\beta$, the Hausdorff distance in this setting reflects the performance with an additional parameters $\alpha$.

Figure 4.10 shows the variance of each model. The $x$ axis represents the six developed models and the constant volume model (rightmost). The $y$ axis indicates the Hausdorff distance: the smaller the distance, the better the model performance. For each model, the figure shows three lines corresponding to the Hausdorff distance between the Markov model and three benchmark curves A, B and C. The top and bottom of each line indicates the largest and the smallest distance, respectively, among all patients all
Table 4.6: Comparison between the 10th percentile sensitivity and specificity of the Markov models at $\alpha = 0$. 

(a) Hausdorff distance

(b) Rank
Table 4.7: Comparison between the $10^{th}$ percentile sensitivity and specificity of the Markov models at $\alpha = 0.3$. 

(a) Hausdorff distance

<table>
<thead>
<tr>
<th>$\Delta_H$</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Markov Model I</td>
<td>0.0364</td>
<td>0.0965</td>
<td>0.0366</td>
<td>0.0366</td>
<td>0.0366</td>
</tr>
<tr>
<td>Markov Model II</td>
<td>0.0564</td>
<td>0.0919</td>
<td>0.0790</td>
<td>0.0620</td>
<td>0.0945</td>
</tr>
<tr>
<td>Markov Model III</td>
<td>0.0515</td>
<td>0.0891</td>
<td>0.0731</td>
<td>0.0592</td>
<td>0.0920</td>
</tr>
<tr>
<td>Markov Model IV</td>
<td>0.0533</td>
<td>0.1084</td>
<td>0.0801</td>
<td>0.0653</td>
<td>0.1004</td>
</tr>
<tr>
<td>Markov Model V</td>
<td>0.0408</td>
<td>0.0641</td>
<td>0.0323</td>
<td>0.0395</td>
<td>0.0336</td>
</tr>
<tr>
<td>Markov Model VI</td>
<td>0.0460</td>
<td>0.0865</td>
<td>0.0700</td>
<td>0.0535</td>
<td>0.0842</td>
</tr>
</tbody>
</table>

(b) Rank

<table>
<thead>
<tr>
<th>Rank</th>
<th>Benchmark Curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I V V I V</td>
</tr>
<tr>
<td>2</td>
<td>V VI I V</td>
</tr>
<tr>
<td>3</td>
<td>VI III VI VI VI</td>
</tr>
<tr>
<td>4</td>
<td>III II III III</td>
</tr>
<tr>
<td>5</td>
<td>IV I II II</td>
</tr>
<tr>
<td>6</td>
<td>II IV IV IV IV</td>
</tr>
</tbody>
</table>
weeks. The circle and the square show the mean value and 10\textsuperscript{th} percentile of the distance of all patients all weeks.

When we put a complete emphasis on sensitivity (see benchmark curve A with black lines), the mean performance (the black circles) of the models are similar. The constance volume model is the worst when we look at the most unpredictable patients (the blue circles), both with A and B, where we compare the models with respect to specificity (see benchmark curve B with blue lines). Model V does not perform as well as others when a small margin is applied (see Table 4.4 and Table 4.6). However, since the Markov model can take advantage of tunable parameters of both $\alpha$ and $\beta$, the large range of values that Model V produce makes it outperform. Benchmark curve C (the red lines) incorporates a balanced evaluation over sensitivity and specificity, under which Markov model I outperforms others in terms of both the mean and 10\textsuperscript{th} percentile values.

The distance between the maximum of Hausdorff distance and the 10\textsuperscript{th} percentile value reflects the performance of the most unpredictable patients. We observe that model I, IV and V have the largest ranges. This implies that when we are provided with information from individual patient, if she is a well behaved patient, then the tumor evolution from $t$ to $t+1$ follows that from $t-1$ to $t$; otherwise, the behavior is random and is more likely to follow the mean behavior of other patients in the past.

The distance between the mean and the 10\textsuperscript{th} percentile value implies whether the model is stable among patients (for both well behaved and poorly behaved). The performances from the Markov models are relatively stable because of a continuous parameter $\beta$, while the constant model can only take discrete values with various $\alpha$. 

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.10}
\caption{Comparison among Markov models}
\end{figure}
Chapter 5

Conclusions and Future Work

Adaptive radiation therapy is a powerful concept in the ongoing battle against cancer. It has been made even more robust with the advent of image guided radiation therapy and intensity modulated radiation therapy. The main objective of this work is to develop a mathematical model to predict the cervical tumor’s evolution during radiation therapy. This model can be extended to the prediction of the evolution of other components in CTV, and hence assist the planning before or during the treatment, depending on the availability of information.

5.1 The Markov assumption

Chapter 3 assumes that a cervical tumor’s evolution during radiation therapy follows a Markov process and develops a Markov model accordingly. We use the scans from \( t - 1 \) and \( t \) to predict the tumor shape at \( t + 1 \), and compare the simulated shape and its actual counterpart. The result indicates that the Markov model performs better than the constant volume model that replicates the method in current clinical practice.

However, we did not formally validate the hypothesis that the evolution follows a Markov process. In order to validate the hypothesis, we need to show that the state of the voxels on the outermost layer of the tumor at time \( t \) will change from state 1 to 0 at time \( t + 1 \) under certain circumstances, regardless of their state at time \( t - 1 \). This requires us to track the exact voxels on the tumor surface. However, since we used a static voxel spatial system, it is difficult to track where the cells on the tumor surface at time \( t - 1 \) move to at time \( t \).

Our study uses weekly MRI scans, which makes the influence from the radiation almost a dominant factor to the tumor shrinkage. Therefore we can justify our assumption that the future state of voxels only depends on the current state of voxels because of such dominating factor in a week. However, if future study gathers more frequent MRI scans, e.g., daily MRI scans, then the assumption may or may not hold.

A model always simplifies the real case, and this study chooses a simple stochastic model to start with. The performance of the developed Markov models indicates that the model functions well and is better than the constant volume model. With the current dataset it is not possible to rigorously validate the assumption that the evolution of a cervical tumor follows a Markov process. An anatomical voxel grid that can capture motion rather than a static voxel grid is needed.
5.2 Value and cost of information

Chapter 3 uses weekly MRI scans to show that the model functions well under the Markov assumption. However, in clinical practice the MRI scans are usually unavailable. Chapter 4 explores the scenarios when only the first MRI scan is available. We also examine the possible improvement we can make with more information that are less expensive, e.g., the tumor volume.

When there is less information available from an incoming patient, we investigate the tumor evolution of the patients in the historical data and draw relationship between the old and new data. Inspired by the classic concept of radio-resistant and radio-responsive, we classified the tumors in the historical data into “fast-shrinking” and “slow-shrinking/growing” groups. Then we observed similarity between the new patient and the existing groups. Further study may explore the possibility of having more groups, which may improve the classification and transition probability estimation that follows.

Chapter 4 shows that the correct use of information that is partially observed may result in a model performance as good as when complete information is known. The Markov model that uses only the an incoming patient’s first MRI scan with updated tumor volume information can be comparable to the model that implements the incoming patient’s weekly MRI scans. Apparently the latter is more expensive than the former, and the former one outperforms because of its cost-effectiveness. However, if the only available information is the first week’s scan of an incoming patient, the Markov model becomes conservative and assumes that the tumor evolves in a similar manner as an average of all tumors in the historical data.

In clinical practice, usually we have the first week’s MRI scan. The tumor volume information is rarely obtained. Further study can perform a sensitivity analysis to explore the influence of a small perturbation from the volume information on the performance of model IV, V and VI. In addition, we can also explore more metrics that can define the tumor shrinkage (other than $V_t$ and $R_t$ presented in this paper).

5.3 Evaluation and performance metrics

We discard the $F$ score in chapter 3, because it assumes an equal importance in tumor coverage and OAR sparing. Instead of evaluating the model performance with a single number, we present both sensitivity and specificity, and therefore the clinicians can decide which model to use based on their own judgement. While chapter 3 defines a “acceptable region” where the clinicians can set their own boundary, chapter 4 introduces the benchmark curves as part of the evaluation process. Since the Hausdorff distance, which can evaluate the model performance, depends on the choice of the benchmark curve, the decision is made by the clinicians eventually.

5.4 Dose-guided radiation therapy

An interesting byproduct of the Markov model is the probability map, which estimates the likelihood of a particular voxel being tumor. This probability map might find application in optimization approaches where voxel-specific doses or weighting factors are used to plan and adapt treatments [39, 40, 41, 42].

This study assumes that the tumor receives uniform dose during the treatment. In current clinical practice, IMRT enables each voxel in the tumor to receive a different dosage by assigning various dosage to beams. The computation of the dose distribution is based on the tumor geometry obtained from the
weekly MRI scans. The predicted tumor geometry from this study can be used as an input of the dose distribution computation, where we can analyze the dosimetric effect and re-optimize the treatment plan. In another application area, we can increase the dose weight associated with a voxel in the tumor if the probability that certain voxel being tumorous is larger than its nearby voxels.

In summary, this study constructs mathematical models to predict the evolution of tumor geometry in cervical cancer during radiation therapy, defines new metrics to measure model performance, and provides insights over the value of acquiring information.
Bibliography


