

Statistics-Based Prediction Analysis for Head and Neck Cancer Tumor Deformation

Maryam Azimi, Ph.D¹, Ali K. Kamrani, Ph.D. P.E.^{1,2,3*} and Hazem J. Smadi, Ph.D⁴

¹Lenovo Corporation, Morrisville, NC, USA

²Design and Free Form Fabrication Laboratory,
University of Houston, Houston, TX, USA

³FARCAMT, Industrial Engineering Department, College of Engineering,
King Saud University, Riyadh, Saudi Arabia

⁴Industrial Engineering Department, Jordan University of Science and Technology,
Irbid, Jordan.

Submitted December 2011. Accepted for publication July, 2012.

ABSTRACT

Most of the current radiation therapy planning systems, which are based on pre-treatment Computer Tomography (CT) images, assume that the tumor geometry does not change during the course of treatment. However, tumor geometry is shown to be changing over time. We propose a methodology to monitor and predict daily size changes of head and neck cancer tumors during the entire radiation therapy period. Using collected patients' CT scan data, MATLAB routines are developed to quantify the progressive geometric changes occurring in patients during radiation therapy. Regression analysis is implemented to develop predictive models for tumor size changes through entire period. The generated models are validated using leave-one-out cross validation. The proposed method will increase the accuracy of therapy and improve patient's safety and quality of life by reducing the number of harmful unnecessary CT scans.

Keywords: H&N cancer, prediction model, regression, prototype models

1. INTRODUCTON

1.1. Head and Neck Cancer

Head and neck (H&N) cancer refers to a variety of malignant tumors that occur in the head and neck region. The term malignant defines these tumors as having the ability to metastasize or spread to other parts of the body. The goal in all cancer treatment is to remove the tumor with as little damage as possible to the important structures in the head and neck.

Most cancers are treated with surgery, radiation therapy (RT), chemotherapy, hormone therapy, or biological therapy. Doctors may decide to use one treatment

*Corresponding author: Ali K. Kamrani, Design and Free Form Fabrication Laboratory, University of Houston, Houston, TX, USA, akamrani@uh.edu. Other authors: mazimi@lenovo.com, hjsmadi@just.edu.jo.

method or a combination of methods. RT is a very important tool in the treatment of cancer. More than half a million cancer patients receive RT each year in US, either alone or in conjunction with surgery, chemotherapy or other forms of therapy. Together with image guided treatment planning, RT is a powerful tool in the treatment of cancer, particularly when the cancer is detected at an early stage. Chemotherapy and RT may also be used together to effectively treat the cancer. Intensity-modulated radiation therapy (IMRT) and Computer Tomography (CT) Imaging are the most widely used techniques for treatment planning [1].

1.2. Objective

The primary objective of radiotherapy is to deliver the correct dose of radiation to cancerous region with minimum damage to surrounding normal tissues in head and neck. Based on pre-treatment Computer Tomography (CT) images, a master plan for the treatment is developed at the beginning of the patient's visit to the hospital. Most types of tumors need more than 60 Gy (a radiation exposure unit) of total dose of radiation. Since human tissue can tolerate only a limited amount of radiation without side effects, the total dose is fractionated for a daily dose of 2 Gy, and the entire period will take 5 to 7 weeks. Most of the current RT planning systems assume that the tumor geometry will not change during the course of treatment. Therefore, based on this assumption, the daily planned radiation dose remains unchanged. However, there is a critical flaw in this assumption, because tumor geometry is shown to be changing over time. The size and shape may shrink or expand due to radiation or some other unknown reasons. Therefore, there is a critical need to track the changes in tumor geometry over time during the radiotherapy treatment.

In the present work, we propose a methodology in order to monitor and predict daily (fraction day) volume and surface changes of head and neck cancer tumors during the entire RT treatment period. The main goal of the proposed methodology is to increase the accuracy of each therapy and quality of life for patients. The objective is to develop a predictive model that captures the changes in the tumor geometry so that doctors and physicians can adjust the amount of planned radiation in each fraction day based on the predicted tumor size. The result of this project will have a positive impact on the head and neck cancer patients by reducing the inconvenient, costly, and time-consuming repetitive CT scans or Magnetic Resonance Imaging (MRI) during the radiation treatment therapy. In other words, using predictive models to monitor tumor deformation in certain fraction days can substitute CT scan in those days. Doctors can just use a few CT scans during the whole therapy to avoid any inaccuracy which may be caused by using the proposed predictive models. The literature survey reveals that no work has been conducted to model the geometrical changes that a tumor might go through after each radiation.

1.3. Background

Modeling and predicting deformations of body organs has gained increasing interest in recent years. These anatomical deformations can be caused by diverse factors such as breathing, tumor growth, etc. [2]. Tumor growth is one of the anatomical deformations

and has critical applications in cancer treatment. The first step in the modeling of tumor growth is to detect the location of the tumor using the computer tomography slices [3]. From modeling point of view, the first step is to acquire the volumetric visualization of the tumor structure. Volumetric objects often come from sampled volume datasets generated with biomedical data acquisition devices such as CT scan, MRI, and confocal microscopy. Volume visualization plays an increasingly important role in biomedical research and applications. The recent model-based volume visualization approach extends modeling-rendering paradigms to a volumetric environment where volumetric modeling and rendering are integrated into one model-based visualization system [4]. Volume visualization can be achieved by image processing techniques. Image data are processed using one or a combination of different segmentation techniques available.

Medical images are of great importance in RT, which became a privileged application field for image processing techniques. There is currently no single segmentation method that yields acceptable results for every medical image. Selection of an appropriate approach to a segmentation problem can therefore be a difficult dilemma. Hohne *et al.* [5] used ray casting algorithm working on a gray scale voxel data. Shen *et al.* [6] presented a deformable model for automatically segmenting brain structures from volumetric MR images and obtaining point correspondences using geometric and statistical information in a hierarchical scheme. Models proposed by Cootes *et al.* [7] and Wang and Staib [8] used equal weights in calculating shape statistical parameters. Cootes *et al.* [7] used two steps for active shape model (ASM) algorithm mainly for image data interrogation and shape approximation. Korn *et al.* [8] investigated the problem of shape retrieval from a large database based on the concept of mathematical morphology. Kervrann and Heitz [10] presented a motion-based segmentation method used for tracking the movement of deformable structures in image sequences. Ferrant *et al.* [11] modeled the biomedical shape changes using physics-based model of the objects the image represents. They used the shape changes of the surfaces of the objects as boundary conditions for the physics-based finite element (FE) model. Miller and Chinzei [12] proposed a constitutive model for brain tissue for modeling and simulation of surgical procedures using FE method based on the strain energy function in polynomial form with time dependent coefficients. Davatzikos *et al.* [13] proposed a framework for modeling and predicting anatomical deformations and tested on simulated images. Some researchers modeled deformation of human organs using biomechanical models. These models are considered to be accurate as they utilize the physical knowledge about the deformed anatomy and its properties, assuming that the material properties, equations governing the deformations and the boundary conditions are known [14–16]. However, these factors are not always known and they are very complex to determine.

Most of the reviewed methods are considered complex and require long computation time. Some of the reviewed methods are also based on some assumptions such as the knowledge available for properties of the deformable objects and their boundary conditions which are not always easy to access. Some of these methods also are implemented on the basis of deformation detection or segmentation using two dimensional (2D) images. The extended review and comparison among these methods

are presented in [17]. We aim to propose a model to remedy these disadvantages, which needs shorter computation time and involves simpler assumptions. Our approach is based on three dimensional geometrical analyses.

Moreover, several papers have highlighted the use of some specific data mining methods and statistical approaches in providing assistance in the field of prognosis, diagnosis, and treatment planning decision makings as well as prediction of medical treatment outcome. Delen *et al.* developed a prediction model for breast cancer survivability using Neural network, decision trees and logistic regression methods to generate the prediction [18]. Oztekin *et al.* adopted approaches such as decision trees, neural network, logistic regression and Cox regression model to predict the heart-lung graft survival time [19]. Kusiak *et al.* aimed to predict survival time for kidney dialysis patients using Rough-set (RS) and decision trees [20]. Pendharkar *et al.* used data envelopment analysis (DEA) (non-linear non-parametric mathematical technique) and artificial neural network for mining the breast cancer data [21]. Su *et al.* attempted to construct a prediction model for Type II diabetes using neural network, decision trees, logistic regression and rough set in order to select the most effective features to predict diabetes [22]. Chou *et al.* applied integration of artificial neural network and multivariate adaptive regression splines (MARS) methods to improve the performance of breast cancer diagnosis [23]. Kuo *et al.* described a novel computer-aided diagnosis system using decision trees approach in order to improve the diagnosis of breast cancer and increase the ability of ultrasonographic (US) technology for differential diagnosis [24]. Jilani *et al.* applied logistic regression to find the significant factors affecting the presence or absence of acute coronary syndrome [25]. Chhatwal *et al.* applied logistic regression in constructing prediction models for early detection of breast cancer [26]. Chen *et al.* applied a classification procedure including three classification techniques: discriminant analysis, support vector machine, and artificial neural network to differentiate innocent musical murmurs from other heart murmur types [27].

Prognostic models using data mining and statistical techniques in predicting patients' outcomes of interest can assist experts in the evaluation of effectiveness and efficiency of health centers services. Hanna and Keizer integrated classification trees and logistic regression methods in order to develop predictive models for patients' mortality in the Intensive Care (IC) [28]. Ture *et al.* applied decision trees algorithms and Cox regression analysis for disease-free survival in breast cancer patients [29]. Osareh and Shadgar utilized support vector machine, k nearest neighbor, neural network, naïve Bayesian, and decision trees to develop models for classification of cancer related gene expression [30]. Jonsdottir *et al.* generated predictive model which is able to classify the new breast cancer patient into either no-event or recurrence-event of cancer five years after diagnosis; Naïve Bayesian classifier, decision trees, and a wide range of meta algorithms were used in this research [31]. Aragonés *et al.* presented a decision support tool by combining neural network and decision trees to identify the most important prognostic factors for breast cancer relapse [32]. Most of the developed predictive models are associated with categorical class labels. No works have been devoted to predict deformation of cancer tumors by applying data mining and statistical techniques.

2. METHOD

2.1. Scope

Figure 1 illustrates the scope of the proposed methodology. The proposed methodology consists of two phases. The first phase constructs the model used for tumor geometrical analysis. The second phase uses proper statistical techniques for deformation prediction of the geometry based on patient's selected attributes (age, weight, stage, etc.). Finally, based on the prediction model, new treatment plans can be developed. The following is a detailed discussion of each phase.

2.2. Data Collection

Clinical patient data were obtained from the University of Texas-MD Anderson Cancer Center (MDACC), Department of Radiation Physics, with the approval by MDACC Institutional Review Board (IRB). Data of 23 patients were collected. In order to perform the analysis, two types of data are needed. The first type is CT scan data which come in the format of x, y, z coordinates or point clouds. The second type of data relates to patients' attributes which are collected through each of the 23 patients' clinical file. For geometrical analysis purpose, the first type of data, CT scan data, is used. In order to generate predictive model, a combination of the results of both geometrical analysis and patients' attributes is used.

2.3. Geometric Analysis

The first stage of the proposed methodology is implemented in two phases: (1) Geometric modeling and analysis, and (2) validation through Rapid Prototyping. The goal the first phase is to use the collected patient data to quantify the progressive

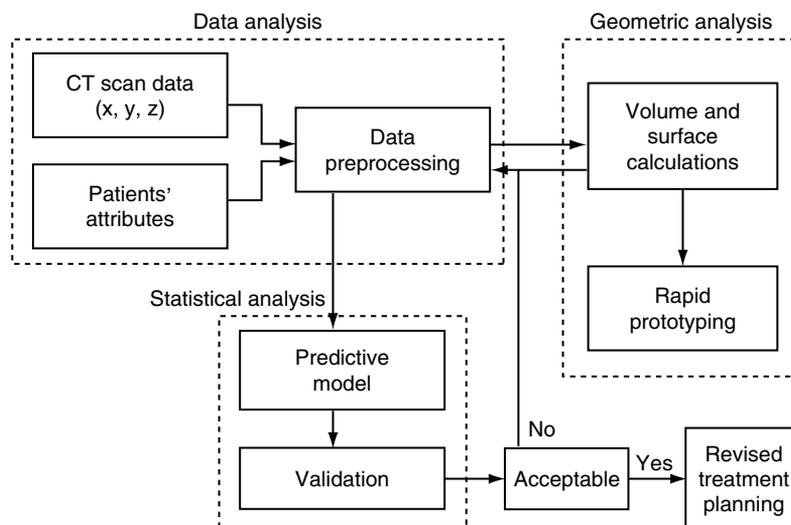


Figure 1. Scope of the proposed methodology.

geometric/anatomic changes occurring in patients treated with RT for head and neck cancer. Steps in this phase include (1) volume construction and analysis, and (2) surface contouring and construction. The CT scan data which are in the x, y, and z coordinates format are used in this phase.

2.3.1. Phase I: Geometric Modeling and Analysis

Two major objectives are associated with the geometry modeling and analysis phase. These two objectives are the following: (1) to visualize the tumor contour and patch the surface of the tumor, and (2) to establish the tumor's selected features such as volume and surface values. A series of MATLAB[®] routines are developed to perform these analysis and required calculations. The surface calculation method consists of the following steps:

- Read CT scan data of a patient.
- Extract data associated with GTV from CT scan.
- The processed data related to gross tumors are sorted by the z-coordinate. Points with the same z-value (same layer) are then used to construct the slice contours.
- Find the number of slices/layers in GTV (in point cloud format).
- Identify GTV data points' coordinates in x-, y-, and z-coordinates separately.
- Sort data based on z-coordinate.
- Identify the space between slices/layers.
- Identify the number of curves in each slice. Each slice may contain more than one curve.
- Using x-, y-, and z-coordinates of each curve, the binary mask is constructed for each layer by applying *roipoly* function. This step is repeated for the entire curves of GTV data.
- The region boundary in a binary image is traced and the number of pixels in the boundary is identified.
- Using the number of pixels and physical dimensions of them, the surrounding surface area of the tumor is calculated.
- Using the same procedure, the surface areas of the first and the last layers are identified and added to the surrounding surface.

In order to calculate the tumor volume, the same approach is applied, yet the number and dimension of voxels are used. The discussed methods are applied to calculate patients' tumor volume and surface values throughout the entire fraction days (33 days).

2.3.2. Phase II: Validation through Rapid Prototyping

The CT images can be exported to build a physical model of the anatomy using the rapid prototyping (RP) process. Rapid prototyping is a manufacturing technique that creates various 3D geometries by means of layer-by-layer construction. Rapid prototyping is impacting the medical sector in several ways. It can be used for surgical planning, prosthesis and implants, direct manufacture of biologically active implants and many others. Using principles of rapid prototyping, CT scan or MRI data of anatomical parts can be translated into physical data and can be used to manufacture a

3D 1:1 scale physical, three-dimensional model. Medical RP models are physical hard copies of a patient's specific anatomy, visualized by three-dimensional scanning techniques. These medical models provide visual and tactile information for diagnosis and operational planning [33, 34]. In the proposed research, RP is employed to build a set of prototype mockups of the tumors. The prototypes are used to inspect and measure the geometry of the tumors as they go through deformation. The tumor prototypes are intended for further validation of the proposed methodology.

2.3.3. Data Preprocessing

The volume/surface database includes volume/surface values' data for 23 patients collected in 33 fraction days. The patients' attributes were collected through the patients' clinical files. The collected attributes can be categorized as characteristic attributes (e.g., sex, marital status, age, etc.), tumor registry attributes (e.g., tumor stage, type of RT, dose, etc.), and patients' personal habits and disease history (e.g., smoke, alcohol, cancer history in family). Efforts were made to select only those attributes that may influence the results. The final selected dataset attributes include the following:

- **Age**
- **Sex**
- **Weight**
- **Height**
- **Primary tumor location:** the primary location of the tumor in the head and neck area.
- **Stage:** The stage of the tumor is identified based on TNM system, in which each cancer is assigned to a T , N , or M category. The T category describes the primary tumor size and the extent the tumor has grown to nearby tissues. It is presented by ordinal numbers of T ($T1$, $T2$, ...); the higher the T number, the larger the tumor and/or the more invasion to nearby tissues. The category N describes the extent the cancer has spread to nearby lymph nodes. It is presented by ordinal numbers ($N1$, $N2$, ...); the higher the N number, the more lymph nodes are involved. The M category describes whether or not the cancer has spread to distant areas of the body. It is presented by $M1$ or $M0$, indicating that cancer has or has not metastasized to distant organs, respectively.
- **SEER Stage:** stage to which a malignancy has developed using Surveillance Epidemiology and end results (SEER) summary staging system.
- **Smoke:** not smoker, former smoker, occasional smoker, and smoker.
- **Alcohol:** using or not using alcohol.
- **Cancer History:** whether the patient has or does not have any cancer history in his/her family.

In order to prepare the attribute dataset for the statistical analysis phase, each attribute was categorized to different levels, and then transformed into *dummy* or *indicator* variables. If the categorical variable has n levels, it is possible to define $n-1$ dummy variables [35]. The implemented method is *reference coding*, which uses zero and one to code the variable.

Based on the tumor volume values, tumor surface value and attribute datasets, two separate datasets were created for the next phase, statistical analysis. For this purpose, the attribute dataset is integrated with volume and surface datasets separately to form two individual datasets associated with volume and surface, respectively.

2.4. Statistical Analysis

In this phase, the aim is to develop a predictive model that can predict tumor deformation during RT for head and neck cancer patients. The preprocessed integrated datasets, consisting of tumor volume/surface values and patients' selected attributes, are treated as training sample data to develop the prediction model for tumor changes. The first two fraction day's tumor volume and surface values are assumed to be given to predict the rest of fraction days' volume and surface values which are in continuous format. Regression analysis, the most appropriate method to predict numerical values, was employed to generate predictive models.

3. RESULTS AND DISCUSSION

3.1. Geometrical Analysis

3.1.1. Tumor Volume and Surface Analysis

Using CT scan data in each fraction day, data related to the gross tumor can be extracted and visualized by patching the surface of the tumor. Figure 2 illustrates a sample of tumor visualization created by MATLAB code. Table 1 and Table 2 present the average tumor volume and surface data, respectively, for all of the 23 patients throughout the entire fraction days. Figure 3 and Figure 4 illustrate changes in tumor volume and surface during the whole RT period for individual patients. The graphs display a general decreasing trend in volume values during the treatment period.



Figure 2. Sample of tumor visualization by MATLAB.

Table 1. Average tumor volume values for all the patients

	Volume (cm ³)											
Case#	1	2	3	4	5	6	7	8	9	10	11	12
Average	27.86	44.18	16.13	9.40	8.82	33.15	10.16	76.78	37.88	57.42	37.52	30.25
Case#	13	14	15	16	17	18	19	20	21	22	23	
Average	27.66	79.30	88.65	12.39	29.99	27.46	11.74	59.30	56.98	27.75	2.54	

Table 2. Average tumor surface values for all the patients

	Surface area (cm ²)											
Case#	1	2	3	4	5	6	7	8	9	10	11	12
Average	71.99	89.07	35.83	26.80	25.05	72.80	29.65	152.77	72.42	111.55	87.82	61.56
Case#	13	14	15	16	17	18	19	20	21	22	23	
Average	51.40	168.45	136.22	35.35	64.51	63.83	30.95	96.06	108.25	74.17	11.84	

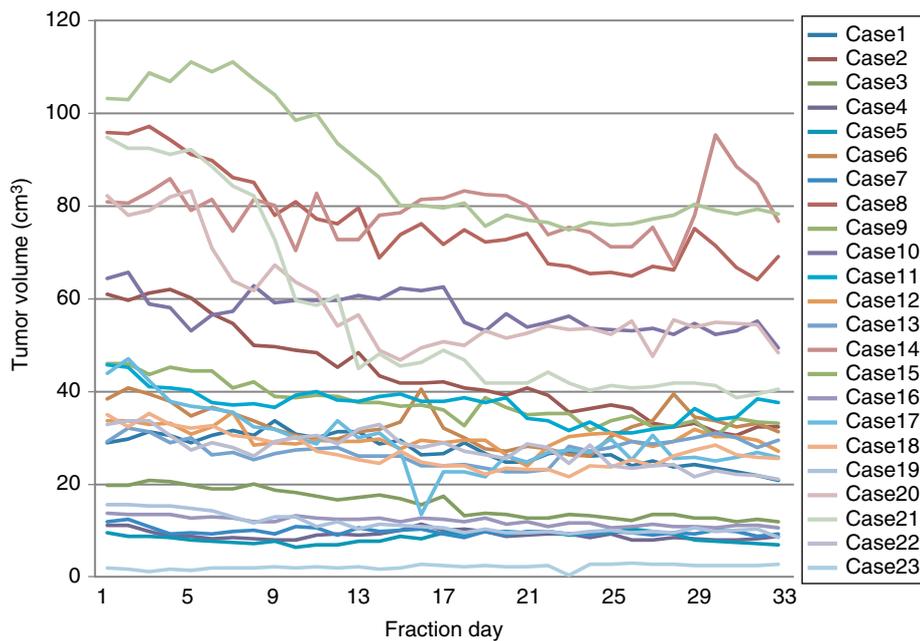


Figure 3. Tumor volume values changes during the whole period of therapy for each patient.

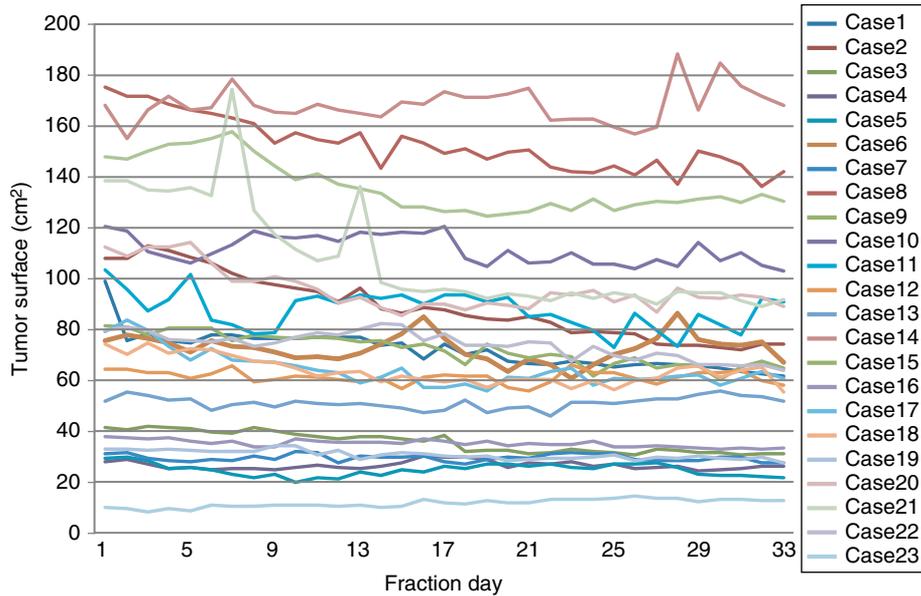


Figure 4. Tumor surface values changes during the whole period of therapy for each patient.

3.1.2. Rapid Prototyping Model

The prototypes are used to inspect and measure the geometry of the tumors as they go through deformation. They are intended for further validation of the proposed methodology. Sample prototype generated from CT scan data is illustrated in Figure 5 where the tumor is clearly visible in the head and neck area.

3.2. Statistical Analysis

To identify the best combination of predictor variables for each 31 day prediction model, the stepwise procedure has been conducted using MATLAB routines. In general, in identifying the best combination of variables for the dependent variable V_i or S_i ($i = 3, 4, \dots, 33$), the independent variables or predictors are patients' selected attribute plus subset of previous volume or surface values, including V_1, V_2, \dots, V_{i-1} or S_1, S_2, \dots, S_{i-1} . Considering the selected variables from stepwise procedure, we find that the most significant and effective predictor variable in prediction of each fraction day's volume/surface value is the previous day's volume/surface value. Therefore, the obtained result is extremely supportive of the idea of applying a simple regression model with one predictor variable, which is the previous day's volume/surface value. For each fraction day's tumor volume/surface, a prediction model is developed using MATLAB software. The main assumption is that the volume/surface values of the first two fraction days for each patient are given. Table 3 and Table 4 summarize the results of regression analysis for tumor volume and surface, respectively. The tables represent the regression coefficients including the intercepts or constant values (α) and variable coefficients (β) for all 31 models developed for the volume/surface value prediction of

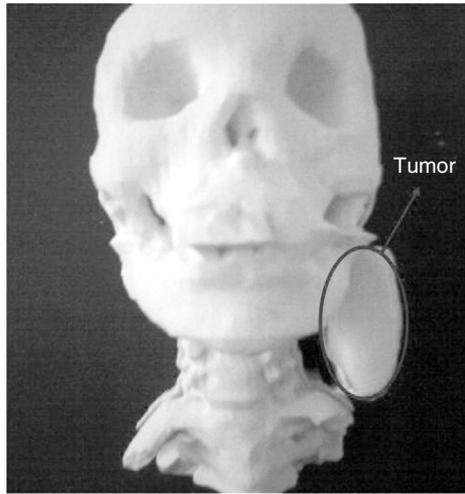


Figure 5. Rapid prototype model generated from CT scan data.

Table 3. Regression coefficients for 31 fraction days' tumor volume values

Day	3	4	5	6	7	8	9	10	11	12	13
α	-1.297	-0.825	-0.963	0.852	0.392	-0.700	1.455	1.583	-0.734	1.519	0.592
β	1.023	1.003	0.997	0.960	0.962	0.998	0.941	0.925	1.020	0.929	0.970
Day	14	15	16	17	18	19	20	21	22	23	24
α	1.250	0.151	-1.282	1.003	-0.983	0.838	-0.771	0.233	1.593	-0.297	0.653
β	0.942	0.984	1.030	0.971	1.008	0.958	1.020	0.983	0.947	0.997	0.980
Day	25	26	27	28	29	30	31	32	33		
α	0.724	-0.316	-0.100	0.816	-1.574	-1.390	0.860	0.488	0.465		
β	0.978	1.004	1.007	0.976	1.084	1.046	0.952	0.984	0.953		

Table 4. Regression coefficients for 31 fraction days' tumor surface values

Day	3	4	5	6	7	8	9	10	11	12	13
α	-2.064	-1.568	0.139	0.287	-3.654	5.422	2.033	1.105	-0.738	-0.029	-0.444
β	1.021	1.010	0.991	0.986	1.067	0.888	0.956	0.988	1.005	0.984	1.033
Day	14	15	16	17	18	19	20	21	22	23	24
α	4.035	-0.970	0.851	-0.185	-0.879	1.150	-1.615	-0.138	3.095	0.561	-0.442
β	0.907	1.022	0.983	1.004	0.986	0.978	1.022	0.999	0.950	0.988	1.001
Day	25	26	27	28	29	30	31	32	33		
α	0.807	0.875	-0.719	-2.681	2.292	-2.877	2.106	1.897	-1.465		
β	0.982	0.992	1.011	1.063	0.965	1.035	0.963	0.973	0.997		

fraction day 3 to 33. Physicians can use the predictive models to predict the tumor deformation for a new patient.

3.3. Validation

3.3.1. Validation by R^2

There are different criteria for checking the quality of the model. One of the most common criteria, coefficient of determination, R^2 , is used in this research. The R^2 value varies from zero to one. The closer this value to one, the better the regression model. Table 5 and Table 6 summarize the values of R^2 for each fraction day's volume and surface prediction model, respectively. For all models, the values of R^2 are above 0.9, and most of them are above 0.99, indicating good levels of fit.

3.3.2. Cross Validation

It is necessary to assess how the results will generalize to an independent dataset. An important point in the construction and evaluation of a prediction model is that they should not be built and tested on the same dataset to avoid the problem of *overfitting*. In our study, a special case of k -fold cross validation, which is called *leave-one-out*, is implemented. This is the same as the k -fold cross validation with k being equal to the number of observations in the original dataset. As the name suggests, it involves reserving a single observation at a time as a testing set and using the remaining observations from the original dataset as training set. This procedure is repeated such that each observation is used once as testing data.

Table 5. R^2 values for 31 volume prediction models

Day	3	4	5	6	7	8	9	10	11	12	13
R^2	0.993	0.997	0.995	0.991	0.994	0.990	0.990	0.985	0.988	0.989	0.975
Day	14	15	16	17	18	19	20	21	22	23	24
R^2	0.982	0.989	0.977	0.984	0.991	0.991	0.995	0.995	0.991	0.992	0.994
Day	25	26	27	28	29	30	31	32	33		
R^2	0.993	0.995	0.987	0.980	0.985	0.975	0.995	0.996	0.987		

Table 6. R^2 values for 31 surface prediction models

Day	3	4	5	6	7	8	9	10	11	12	13
R^2	0.992	0.997	0.996	0.990	0.967	0.964	0.998	0.993	0.998	0.998	0.982
Day	14	15	16	17	18	19	20	21	22	23	24
R^2	0.967	0.992	0.993	0.995	0.993	0.996	0.995	0.996	0.994	0.993	0.993
Day	25	26	27	28	29	30	31	32	33		
R^2	0.995	0.992	0.993	0.974	0.973	0.989	0.997	0.990	0.994		

In this research, the ratio of PRESS (predicted residual sum of squares) to SSE (error sum of square) is taken as an indicator for the level of fit. Figure 6 and Figure 7 display the graph of PRESS/SSE values in all fraction days for volume and surface predictions, respectively. The results reveal that for most of the fraction days, this ratio is close to 1.2, verifying a good level of fit. However, as the volume graph (Figure 6) shows, this ratio is higher than the normal value on fraction day 30, and surface graph (Figure 7) shows that this ratio is higher than the normal value on days 28, 29, and 30. In this case or similar cases where the results show some deviations from the accepted range, it is necessary to check the possible existence of outliers or highly influential observations in the model.

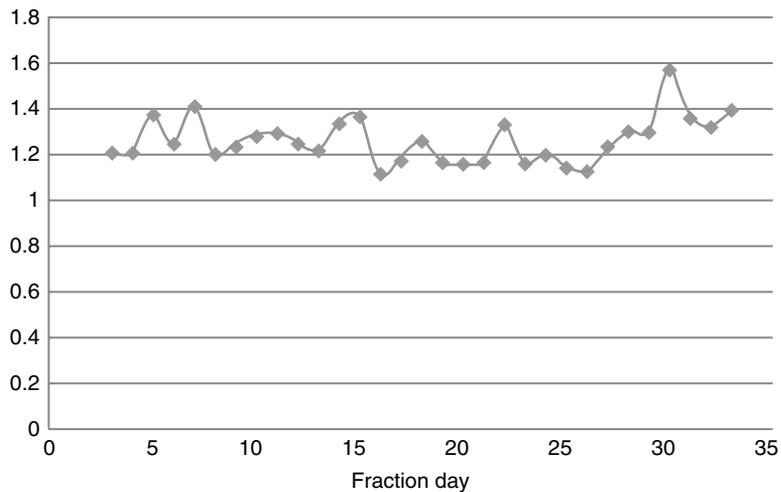


Figure 6. PRESS to SSE ratio for tumor volume.

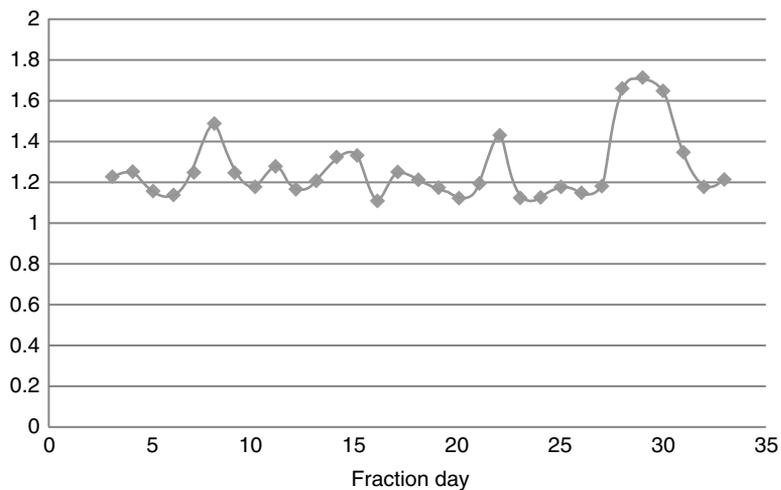


Figure 7. PRESS to SSE ratio for tumor surface.

The existence of any outliers or highly influential observations can significantly impact the final results. We used studentized residual values as the outlier detector, and identified one patient's data as a possible outlier. There are different reasons that could explain this result such as erroneous CT scan data in some fraction days of this patient.

3.4. Significance

Using the proposed method, we are able to measure tumor size using CT scan data and generate prediction model during the fraction days. Many previous research works were devoted to mathematically modeling cancer tumors disregarding the stages of deformation as the tumors go through deformation. Moreover, studies that were focused on tumor deformation and modeling used biological condition and information of the organs itself. For example, Kima *et al.* proposed and designed patient-specific RT strategies based on the patient's biological responses over several treatment sessions [36]. The proposed mathematical method used the patient's biological condition to model the system state and the beam intensities as controls [36]. Other researchers such as Kim *et al.* proposed a method to mathematically explore the potential benefit of fractionated RT using Markov decision process (MDP) model [37], and Ferris and Voelker formulate a model for the day-to-day RT planning using heuristics and neuro-dynamic programming and show that the results can be achieved by incorporating uncertainty about errors during treatment planning [38]. The tool developed by Swanson uses MRI data to predict where a patient's tumor can grow a few months ahead of time [39]. The software is based on MRI data to simulate how fast a tumor is likely to grow and how long a patient is likely to live under different scenarios [39]. These methods, although may be valid, are mathematically complex and require long computation time. In the proposed method, CT scan data are used to visualize 3D image of the tumor and the information from these data is used for calculation of geometry. This process does not require complicated image analysis methods. Statistical and regression analyses are applied for effective and accurate deformation analysis and prediction, which is used by the clinicians to better plan the RT area for better accuracy and safety.

4. CONCLUSIONS

In this research, a methodology is proposed to monitor and predict daily (fraction day) volume and surface changes of head and neck cancer tumors during RT. The proposed method enables using geometrical and statistical analyses rather than repetitive CT scans to predict the tumor deformation during RT. Data of real patients were obtained from MD Anderson Cancer Center, including patient clinical attributes and CT scan data of 33 fraction days of RT. CT scan data were used to develop MATLAB routines to calculate tumor volume and surface. Patients' selected clinical attributes were used to build the training dataset for regression analysis implemented to develop predictive models for tumor volume and surface changes throughout the fraction days. The generated models were validated through leave-one-out cross validation. The predictive models may help physicians to predict the entire fraction days (generally 33 fraction day) for a new patient based on patient's selected attributes which are usually available after the first patient visit and the first two fraction days' CT scan data. The proposed methodology may help

increase the accuracy of RT and avoid delivery of radiation to surrounding normal tissues in head and neck area. It may also reduce the number of harmful, repetitive CT scans and thus improving patient safety and quality of life.

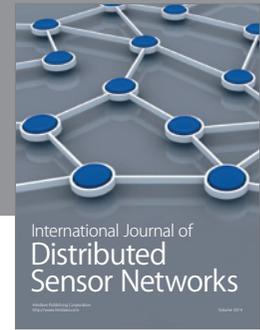
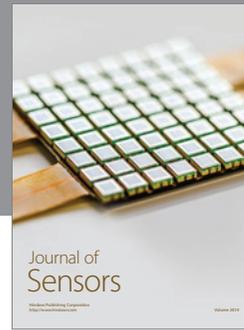
CONFLICT OF INTEREST

The authors indicated no potential conflicts of interest.

REFERENCES

- [1] Kamrani A, Azimi M. Geometrical modeling of H&N cancer tumor. *Journal of Rapid Prototyping*. 2011, 17(1): 55–63.
- [2] Mohamed A, Kyriacou KS, Davatzikos C. A statistical approach for estimating brain tumor induced deformation. *IEEE workshop on Mathematical Methods in Biomedical Image Analysis*. 2001, 52–59.
- [3] Zizzari A, Seiffert U, Michaelis B, Gademann G, Swiderski S. Detection of tumor in digital images of the brain. *Proceedings of the IASTED International Conference on Signal Processing, Pattern recognition and Applications*, Rhodes, Greece, 2001, 132–137.
- [4] Kamrani A, Heramb I, George R. Rapid Prototyping Applications in Cancer Tumor Deformation Modeling. *34th International Conference on Computers and Industrial Engineering*, 2004, San Francisco, CA.
- [5] Hohne K.H, Bomans M, Pommert A, Riemer M, Schiers C, Tiede U, Wiebecke G. 3D-Visualization of tomographic volume data using the generalized voxel model. *The Visual Computer*, 1990, 6: 28–36.
- [6] Shen D, Herskovits EH, Davatzikos C. An adaptive-focus statistical shape model for segmentation and shape modeling of 3-D brain structures. *IEEE transactions on medical imaging*, 2001, 20: 257–270.
- [7] Cootes TF, Cooper D, Taylor CJ, Grahant J. Active shape models-their training and application. *Computer Vision and Image Understanding*. 1995, 61: 38–59.
- [8] Wang Y, Staib LH. Elastic model based non-rigid registration incorporating statistical shape Information. *Proceeding of MICCAI*. 1998, 1162–1173.
- [9] Korn P, Sidiropoulos N, Faloutsos C, Siegel E, Protopapas Z. Fast and effective retrieval of medical tumor shapes. *IEEE transactions on knowledge and data engineering*. 1998, 10: 889–904.
- [10] Kervrann C, Heitz F. Statistical deformable model-based segmentation of image motion. *IEEE transactions on image processing*. 1999, 8: 583–588.
- [11] Ferrant M, Macq B, Nabavi A, Warfield SK. Deformable Modeling for Characterizing Biomedical Shape Changes. *Proceedings of 9th international conference on Discrete Geometry for Computer Imagery*. Springer, Berlin, Heidelberg, 2001, 235–248.
- [12] Miller K, Chinzei K. Constitutive modeling of brain tissue. *Experiment and Theory. J. Biomechanics*. 1997, 30: 1115–1121.
- [13] Davatzikos D, Shen D, Mohamed A, Kyriacou S. A framework for predictive modeling of anatomical deformations. *IEEE Transactions on Medical Imaging*. 2001, 20: 836–843.
- [14] Andersen PR, Bookstein FL, Conradsen K, Ersboll BK, Marsh JL, Kreiborg S. Surface bounded growth modeling applied to human mandibles. *IEEE transactions on medical imaging*. 2000, 19: 1053–1063.
- [15] Asachenkov AA. Nonlinear models and data analysis of cancer patients. *IEEE international symposium on circuits and systems*. 1989, 3: 1617–1619.
- [16] Garbey M, Zouridakis G. Modeling tumor growth: From differential deformable models to growth prediction of tumors detected in PET images. *Proceedings of the 25th Annual International Conference on IEEE EMBS*. Cancun, Mexico, 2003, 3: 2687–2690.
- [17] Kamrani A, Azimi M. Geometrical Modeling of H&N Cancer Tumor. *Journal of Rapid Prototyping*. 2011, 17 (1): 55-63.
- [18] Delen D, Walker G, Kadam A. Predicting breast cancer survivability: a comparison of three data mining methods. *Artificial Intelligence in Medicine*. 2005, 34: 113–127.

- [19] Oztekin A, Delen D, Kong ZJ. Predicting the graft survival for heart-lung transplantation patients: an integrated data mining methodology. *International Journal of Medical Informatics*. 2009, 78: e84–e96.
- [20] Kusiak A, Dixon B, Shah S. Predicting survival time for kidney dialysis patient: a data mining approach. *Computers in Biology and Medicine*. 2005, 35: 311–327.
- [21] Pendharkar PC, Rodger JA, Yaverbaum GJ, Herman N, Benner M, Association, statistical, mathematical and neural approaches for mining breast cancer patterns. *Expert System with Applications*. 1999, 17: 223–232.
- [22] Su CT, Yang CH, Hsu KH, Chiu WK. Data mining for diagnosis of type II diabetes from three-dimensional body surface anthropometrical scanning data. *Computers and Mathematics with Applications*. 2006, 51: 1075–1092.
- [23] Chou SM, Lee TS, Shao YE, Chen IF. Mining the breast cancer pattern using artificial neural networks and multivariate adaptive regression splines. *Expert Systems with Application*. 2004, 27: 133–142.
- [24] Kuo WJ, Chang RF, Chen DR, Lee CC. Data mining with decision trees for diagnosis of breast tumor in medical ultrasonic images. *Breast Cancer Research and Treatment*. 2001, 66: 51–57.
- [25] Jilani TA, Yasin H, Yasin M, Ardil C. Acute coronary syndrome prediction using data mining techniques—An application. *World Academy of Science, Engineering and Technology*. 2009, 59: 474–478.
- [26] Chhatwal J, Alagoz O, Lindstrom MJ, Kahn CE, Shaffer KA, Burnside ES. A logistic regression model based on national mammography database format to aid breast cancer diagnosis. *AJR Am J Roentgenol*. 2009, 192: 1117–1127.
- [27] Chen Y, Wang S, Shen CH, Choy FK. Intelligent identification of childhood musical murmurs. *Journal of Healthcare Engineering*. 2012, 3(1): 125–140.
- [28] Hanna AA, Keizer ND. Integrating classification trees with local logistic regression in intensive Care prognosis. *Artificial Intelligence in Medicine*. 2003, 29: 5–23.
- [29] Ture M, Tokatli F, Omurlu IK. The comparisons of prognostic indexes using data mining techniques and Cox regression analysis in the breast cancer data. *Expert Systems with Applications*. 2009, 36: 8247–8254.
- [30] Osareh A, Shadgar B. Classification and diagnostic prediction of cancers using gene microarray data analysis. *Journal of Applied Sciences*. 2009, 9: 459–468.
- [31] Jonsdottir T, Hvannberg ET, Sigurdsson H, Sigurdsson S. The feasibility of constructing a predictive outcome model for breast cancer using the tools of data mining. *Expert System with Applications*. 2008, 34: 108–118.
- [32] Jerez-Aragones JM, Gomez-Ruiz JA, Ramos-Jimenez G, Munoz-Perez J, Alba-Cpnejo E. A combined neural network and decision trees model for prognosis of breast cancer relapse. *Artificial Intelligence in Medicine*. 2003, 27: 45–63.
- [33] Inamdar H, Kamrani A, George R. Rapid prototyping application in cancer tumor deformation modeling. *Proceeding of the 34th international conference on computers and industrial engineering*. 2004.
- [34] Olszewski R. Surgical engineering in Cranio-Maxillofacial surgery: A literature review. *Journal of Healthcare Engineering*. 2012, 3(1): 53–86.
- [35] Mickey RM, Dunn OL, Clark VA. *Applied Statistics Analysis of Variance and Regression*, 3rd ed. Wiley and Sons Inc, New Jersey, 2004.
- [36] Kima M, Ghateb A, Phillips MH. A stochastic control formalism for dynamic biologically conformal radiation therapy. *European Journal of operation research*. 2012, 219: 541–556.
- [37] Kim M, Ghate A, Phillips MH. A Markov decision process approach to temporal modulation of dose fractions in radiation therapy planning. *Physics in Medicine and Biology*. 2009, 54: 4455–4476.
- [38] Ferris MC, Voelker MM. Fractionation in radiation treatment planning. *Math Programming*. 2004, 101: 387–413.
- [39] Swanson KR. Can math cure cancer?. *Forbes magazine*. 2008.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

