Int. J. Radiation Oncology Biol. Phys., Vol. 69, No. 1, pp. 230–239, 2007

0360-3016/07/$–see front matter

ELSEVIER

doi:10.1016/j.ijrobp.2007.04.066

PHYSICS CONTRIBUTION

PRINCIPAL COMPONENT ANALYSIS-BASED PATTERN ANALYSIS OF DOSE–VOLUME HISTOGRAMS AND INFLUENCE ON RECTAL TOXICITY

MATTHIAS SÖHN, DIPL.PHYS.,* MARKUS ALBER, PH.D.,* AND DI YAN, D.SC.†

*Section of Biomedical Physics, University Hospital for Radiation Oncology, Tübingen, Germany; and †Department of Radiation Oncology, William Beaumont Hospital, Royal Oak, MI

Purpose: The variability of dose–volume histogram (DVH) shapes in a patient population can be quantified using principal component analysis (PCA). We applied this to rectal DVHs of prostate cancer patients and investigated the correlation of the PCA parameters with late bleeding.

Methods and Materials: PCA was applied to the rectal wall DVHs of 262 patients, who had been treated with a four-field box, conformal adaptive radiotherapy technique. The correlated changes in the DVH pattern were revealed as "eigenmodes," which were ordered by their importance to represent data set variability. Each DVH is uniquely characterized by its principal components (PCs). The correlation of the first three PCs and chronic rectal bleeding of Grade 2 or greater was investigated with uni- and multivariate logistic regression analyses.

Results: Rectal wall DVHs in four-field conformal RT can primarily be represented by the first two or three PCs, which describe ~94% or 96% of the DVH shape variability, respectively. The first eigenmode models the total irradiated rectal volume; thus, PC1 correlates to the mean dose. Mode 2 describes the interpatient differences of the relative rectal volume in the two- or four-field overlap region. Mode 3 reveals correlations of volumes with intermediate doses (~40–45 Gy) and volumes with doses >70 Gy; thus, PC3 is associated with the maximal dose. According to univariate logistic regression analysis, only PC2 correlated significantly with toxicity. However, multivariate logistic regression analysis with the first two or three PCs revealed an increased probability of bleeding for DVHs with more than one large PC.

Conclusions: PCA can reveal the correlation structure of DVHs for a patient population as imposed by the treatment technique and provide information about its relationship to toxicity. It proves useful for augmenting normal tissue complication probability modeling approaches.

Prostate cancer, Rectal toxicity, Dose–volume histograms, Principal component analysis, Normal tissue complication probability.

INTRODUCTION

Advances in modern radiotherapy (RT), such as three-dimensional conformal RT (3D-CRT) and intensity-modulated RT, have enabled accurate tailoring of dose distributions to target volumes and better sparing of adjacent normal structures, thereby facilitating increased target doses. Safe dose escalation, however, requires reliable information about normal tissue complications and its dependence on dose and volume.

For prostate RT, the essential dose-limiting organs are the bladder and rectum, with chronic rectal bleeding as one of the most relevant side effects. A first step to correlating complications with the applied doses is a reduction of the 3D dose distributions to dose–volume histograms (DVHs) at the expense of losing information about location. Several studies have found significant correlations between the parameters derived from rectal DVHs and the incidence of bleeding (1–8). However, the studies differ in the parameters used as summary measures of the DVH. In one approach, correlations of single DVH features such as the maximal dose or the value of a single DVH point (*i.e.*, the volume V_{D_c} receiving doses greater than a cutoff dose level D_c or the dose D_{V_c} to a cutoff volume V_c) and toxicity are investigated (1, 2, 5–8). In another approach, a comprehensive surrogate value such as an effective volume (9, 10), the mean dose, or another generalized equivalent uniform dose (EUD) (11, 12) is calculated and tested for its correlation with toxicity (3, 6–8). Some of these studies aimed to quantify complication risk

Reprint requests to: Matthias Söhn, Dipl.Phys., Section of Biomedical Physics, University Hospital for Radiation Oncology, Hoppe-Seyler-Str. 3, 72076 Tübingen, Germany. Tel: (++49) 7071-2986061; Fax: (++49) 7071-295920; E-mail: Matthias.Soehn@med.uni-tuebingen.de

Supported in part by Deutsche Krebshilfe e.V. Grant 106280 and National Institutes of Health Grant RO1 CA091020.

Presented in part at the 47th Annual Meeting of the American Society for Therapeutic Radiology and Oncology (ASTRO), Denver, CO, October 16–20, 2005.

Conflict of interest: none.

Received Sept 20, 2006, and in revised form Feb 14, 2007. Accepted for publication April 14, 2007.

in terms of normal tissue complication probability (NTCP) models and determine the corresponding model parameters (3, 4, 6, 8).

An inherent problem of outcome modeling, especially of direct dose–volume-based approaches (*i.e.*, models that consider only single DVH features), is the influence of the treatment technique on the results of modeling. For a given patient population, the treatment technique used induces correlations between DVH bins of different dose levels owing to the interaction of the given beam directions and shapes with the variability of the patient geometries. Because this can deteriorate the prospective use of such models for other patient populations treated using different techniques, these correlations should be considered when interpreting and comparing the results of different studies. Thus, a method that explicitly reveals the correlation structure of a DVH data set is desirable.

Recently, Dawson *et al.* (13) and Bauer *et al.* (14, 15) proposed the use of principal component analysis (PCA) to analyze the partial volume effects of normal tissues to RT and applied it to DVHs of the liver and parotid gland (13) and rectal wall (15), respectively. With this multivariate approach, the correlated variability of DVH shapes in a given patient population can be quantitatively described in terms of ''eigenmodes'', which provide information about the correlation structure inherent in the DVH data set. Moreover, PCA allows characterization of individual DVHs using a few parameters, the ''principal components'' (PCs) (formal definition given below in the subsection ''PCA of DVH data''). Regarding the PCs as a summary measure of individual DVH morphology, correlations with toxicity can be assessed using logistic regression models in a purely phenomenologic manner. However, the value of these models is restricted to the treatment technique for which they were derived.

Bauer *et al.* (14, 15) analyzed two data sets comprising 52 and 119 rectal wall DVHs of patients treated with a six-field 3D-CRT technique with a prescription dose of 70.2 Gy and 75.6 Gy, respectively, and found correlations of some of the dominating PCs with rectal bleeding of Grade 2 or greater.

In the present study, PCA was applied to rectal wall DVHs of 262 prostate cancer patients treated with a four-field box, 3D-CRT adaptive RT (ART) technique, thereby revealing correlations of different dose levels for the population as imposed by this treatment technique. Correlation of the first three PCs and chronic rectal bleeding of Grade 2 or greater was then investigated with logistic regression analysis. The same patient population was used in a parallel study (4), in which six EUD- and dose–volume-based NTCP models were applied to the data. This allowed a comparison of the different approaches with respect to their power to describe the data and make predictions.

METHODS AND MATERIALS

Patient data

The patient data set used comprises 262 prostate cancer patients treated between 1999 and 2002 at the William Beaumont Hospital. This patient data set, with a minimal follow-up time of 1 year (see below), represents a subgroup of a patient population used in a parallel study (5). For additional details, we refer to that study and limit the present description to information relevant for the following investigation. The clinical characteristics of the patient population have been previously described (7, 16).

The patients were part of a phase II dose-escalation study and underwent 3D-CRT with image-guided off-line correction under an ART protocol (17, 18). In brief, a four-field box technique was used for the initial treatment plan of the first week and the adapted plan. The initial planning target volume (PTV) was defined based on the clinical target volume plus a 1-cm uniform margin. For the adapted plan, information from daily portal and computed tomography (CT) imaging was used to form a patient-specific confidence-limited PTV (cl-PTV), thereby considering random and systematic errors as estimated from the first week. Beam apertures for the initial/adapted plan were defined according to the PTV/cl-PTV in the beam's eye view, with a PTV-to-field edge margin of 7 mm everywhere, but 11 mm at the superior and inferior edges of the cl-PTV. The final dose to the cl-PTV was limited by the dose–volume constraints of rectal wall and bladder. For the rectal wall, these were $D_{30\%} = 75.6$ Gy for the minimal dose received by 30% of the target volume and 82 Gy for the minimal dose received by $D_{5\%} = 82$ Gy of the target volume. For each patient, the dose level (minimal cl-PTV dose) was chosen individually so as to meet the rectum and bladder constraints and was one of the following doses as defined by the study protocol: 70.2, 72, 73.8, 75.6, 77.4, and 79.2Gy.

Cumulative DVHs. For DVH calculation, a composite planning dose was used. It included the initial treatment plan for the first week before correcting the systematic error and then the adaptive plan for the rest of the treatment after beam aperture correction according to the cl-PTV. The dose distributions of both plans were calculated using Pinnacle 6.2b (ADAC Laboratories, Milpitas, CA) according to the CT geometry (density information) of the planning CT scan. The overall dose was defined as the sum of the initial and adapted (physical) dose distributions. An in-house–developed software program was used to calculate the DVHs of the rectal wall, which was defined according to the solid rectum contours of the planning CT with 3–4-mm wall thickness. The dose bin size used for calculation of the DVHs was 0.1 Gy, with the volume defined as relative (percentage) volume.

Follow-up information. The toxicity variable regarded in this study was chronic rectal bleeding, for which the grading was determined using the Common Terminology Criteria for Adverse Events (CTCAE), version 3.0 (19). Although the full patient population, as described in our previous study (4), comprised 319 patients, in the present study we considered only those patients with a follow-up time of $\geq$1 year. This resulted in a median clinical follow-up for the remaining 262 patients of 3.2 years (range, 1.0–6.4 years), with an interquartile range of 2.3–4.2 years (25th to 75th percentile).

Principal component analysis of DVH data

Given n observations of p variables, a method from multivariate statistics, PCA, can be used to analyze and describe the correlated variability of the p variables in a data set. First applications of PCA in the field of RT have been proposed only recently and encompass automatic model-based organ segmentation (20) and modeling of organ deformations (21). Regarding the volume values of the dose bins as variables, PCA can also be applied to a set of DVHs to efficiently describe the variance of organ DVH shapes present in a patient population as first proposed by Bauer *et al.* (14, 15) and Dawson *et al.* (13). For our study, we used $p = 850$