Coverage-based treatment planning to accommodate delineation uncertainties in prostate cancer treatment

Huijun Xu^{a)}

Department of Radiation Oncology, Virginia Commonwealth University, Richmond, Virginia 23298
and Department of Radiation Oncology, University of Maryland, Baltimore, Maryland 21201

J. James Gordon

Department of Radiation Oncology, Henry Ford Health System, Detroit, Michigan 48202

Jeffrey V. Siebers

Department of Radiation Oncology, Virginia Commonwealth University, Richmond, Virginia 23298
and Department of Radiation Oncology, University of Virginia, Charlottesville, Virginia 22908

(Received 24 April 2015; revised 23 July 2015; accepted for publication 28 July 2015; published 21 August 2015)

Purpose: To compare two coverage-based planning (CP) techniques with fixed margin-based (FM) planning for high-risk prostate cancer treatments, with the exclusive consideration of the dosimetric impact of delineation uncertainties of target structures and normal tissues.

Methods: In this work, 19-patient data sets were involved. To estimate structure dose for each delineated contour under the influence of interobserver contour variability and CT image quality limitations, 1000 alternative structures were simulated by an average-surface-of-standard-deviation model, which utilized the patient-specific information of delineated structure and CT image contrast. An IMRT plan with zero planning-target-volume (PTV) margin on the delineated prostate and seminal vesicles [clinical-target-volume (CTV_{prostate}) and CTV_{SV}] was created and dose degradation due to contour variability was quantified by the dosimetric consequences of 1000 alternative structures. When D_{98} failed to achieve a 95% coverage probability objective $D_{98,95} \geq 78$ Gy (CTV_{prostate}) or $D_{98,95} \geq 66$ Gy (CTV_{SV}), replanning was performed using three planning techniques: (1) FM (PTV_{prostate} margin = 4,5,6 mm and PTV_{SV} margin = 4,5,7 mm for RL, PA, and SI directions, respectively), (2) CP_{OM} which optimized uniform PTV margins for CTV_{prostate} and CTV_{SV} to meet the $D_{98,95}$ objectives, and (3) CP_{COP} which directly optimized coverage-based objectives for all the structures. These plans were intercompared by computing percentile dose-volume histograms and tumor-control probability/normal tissue complication probability (TCP/NTCP) distributions.

Results: Inherent contour variability resulted in unacceptable CTV coverage for the zero-PTV-margin plans for all patients. For plans designed to accommodate contour variability, 18/19 CP plans were most favored by achieving desirable $D_{98,95}$ and TCP/NTCP values. The average improvement of probability of complication free control was 9.3% for CP_{COP} plans and 3.4% for CP_{OM} plans.

Conclusions: When the delineation uncertainties need to be considered for prostate patients, CP techniques can produce more desirable plans than FM plans for most patients. The relative advantages between CP_{COP} and CP_{OM} techniques are patient specific. © 2015 American Association of Physicists in Medicine. [<http://dx.doi.org/10.1118/1.4928490>]

Key words: coverage, treatment planning, IMRT optimization, delineation uncertainties, prostate cancer

1. INTRODUCTION

With the development of image guided radiation therapy (IGRT), delineation (contouring) uncertainties (DUs) are becoming the dominant geometric uncertainties to be accounted for in treatment planning.^{1,2} Uncertainty regarding the microscopic spread of disease and the inability to reproducibly distinguish gross tumor—normal tissue boundaries in anatomical images are manifested as DUs. DUs may be influenced by many factors, such as limited image quality of employed imaging modalities, interobserver delineation uncertainties (different clinical judgments of different observers), and intraobserver delineation uncertainties (different clinical judgments of the same observer in different trials).

Autocontouring tools and deformable contour mapping used in adaptive therapy are also prone to DUs.^{3,4} Though site and application-specific, DUs may lead to significant bias of the delineated ROI from its truth and cause a non-negligible systematic offset in the true target for treatment⁵ and/or cause inaccurate assessment of normal tissue complication risk. While tumor DUs can be reduced by using functional imaging, anatomical imaging remains the norm, and DUs will contribute to planning uncertainties for the foreseeable future.

Many researchers have documented measurements^{2,6–10} of interobserver contour deviations with different imaging modalities. Weiss *et al.*⁹ showed that the average standard deviations (SDs) from average surface meshes of five repeated delineations on FBCT vs CBCT were 1.5 vs 2.1 mm for

prostate, 0.7 vs 1.4 mm for bladder, and 1.3 vs 1.5 mm for rectum. Bhardwaj *et al.*⁶ found statistically significant differences in dose-volume histograms (DVHs), tumor-control probability (TCP), and normal tissue complication probability (NTCP) values for differing observer's contours with prostate margins. An adequate accommodation of delineation uncertainties is required to achieve the anticipated treatment goal.

In current treatment planning, DUs are implicitly accommodated via the gross-tumor-volume (GTV) to clinical-target-volume (CTV) expansion. Plans in which the CTV is equal to the GTV, such as for the prostate, lack this implicit consideration. Independent of CTV expansion, DUs and their clinical impact are not explicitly considered in current plans. One might argue that the CTV to planning-target-volume (PTV) margin provides an additional route for implicit accommodation of delineation uncertainties, although this falls outside the scope of the ICRU PTV definition^{11,12} (e.g., ICRU 62, Table 2.1). Conceptually, one might consider DU accommodation via a margin approach, perhaps even via use of an empirical margin recipe^{5,13,14} such as those suggested for CTV-to-PTV or organ-at-risk-to-planning-organ-at-risk-volume (OAR-to-PRV) margin, with the understanding that parameters required for margin recipe application are known or can be measured, and that such parameters are applicable across a population. Note, margin recipes such as the van Herk margin formula¹⁵ (VHMF) intend to ensure a certain percentage of coverage on a population basis. However, such population-based margin may not adapt to patient-specific coverage probability in the presence of delineation uncertainties.¹⁶ Alternative formulas suggested for PRVs (Ref. 17) have the same limitation.

The patient-specific residual errors due to delineation uncertainties may be compensated by using coverage-based planning (CP) techniques. The coverage probability is computed by sampling the relevant geometric uncertainties, in this case alternative ROI structures. The dose distribution is optimized to achieve prespecified coverage probabilities as opposed to simply conforming to a heuristic PTV. In our lab, two CP techniques—optimized-margin planning (CP_{OM})¹⁸ and coverage-optimized planning (CP_{COP})¹⁹ have been developed. CP_{OM} iteratively increases PTV margins until the predefined coverage criteria are satisfied, and CP_{COP} directly modifies the fluence to achieve the specified coverage criteria. Our previous studies^{18–20} showed that both CP_{OM} and CP_{COP} can spare more normal tissue dose relative to conventional PTV-based planning when (rigid-body) setup errors or organ deformable motions were accounted for. However, the application of CP_{OM} or CP_{COP} to accommodate the delineation uncertainties has yet to be studied.

To use CP techniques, a mathematical model that represents possible variations of DUs is required. As “the weakest link in the search for accuracy,”¹⁶ DUs are not easily modeled. Though recent developed autocontouring tools allow fast and consistent contours between trials, the tedious and time-consuming manual contours are usually still required for validation and corrections.³ Based on a limited number of manual contours, most researchers have been narrowly focused on evaluation of boundary visualization^{2,9,21–23} or development

of 3D analysis tools.^{24–26} In this study, a three-dimensional voxel-based model was developed and incorporated into CP techniques. To simulate the possible DUs for prostate cancer treatment, this model utilized the information of patient-specific image contrast and a physician's delineation.

We studied contour delineation variations in isolation from patient positioning and organ shape variations. In clinical practice, such uncertainties would need to be considered, yet understanding DU impact in isolation provides insight into how to best clinically consider them, especially in light of the fact that image guided adaptive therapy is reducing positioning and shape error components.

Here, a DU model was incorporated in CP_{OM} and CP_{COP} techniques to optimize the dose distribution for each prostate patient. The potential clinical benefit of these two CP techniques was investigated by comparing target coverage and normal tissue sparing with the conventional fixed margin-based (FM) planning.

2. METHODS

The base flow of this study is as follows: for each of the 19 patients (see Sec. 2.A), an IMRT plan using a zero PTV margin was created. The 95% coverage probabilities of D_{98} of both prostate and seminal vesicles were evaluated based on the dosimetric consequences of 1000 alternative ROI structure sets simulated by a patient-specific DU model (see Sec. 2.C), which considers interobserver contouring difference and CT image quality. This established the need to consider DUs during treatment planning. To achieve the desired coverage probabilities, treatment plans were created using three different techniques: FM, CP_{OM}, and CP_{COP} (see Sec. 2.B.2). Then, these plans were compared with each other to determine which technique can best accommodate contouring variability (see Sec. 2.D.2). The plan comparison was based on the probability of achieving the desired dose-volume metrics and the probability for complication-free control ($P+$). Each aspect is described in further detail below.

2.A. Prostate plans

A 19 patient cohort obtained from Netherlands Cancer Institute (NKI) was involved in this study. There were 8–13 fan-beam CT images for each patient and the image resolution was $0.094 \times 0.094 \times 3$ mm³. More details of this patient cohort have been described in earlier published work.^{20,27,28} A single physician at VCU contoured the structures of prostate (CTV_{prostate}), seminal vesicles (CTV_{SV}), rectum, bladder, left femur, and right femur on each image. Contours of CTV_{SV}, rectum, and bladder were modified slightly to eliminate overlapping region with CTV_{prostate} or with other structures.

For each patient, the first CT was selected as the reference image for IMRT treatment planning to prescribe 78 Gy to CTV_{prostate} and 66 Gy to CTV_{SV} in 30 fractions. As we are studying delineation uncertainty in isolation, the other image sets were not utilized in this study. Seven coplanar (transverse) photon beams with gantry angles 30°, 80°, 130°, 180°, 230°, 280°, 330° were utilized with the beam isocenter

at the prostate centroid. The in-house IMRT criteria used for CP_{OM} , CP_{COP} , and FM were the same as a recent published work.²⁰ For the CP techniques, the coverage probability objectives³⁴ were $CTV_{prostate} D_{98,95} \geq 78$ Gy, $D_{2,5} \leq 81$ Gy, $CTV_{SV} D_{98,95} \geq 66$ Gy, bladder $D_{70,5} \leq 18$ Gy, $D_{50,5} \leq 36$ Gy, $D_{30,5} \leq 57$ Gy, $D_{20,5} \leq 66$ Gy, $D_{14,5} \leq 69$ Gy, $D_{9,5} \leq 75$ Gy, $D_{2,5} \leq 81$ Gy, and rectum $D_{50,5} \leq 36$ Gy, $D_{30,5} \leq 51$ Gy, $D_{20,5} \leq 66$ Gy, $D_{5,5} \leq 69$ Gy, $D_{2,5} \leq 75$ Gy. To achieve the static dose-volume criteria (used for FM or CP_{OM}) for 95% of the alternative structure sets simulated, $q = 95\%$ and $q = 5\%$ coverage values were selected for minimum (target) or maximum (OAR) percentile DVH (pDVH) criteria used for CP_{COP} .³⁵ All of the optimizations were performed using direct machine parameter optimization (DMPO) in the PINNACLE³ (research version 9.100) TPS (Philips Medical Systems, Fitchburg, WI). The optimizer was run for maximal 50 iterations and dose grid resolution was $2 \times 2 \times 2$ mm³ for both evaluation and optimization.

2.B. Coverage evaluation and coverage-based optimization

2.B.1. Coverage evaluation—pDVH

Coverage probability for each potential treatment plan was evaluated by mimicking dose delivery for 1000 alternative ROI structure sets. Each alternative structure of each set was independently sampled from the DU model, yielding different dose evaluations for each ROI. The 1000 alternative DVHs per ROI rendered a distribution from which the likelihood of achieving the specified dose-volume metric was determined.

From the DVH distribution, a percentile DVH at a specific coverage probability q was constructed by finding the line with equal coverage probability as a function of dose. Readers are referred to the previous^{19,20} work for details. The pDVHs were used in not only plan evaluation to estimate the chance that dose volume criteria may be met, but also in the CP_{COP} dose optimization to drive the pDVHs toward the desired coverage metrics.

2.B.2. Planning techniques: FM, CP_{COP} , and CP_{OM}

FM is a basic PTV-based IMRT planning technique using empirically predefined PTV margins. Here, the fixed PTV margins utilized by FM for all the patients in RL, PA, and SI were 4, 5, and 6 mm for $CTV_{prostate}$ and 4, 5, and 7 mm for CTV_{SV} . The margins M were calculated according to the VHM— $M = 2.5\Sigma$, where Σ refers to the SD values (Table I) of systematic error as a result of contouring variability.

CP_{COP} optimized the beam fluence and dose with probabilistic dose-volume criteria obtained from pDVHs. CP_{COP} included all the voxels that covered the dynamic ROI as sampled from the DU model for the different alternative structures, and weighted these voxels according to their potential contribution to the ROI dose. Then, the ROI dose was optimized based on Newton's method and minimizing the sum of a series of pDVH objectives [e.g., like Xu *et al.*²⁰ but with the DU model (see Sec. 2.C) used in the pDVH computation as opposed to the PCA organ deformation

TABLE I. SD values in mm of in RL(x), PA(y), and SI(z) directions used in ASSD model for ROIs of prostate cancer patients.

ROI	SD _x	SD _y	SD _z	Reference
$CTV_{prostate}$	1.7	2	2.5	van Herk (Ref. 5) and Rasch <i>et al.</i> (Ref. 22)
CTV_{SV}	1.7	2	3	Fiorino <i>et al.</i> (Ref. 7) and Deurloo <i>et al.</i> (Ref. 27)
Bladder	0.7	0.7	3	Weiss <i>et al.</i> (Ref. 9)
Rectum	1.3	1.3	3	Weiss <i>et al.</i> (Ref. 9)

model]. The probabilistic dose volume metrics included $D_{98,95}$ for CTV structures and $D_{2,5}$ for bladder and rectum drove plan optimization to maximally achieve 95% coverage probability for target D_{98} and 5% coverage probability for OAR D_2 criteria. The details on the CP_{COP} algorithm can be found in our previous work.^{19,20} As in our prior work, dose shift invariance²⁸ was assumed to reduce required computations during the CP_{COP} optimization.

CP_{OM} utilizes the method developed by Gordon and Siebers,¹⁸ starting with the IMRT plan using 0 PTV margins for $CTV_{prostate}$ and CTV_{SV} and same dose-volume objectives as those used in FM technique. PTV margin adequacy is judged by the 1000 alternative ROI set coverage evaluation. When the evaluation showed insufficient coverage, the size of PTV margin was iteratively expanded uniformly by a 1 mm increment until the margin was sufficient such that the prescribed $D_{98,95}$ for both CTVs is achieved.

2.C. Delineation uncertainties model—average-surface-of-standard-deviation (ASSD)

Coverage probability evaluations require a DU model. Creation of such a model poses practical challenges. One might create such a model by having multiple observers contour patient image sets multiple times, then use the resultant contour sets as input to a model to determine a mean contour with some variance. Besides being resource intensive, such a model might poorly represent uncertainties as those contouring may have similar training and bias.

To execute this study, we created an ASSD model which estimates the patient-specific DUs to determine the alternative ROI structures relative to the delineated contours. In other words, DUs for a ROI were quantified by the deviation between the physician contoured volume and alternative possible volumes. The contoured to alternative ROI variation was incorporated as a deformation, contraction, or expansion, corresponding to the delineator who delineated in a conservative versus liberal manner. The ASSD model assumed that image contrast perception was the source of contouring delineation variations, with the deformation being a function of image contrast in voxels adjacent to the contoured surface. The model seems reasonable for contouring prostate, seminal vesicles, bladder, and rectum in pelvis, based on the work from Baudry *et al.*²⁹

In the ASSD model, for each ROI voxel, the magnitude of the DUs was scaled by a voxel-specific delineated-to-

alternative-location displacement vector $\vec{D}$ as

$$\vec{D} = (D_x, D_y, D_z), \quad (1)$$

where D_x, D_y, D_z are x [= right-left (RL)], y [= posterior-anterior (PA)], and z [= superior-inferior (SI)] components of $\vec{D}$. With r representing (x, y, z) , $(D_r)_{i,k}$ is the DUs' component of the ROI voxel i for alternative ROI structure with index k . $(D_r)_{i,k}$ is written as

$$(D_r)_{i,k} = (F_{SD})_{r,k} \cdot (F_{CT})_{r,i}, \quad (2)$$

where $(F_{SD})_{r,k}$ is a variable representing the base contouring uncertainty and $(F_{CT})_{r,i}$ is an individualized factor determined by the CT image intensity gradient on the location at voxel i . $(F_{SD})_{r,k}$ is assumed to vary with the different alternative structure as

$$(F_{SD})_{r,k} = \text{norminv}((p_{\text{Rand}})_k, 0, SD_r), \quad (3)$$

where $\text{norminv}()$ is an inverse standard normal cumulative density function and $(p_{\text{Rand}})_k$ is a random probability sampled from a uniform distribution with range from 0 to 1 for the alternative structure k . 0 and SD_r are the Gaussian mean and the SD in r direction. The mean of the Gaussian distribution was zero because the average of the delineated ROI surface is assumed the best available estimate of the true ROI surfaces. The SD values of the Gaussian distribution $RL(x)$, $PA(y)$, and $SI(z)$ are determined empirically based on the literature in Table I.

$(F_{CT})_{r,i}$ is a function of CT image intensity gradient for voxel i . The expression of $(F_{CT})_{r,i}$ depended on voxel location. If voxel i is on the surface of ROI, $(F_{CT})_{r,i}$ in the r direction is written as

$$(F_{CT})_{r,i} = \frac{a}{|(\text{grad}_{CT})_{r,i}| + a}, \quad (4)$$

where $(\text{grad}_{CT})_{r,i}$ is the CT number gradient of voxel i in r direction, $|\cdot|$ denotes its absolute value, and a is a constant gradient use to normalize the maximum $(F_{CT})_{r,i}$ to 1. For extremely poor contrast with zero $|(\text{grad}_{CT})_{r,i}|$, $(F_{CT})_{r,i} = 1$. For ideal image contrast with sharp gradient $(|(\text{grad}_{CT})_{r,i}| \gg a)$, $(F_{CT})_{r,i} = 0$. If a is too small, e.g., 1,

F_{CT} will be only sensitive to very low gradient, e.g., up to only 20 HU/mm, while in fact, the delineator's eyes may not catch the difference within this low range of gradient very well. If a is too large such as 10 000, the F_{CT} may not change with most gradients of interest since they are negligible to a . Here, $a = 50$ was selected as a reasonable base gradient. The constant a was selected such that it fell in a range where a did not affect the sensitivity of $(F_{CT})_{r,i}$ to various $|(\text{grad}_{CT})_{r,i}|$.

The voxel-specific $(\text{grad}_{CT})_{r,i}$ is obtained from equation

$$(\text{grad}_{CT})_{r,i} = \frac{CT_{r,i}^+ - CT_{r,i}^-}{r_i^+ - r_i^-} = \frac{CT_{r,i}^+ - CT_{r,i}^-}{2 \cdot \text{VoxelSize}_r}, \quad (5)$$

where $(\text{grad}_{CT})_{r,i}$ is the ratio of CT number difference $CT_{r,i}^+ - CT_{r,i}^-$ and position difference $r_i^+ - r_i^-$ of two neighboring voxels in positive(+) r and negative(-) r sides of voxel i . $r_i^+ - r_i^-$ is equivalent to two times the voxel size in the r direction. Analysis of our patient data, we find that in the rectum-CTV_{prostate} boundary region, the $(\text{grad}_{CT})_{r,i}$ ranges from 0 to higher than 500 HU/mm, with average about 90, 90, 40 HU/mm in x, y, z directions.

When voxel i is an interior ROI voxel, its $(F_{CT})_{r,i}$ is interpolated from its two nearest surface voxels, denoted as $i_{\text{surf}+}$ and $i_{\text{surf}-}$, in the positive and the negative r directions,

$$(F_{CT})_{r,i} = (F_{CT})_{r,i_{\text{surf}-}} + \frac{((F_{CT})_{r,i_{\text{surf}+}} - (F_{CT})_{r,i_{\text{surf}-}})}{r_{i_{\text{surf}+}} - r_{i_{\text{surf}-}}} \cdot (r_i - r_{i_{\text{surf}-}}). \quad (6)$$

An illustration of F_{CT} -dependent $\vec{D}$ vectors (with F_{SD} vectors of magnitude of one in an expansion direction) on a transverse slice of a rectum is shown in Fig. 1. Lower gradient (= poor image contrast) is associated with a larger DU displacement vector.

2.D. Treatment planning decision flow

A treatment planning decision flow was designed to (1) validate that an uncertainty compensation strategy (e.g., FM, CP_{OM}, or CP_{COP}) was needed to accommodate the dosimetric

FIG. 1. CT image slice in transverse plane for rectum (solid-line contour) and corresponding $(D_r)_{i,k} = (F_{SD})_{r,k} \cdot (F_{CT})_{r,i}$ with $|(F_{SD})_{r,k}| = 1$ for display purposes.

variability caused by contour DUs and (2) address which plan was the most preferred among the FM, CP_{OM}, and CP_{COP} plans.

2.D.1. Zero margin evaluation

To determine the dosimetric effect of ASSD-modeled DUs, a zero-PTV-margin IMRT plan with PTV1 = CTV_{prostate} and PTV2 = CTV_{SV} was evaluated in terms of $D_{98,95}$ for both CTV structures via the alternative ROI simulations. Denote $D_{98,95,Rx}$ as the prescribed dose (78 Gy for CTV_{prostate} and 66 Gy for CTV_{SV}). If the inherent dosimetric-margin³⁰ of the treatment plan exceeds the ASSD-modeled CTV variability for both CTV_{prostate} and CTV_{SV}, then $D_{98,95} \geq D_{98,95,Rx}$, the optimal margin would be zero, and a contour uncertainty compensation strategy would not be needed. In addition to coverage probabilities, two additional quantities were computed for the CTVs of all the zero-PTV plans. One was the percentage dose degradation $\% \Delta D_{98,95}$, which was the percentage difference of $D_{98,95}$ relative to $D_{98,95,Rx}$ when $D_{98,95} < D_{98,95,Rx}$ or 0 when $D_{98,95} \geq D_{98,95,Rx}$. The other was the percentage DVH variability $\% \Delta DVH_{98,5-95}$, the percentage difference between $D_{98,95}$ and $D_{98,5}$. This quantity indicates the spread of DVHs obtained from different alternative ROI structures.

2.D.2. FM vs CP_{OM} vs CP_{COP}

FM, CP_{OM}, and CP_{COP} plans were created and compared to each other to determine which planning technique was most preferred. The primary criteria of the most preferred plan were the target coverage probabilities corresponding to the optimization objectives, i.e., CTV_{prostate} $D_{98,95} \geq 78$ Gy, CTV_{SV} $D_{98,95} \geq 66$ Gy. Considering the uncertainty of coverage calculation (see Sec. 4), an optimized dose slightly lower than coverage objective is regarded as acceptable (within 1%). For each final plan for each method, the $P+$ (probability of complication free control) value representing TCP/NTCP distributions of the 1000 virtual treatment courses was examined for a secondary comparison. The equations and parameters to calculate $P+$ can be found in Xu *et al.*²⁰

3. RESULTS

3.A. Zero margin evaluation

As Fig. 2 illustrates, the inherent dosimetric margin was insufficient to yield acceptable coverage probability for all of the zero-PTV-margin IMRT plans, although it did yield sufficient dose spillage to ensure CTV_{SV} coverage for three patients (ID 6–8). The dose degradation $\% \Delta D_{98,95}$ due to DUs ranged from 3.2% to 12.2% for CTV_{prostate} and from 0% to 12.9% for CTV_{SV}. The resulting $\% \Delta DVH_{98,5-95}$ was 2.9%–13.6% for CTV_{prostate} and 3.8%–13.8% for CTV_{SV}.

3.B. FM vs CP_{OM} vs CP_{COP}

For each of the 19 patients, FM, CP_{COP}, and CP_{OM} plans were created. The resultant patient-specific uniform margins

FIG. 2. (a) Percentage degraded dose ($\% \Delta D_{98,95}$) and (b) DVH variability ($\% \Delta DVH_{98,5-95}$) at prescribed dose of CTV_{prostate} (black bins) and CTV_{SV} (upward diagonal bins) for 19 zero-PTV-margin plans with delineation uncertainties considered using the ASSD model.

from CP_{OM} are listed in Table II. The optimized PTV margin for the target structures varies with patient: 2–5 mm for CTV_{prostate} and 0–4 mm for CTV_{SV}.

The comparison results of CP_{COP}, CP_{OM}, and FM plans in terms of their target coverage probabilities and $P+$ values are summarized in Table III. CP_{COP} plans were preferred for 11/19

TABLE II. The optimized PTV margins of CTV_{prostate} (P) and CTV_{SV} (S) obtained by CP_{OM} technique for 19 prostate cancer patients. The margins are quite different from the fixed PTV margins used in FM technique, which are 4 mm (RL), 5 mm (PA), 6 mm (SI) for CTV_{prostate} and 4 mm (RL), 5 mm (PA), 7 mm (SI) for CTV_{SV}.

Patient ID	COP _{OM} PTV (mm)
1	$P: 3 \quad S: 2$
2	$P: 5 \quad S: 2$
3	$P: 4 \quad S: 3$
4	$P: 3 \quad S: 2$
5	$P: 3 \quad S: 2$
6	$P: 2 \quad S: 2$
7	$P: 3 \quad S: 0$
8	$P: 3 \quad S: 2$
9	$P: 5 \quad S: 2$
10	$P: 5 \quad S: 2$
11	$P: 3 \quad S: 3$
12	$P: 5 \quad S: 2$
13	$P: 5 \quad S: 2$
14	$P: 3 \quad S: 2$
15	$P: 4 \quad S: 3$
16	$P: 3 \quad S: 2$
17	$P: 3 \quad S: 2$
18	$P: 5 \quad S: 4$
19	$P: 2 \quad S: 2$

TABLE III. Patient ID, best planning technique and its gain relative to the other two plans in terms of target dose coverage ($D_{98,95}$), and probability of complication free control $P+$.

ID	Best plan	Gain relative to the other plans	
1	CP _{OM}	CP _{COP} (+10.0% $P+$)	FM (+6.4% $P+$)
2	CP _{COP}	CP _{OM} (+17.5% $P+$)	FM (+16.5% $P+$)
3	CP _{COP}	CP _{OM} (+4.8% $P+$)	FM (+9.2% $P+$)
4 ^a	CP _{OM}	CP _{COP} (+3.5% CTV _{prostate} $D_{98,95}$)	FM (+6.9% $P+$)
5	CP _{OM}	CP _{COP} (+1.0% $P+$)	FM (+7.1% $P+$)
6	CP _{OM}	CP _{COP} (+2.4% $P+$)	FM (+22.0% $P+$)
7 ^a	CP _{OM}	CP _{COP} (+2.5% CTV _{SV} $D_{98,95}$)	FM (+10.8% $P+$)
8	CP _{COP}	CP _{OM} (+6.5% $P+$)	FM (+12.6% $P+$)
9	CP _{COP}	CP _{OM} (+15.1% $P+$)	FM (+15.1% $P+$)
10	CP _{COP}	CP _{OM} (+14.2% $P+$)	FM (+16.2% $P+$)
11	CP _{OM}	CP _{COP} (+4.6% $P+$)	FM (+7.9% $P+$)
12	CP _{COP}	CP _{OM} (+9.0% $P+$)	FM (+7.0% $P+$)
13	CP _{COP}	CP _{OM} (+13.4% $P+$)	FM (+14.4% $P+$)
14	CP _{COP}	CP _{OM} (+5.0% $P+$)	FM (+16.1% $P+$)
15	CP _{COP}	CP _{OM} (+5.7% $P+$)	FM (+8.6% $P+$)
16	CP _{COP}	CP _{OM} (+7.0% $P+$)	FM (+16.4% $P+$)
17	CP _{OM}	CP _{COP} (+0.2% $P+$)	FM (+10.4% $P+$)
18 ^a	FM	CP _{COP} (+2.4% CTV _{prostate} $D_{98,95}$)	CP _{OM} (+4.7% $P+$)
19	CP _{COP}	CP _{OM} (+4.6% $P+$)	FM (+10.3% $P+$)

^aCOP plan that fails to achieve target $D_{98,95}$.

cases, followed by CP_{OM} (7/19) and FM (1/19). Preference was mainly based on improved $P+$ indicating superior normal tissue sparing.

Comparing to only the FM plans, CP plans were preferred plans because of the $P+$ gain, which ranged from 6.4% to 22.0%. This indicated that the PTV margins used in FM were oversized. By shaving the redundant margin using coverage-based objectives, CP plans in most cases successfully reduced the OAR dose while maintaining the desirable target coverage. The only exceptions were 3/19 plans (patient IDs 4, 7, and 18) which failed to achieve CTV_{prostate} or CTV_{SV} coverage objective. The resultant $D_{98,95}$ was up to 3.5% lower than the objective dose. CP_{COP}'s coverage failure is due to the fact that this method alone considered the OAR coverage probabilities in the objective function. The objective-weighted resultant plan sacrificed target coverage in favor of normal tissue sparing. The CTV coverage objectives could be met by increasing their relative weight in the cumulative objective function. In direct comparisons between CP_{COP} and CP_{OM}, CP_{COP} reduced the OAR dose(s) further as a result of balancing the weighted OAR objectives. The average $P+$ gain for the 11 preferred CP_{COP} plans was 9.3%. This gain was larger than that from CP_{OM} plans; the average $P+$ improvement for the seven preferred CP_{OM} plans was 3.4%. Though less beneficial in normal tissue sparing, CP_{OM} plans guaranteed target coverage since achieving coverage indicated PTV margin sufficiency. In general, the relative advantages between three plans were patient-specific for the 19 patients.

4. DISCUSSION

The evaluation or optimization of treatment planning should include the information of the coverage probability of achieving isodose coverage in the presence of geometric

uncertainties. By using a CP technique, a patient treatment plan can be quality-assured with respect to the plan's robustness to geometric uncertainties and altered so that the desired target coverage is achieved while sparing adjacent normal tissue sparing.

This work proposed CP techniques for treatment planning evaluation and optimization to account for DUs for prostate cancer patients. The dosimetric influence of the geometric uncertainties other than DUs is beyond the scope of this paper. With the advancement of immobilization methods and IGRT techniques, the geometric uncertainties other than DUs (such as setup uncertainties and interfraction organ motions) should continue to decrease, leaving the contour delineation as an increasingly important uncertainty source. This work addressed the question as to what can be done to savvily account for DUs in the foreseeable scenario that other geometric uncertainties are well compensated.

This work utilized a new ASSD model to account for DUs. To overcome the lack of sufficient repeatedly delineated contours, the quasitrue ROI contour was assumed to follow population-based Gaussian distribution near the delineated contour. The Gaussian distribution utilized parameters from published papers on contouring difference by other researchers^{5,7,9,22,27} where statistics were more representative than could be obtained by the limited repeatedly delineated contours. Similar to a study which evaluated the delineation accuracy (particularly for small and low contrast lesions) under the influence of imaging quality from using different CT scanners,³¹ our ASSD model assumed that the DUs of each voxel were associated with the CT image contrast. The CT gradient near the voxel was calculated to scale the magnitude of DUs. Larger CT gradient (and equivalently better image contrast) rendered smaller deviation from the delineated to the true anatomies.¹⁶ In the future, this model could be further refined

FIG. 3. Axial, sagittal, and coronal dose distributions of CP_{COP} versus CP_{OM} versus FM plans on one slice for patient 8 when delineation uncertainties are considered. The thick isodose surfaces are the 78 Gy (seashell) TV for the $CTV_{prostate}$ (red colorwash) and the 66 Gy (aquamarine) TV for CTV_{SV} (green colorwash). The yellow and magenta colorwashes are the bladder and rectum. The thin isodose surfaces are 75 Gy (maroon), 69 Gy (slateblue), 57 Gy (lightblue), and 40 Gy (green). Compared to CP_{OM} and FM plans, the CP_{COP} dosimetric margin (the distance between the $CTV_{prostate}$ and the 78 Gy TV) on the sides of bladder and rectum is, in general, smaller due to CP_{COP} 's inclusion of OAR uncertainties in the objective function. This explains why a CP_{COP} plan is more likely to have better normal tissue sparing and sometimes lower target coverage.

by using, e.g., three-dimensional-orientation-dependent SD for Gaussian-based interobserver contour differences.

Inherent limitations of our methods include (a) the assumption that the quasitru ROI contour represents the true, or more appropriately mean ROI from multiple contours and (b) validity of the ASSD model. Assumption (a) is justified by the fact that we have no other method to estimate the true ROI; therefore, we apply the common technique of using the available measurement as the best estimate of the mean. With respect to the validity of the ASSD model, the ASSD only needs to follow the trend of contour variability for our conclusions to be useful. Specifically, if ASSD underestimates true DUs, the dosimetric influence of DUs is larger (and more significant) than what we report. On the other hand, if ASSD overestimates DUs, the clinical benefit of CP technique relative to FM techniques in terms of normal tissue sparing will be more pronounced, as a result of the oversized margin used in FM. Regardless, the ASSD model has its value in

allowing an initial study on DUs and the advanced planning technique to accommodate them.

In isolation, the dosimetric impact of DUs based on the ASSD model is non-negligible and degrades the patient treatment plan. While we have not evaluated the contribution of contouring uncertainties in relation to other geometric uncertainties and their compensation strategies, we acknowledge that it is reasonable to assume that some level of protection may come from the margins for other geometric uncertainties. However, as other geometric uncertainties are reduced, the fractional contribution from contouring DUs will increase, therefore, requiring more attention during the treatment planning process. For example, assume that DUs, deformable organ motions, and setup uncertainties are independent and their effects can be added in quadrature. If the margin from the VHMF, applied for setup-errors with standard deviations $\sigma = 2$ mm (random) and $\Sigma = 2$ mm (systematic), combines in quadrature with an OM = 4 mm for the possible

organ deformations,²⁰ and OM = 4 mm for contouring DUs, the combined margin increases only from 8 to 9 mm. If, however, setup error $\sigma = 1$ mm and $\Sigma = 1$ mm and deformable organ motion OM = 2 mm, the margin increases by more than 40% from 3.8 to 5.4 mm to accommodate DUs. Our future work is to study the dosimetric consequences of all the geometric uncertainties—setup uncertainties, organ motion and deformation, and delineation uncertainties.

From Table III, CP techniques turned out to be beneficial in adapting DUs to patient-specific coverage probabilities for high-risk prostate cancer treatment. The VHMF-based PTV margins used in FM, in contrast, were often much larger than needed on per patient basis. The inadequacy of VHMF-based PTV margins has also been reported for other geometric uncertainties (e.g., setup uncertainties^{18,32,33} and deformable interfraction motions²⁰).

The relative advantages between CP_{COP}, CP_{OM}, and FM were patient-specific. The characteristics of different dose distributions of these three plans for one patient are illustrated in Fig. 3. For the CP_{COP} plan, the dosimetric margin between CTV_{prostate} and the 78 Gy treated volume (TV) was smallest in the region of bladder due to the competing target and OAR objectives. CP_{OM} had less burden of OAR sparing and the optimized PTV can be continuously enlarged when it is needed to boost the current prescribed target coverage.

Since the initial planning objectives were based on a FM clinical standard, alternative CP optimization objectives and weights may show more potential clinical benefit. The complexity of the plan segments needed for delivery of FM, CP_{COP}, or CP_{OM} plan has been reviewed by comparing the fractional monitor units of each plan (Table IV). A larger

number indicates an increased complexity (number monitor units and/or segments) needed for delivery. In most cases, CP_{OM} plan was least complex among the three, while CP_{COP} and FM plans were comparable for most of the patients.

This coverage probability estimation based on simulating 1000 alternative ROI structures was assumed to be sufficient. To test this assumption, tests were performed on two arbitrarily selected patients. The accuracy was checked by calculating $\Delta D_{v,q}$, the absolute percentage difference of dose volume metric D_v corresponding coverage probability q using 10, 100, 200, and 500 alternative ROI structures to see how q converges relative to that using 1000 courses. For CTV_{prostate} and CTV_{SV}, $\Delta D_{98,95}$ decreased to 0.2% as the number of alternative structures increases from 10 to 500. For the bladder and rectum, the absolute percentage dose difference was mostly below 0.5%. The precision was tested by reproducing each $D_{v,q}$ based on 1000 alternative ROI sets for 7–10 times. The maximum absolute percentage difference among all the repeated estimated values was 0.3% and 0.2% for CTV_{prostate} and CTV_{SV}, respectively, and less than 1% for the OARs. These test cases supported that coverage estimation using 1000 alternative structures was adequate.

5. CONCLUSION

To ensure desirable target coverage and normal tissue sparing, the dosimetric variations of delineation uncertainties should be considered in treatment planning for the high-risk prostate patients treated by external beam radiation therapy. An ASSD model to simulate delineation uncertainties gave a quantitative evaluation of the plan quality. With this model incorporated, CP_{COP} or CP_{OM} techniques resulted in either better target dose coverage probability or lower normal tissue doses than the FM technique that uses empirical fixed PTV margins, thereby providing a theoretical benefit in terms of complication free control.

ACKNOWLEDGMENTS

This work is supported in part by NIH Grant No. P01CA116602 and by a research contract with Philips Medical Systems. The authors thank the Netherlands Cancer Institute for providing the image sets, Ford Sleeman and Dr. Mirek Fatyga for the initial processing of the image sets, and Dr. Elizabeth Weiss for contouring the image sets.

^{a)}Author to whom correspondence should be addressed. Electronic mail: hjxu111@gmail.com; Telephone: 443-546-1316; Fax: 443-546-1303; Present address: Department of Radiation Oncology, Central Maryland Radiation Oncology Center, University of Maryland School of Medicine, 10710 Charter Dr., Suite G030, Columbia, Maryland 21044.

¹C. F. Njeh, L. Dong, and C. G. Orton, "IGRT has limited clinical value due to lack of accurate tumor delineation," *Med. Phys.* **40**, 040601 (4pp.) (2013).

²E. Weiss and C. F. Hess, "The impact of gross tumor volume (GTV) and clinical target volume (CTV) definition on the total accuracy in radiotherapy theoretical aspects and practical experiences," *Strahlenther. Onkol.* **179**, 21–30 (2003).

³D. P. Huyskens, P. Maingon, L. Vanuytsel, V. Remouchamps, T. Roques, B. Dubray, B. Haas, P. Kunz, T. Coradi, R. Bühlman, R. Reddick, A. V.

TABLE IV. The monitor units (MUs) per fraction for CP_{COP} plans and percentage deviation from CP_{COP} for the CP_{OM} and FM plans. Summary statistics indicate that averaged over all patients, the intermethod MUs are well one standard deviation.

ID	CP _{COP}	CP _{OM} (%)	FM (%)
1	572	−28	−22
2	662	−21	−17
3	601	−7	5
4	552	−14	−9
5	589	−22	−15
6	442	−14	4
7	401	−4	3
8	400	−6	5
9	477	−6	−4
10	498	−3	1
11	544	−4	10
12	559	−4	−9
13	601	−19	−17
14	461	−4	10
15	441	0	2
16	451	−4	−2
17	600	−16	−4
18	567	−5	−4
19	417	−3	4
Avg. ± SD(MU)	518 ± 80	464 ± 57	497 ± 61

- Esch, and E. Salamon, "A qualitative and a quantitative analysis of an auto-segmentation module for prostate cancer," *Radiother. Oncol.* **90**, 337–345 (2009).
- ⁴X. Gu, B. Dong, J. Wang, J. Yordy, L. Mell, X. Jia, and S. B. Jiang, "A contour-guided deformable image registration algorithm for adaptive radiotherapy," *Phys. Med. Biol.* **58**, 1889–1901 (2013).
- ⁵M. van Herk, "Errors and margins in radiotherapy," *Semin. Radiat. Oncol.* **14**, 52–64 (2004).
- ⁶A. K. Bhardwaj, T. S. Kehwar, S. K. Chakarvarti, G. J. Sastri, A. S. Oinam, G. Pradeep, V. Kumar, M. Indranil, and S. C. Sharma, "Variations in inter-observer contouring and its impact on dosimetric and radiobiological parameters for intensity-modulated radiotherapy planning in treatment of localised prostate cancer," *J. Radiother. Pract.* **7**, 77–88 (2008).
- ⁷C. Fiorino, M. Reni, A. Bolognesi, G. M. Cattaneo, and R. Calandrino, "Intra- and inter-observer variability in contouring prostate and seminal vesicles: Implications for conformal treatment planning," *Radiother. Oncol.* **47**, 285–292 (1998).
- ⁸P. Petric, J. Dimopoulos, C. Kirisits, D. Berger, R. Hudej, and R. Pötter, "Inter- and intraobserver variation in HR-CTV contouring: Intercomparison of transverse and paratransverse image orientation in 3D-MRI assisted cervix cancer brachytherapy," *Radiother. Oncol.* **89**, 164–171 (2008).
- ⁹E. Weiss, J. Wu, W. Sleeman, J. Bryant, P. Mitra, M. Myers, T. Ivanova, N. Mukhopadhyay, V. Ramakrishnan, M. Murphy, and J. Williamson, "Clinical evaluation of soft tissue organ boundary visualization on cone-beam computed tomographic imaging," *Int. J. Radiat. Oncol., Biol., Phys.* **78**, 929–936 (2010).
- ¹⁰H. Vorwerk, G. Beckmann, M. Bremer, M. Degen, B. Dietl, R. Fietkau, T. Gsänger, R. M. Hermann, M. K. Alfred Herrmann, U. Höller, M. van Kampen, W. Körber, B. Maier, T. Martin, M. Metz, R. Richter, B. Siekmeyer, M. Steder, D. Wagner, C. F. Hess, E. Weiss, and H. Christiansen, "The delineation of target volumes for radiotherapy of lung cancer patients," *Radiother. Oncol.* **91**, 455–460 (2009).
- ¹¹ICRU, "Prescribing, recording and reporting photon beam therapy," ICRU Report 50 (ICRU Publications, Bethesda, MD, 1994).
- ¹²ICRU, "Prescribing, recording and reporting photon beam therapy (Supplement to ICRU Report 50)," ICRU Report 62 (ICRU Publications, Bethesda, MD, 2000).
- ¹³J. C. Stroom, P. C. Koper, G. A. Korevaar, M. van Os, M. Janssen, H. C. de Boer, P. C. Levendag, and B. J. Heijmen, "Internal organ motion in prostate cancer patients treated in prone and supine treatment position," *Radiother. Oncol.* **51**, 237–248 (1999).
- ¹⁴J. Stroom and B. Heijmen, "Limitations of the planning organ at risk volume (PRV) concept," *Int. J. Radiat. Oncol., Biol., Phys.* **66**, 279–286 (2006).
- ¹⁵M. van Herk, P. Remeijer, C. Rasch, and J. V. Lebesque, "The probability of correct target dosage: Dose-population histograms for deriving treatment margins in radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **47**, 1121–1135 (2000).
- ¹⁶C. Njeh, "Tumor delineation: The weakest link in the search for accuracy in radiotherapy," *J. Med. Phys.* **33**, 136–140 (2008).
- ¹⁷A. McKenzie, M. van Herk, and B. Mijnheer, "Margins for geometric uncertainty around organs at risk in radiotherapy," *Radiother. Oncol.* **62**, 299–307 (2002).
- ¹⁸J. J. Gordon and J. V. Siebers, "Coverage-based treatment planning: Optimizing the IMRT PTV to meet a CTV coverage criterion," *Med. Phys.* **36**, 961–973 (2009).
- ¹⁹J. J. Gordon, N. Sayah, E. Weiss, and J. V. Siebers, "Coverage optimized planning: Probabilistic treatment planning based on dose coverage histogram criteria," *Med. Phys.* **37**, 550–563 (2010).
- ²⁰H. Xu, D. J. Vile, M. Sharma, J. J. Gordon, and J. V. Siebers, "Coverage-based treatment planning to accommodate deformable organ variations in prostate cancer treatment," *Med. Phys.* **41**, 101705 (14pp.) (2014).
- ²¹S. M. Zhou, G. C. Bentel, C. G. Lee, and M. S. Anscher, "Differences in gross target volumes on contrast vs. noncontrast CT scans utilized for conformal radiation therapy treatment planning for prostate carcinoma," *Int. J. Radiat. Oncol., Biol., Phys.* **42**, 73–78 (1998).
- ²²C. Rasch, I. Barillot, P. Remeijer, A. Touw, M. van Herk, and J. V. Lebesque, "Definition of the prostate in CT and MRI: A multi-observer study," *Int. J. Radiat. Oncol., Biol., Phys.* **43**, 57–66 (1999).
- ²³Z. Gao, D. Wilkins, L. Eapen, C. Morash, Y. Wassef, and L. Gerig, "A study of prostate delineation referenced against a gold standard created from the visible human data," *Radiother. Oncol.* **85**, 239–246 (2007).
- ²⁴P. Remeijer, C. Rasch, J. V. Lebesque, and M. van Herk, "A general methodology for three-dimensional analysis of variation in target volume delineation," *Med. Phys.* **26**, 931–940 (1999).
- ²⁵J. G. Korporaal, C. A. T. van den Berg, G. Groenendaal, M. R. Moman, M. van Vulpen, and U. A. van der Heide, "The use of probability maps to deal with the uncertainties in prostate cancer delineation," *Radiother. Oncol.* **94**, 168–172 (2010).
- ²⁶J. Wu, M. J. Murphy, E. Weiss, W. C. Sleeman, and J. Williamson, "Development of a population-based model of surface segmentation uncertainties for uncertainty-weighted deformable image registrations," *Med. Phys.* **37**, 607–614 (2010).
- ²⁷K. E. I. Deurloo, R. J. H. M. Steenbakkers, L. J. Zijp, J. A. de Bois, P. J. C. M. Nowak, C. R. N. Rasch, and M. van Herk, "Quantification of shape variation of prostate and seminal vesicles during external beam radiotherapy," *Int. J. Radiat. Oncol., Biol., Phys.* **61**, 228–238 (2005).
- ²⁸M. Sharma, E. Weiss, and J. V. Siebers, "Dose deformation-invariance in adaptive prostate radiation therapy: Implication for treatment simulations," *Radiother. Oncol.* **105**, 207–213 (2012).
- ²⁹M. Baudry, A. Benlakhel, and N. Vincent, A Contour Extraction Method Based on Uncertainty Management and Fuzzy Modeling (2001), pp. 1461–1466, <http://www.wseas.us/e-library/conferences/malta2001/papers/146.pdf>.
- ³⁰J. J. Gordon and J. V. Siebers, "Evaluation of dosimetric margins in prostate IMRT treatment plans," *Med. Phys.* **35**, 569–575 (2008).
- ³¹Y. W. Ho, W. K. R. Wong, S. K. Yu, W. W. Lam, and H. Geng, "Accuracy in contouring of small and low contrast lesions: Comparison between diagnostic quality computed tomography scanner and computed tomography simulation scanner-a phantom study," *Med. Dosim.* **37**, 401–405 (2012).
- ³²J. J. Gordon, A. J. Crimaldi, M. Hagan, J. Moore, and J. V. Siebers, "Evaluation of clinical margins via simulation of patient setup errors in prostate IMRT treatment plans," *Med. Phys.* **34**, 202–214 (2007).
- ³³H. Xu, J. J. Gordon, and J. V. Siebers, "Sensitivity of postplanning target and OAR coverage estimates to dosimetric margin distribution sampling parameters," *Med. Phys.* **38**, 1018–1027 (2011).
- ³⁴Format $D_{V\%q}$ with $V\%$ being the volume percentage and q being the coverage probability.
- ³⁵Since q is the probability that the dose is greater than or equal to the objective value, $q = 5\%$ for OARs is equivalent to $q' = 95\%$, where q' is the confidence that the dose is less than or equal to the objective value.