Sensitivity of postplanning target and OAR coverage estimates to dosimetric margin distribution sampling parameters

Huijun Xu,[a] J. James Gordon, and Jeffrey V. Siebers
Department of Radiation Oncology, Virginia Commonwealth University, Richmond, Virginia 23298

(Received 28 July 2010; revised 21 December 2010; accepted for publication 23 December 2010; published 31 January 2011)

Purpose: A dosimetric margin (DM) is the margin in a specified direction between a structure and a specified isodose surface, corresponding to a prescription or tolerance dose. The dosimetric margin distribution (DMD) is the distribution of DMs over all directions. Given a geometric uncertainty model, representing inter- or intrafraction setup uncertainties or internal organ motion, the DMD can be used to calculate coverage Q, which is the probability that a realized target or organ-at-risk (OAR) dose metric D_v exceeds the corresponding prescription or tolerance dose. Postplanning coverage evaluation quantifies the percentage of uncertainties for which target and OAR structures meet their intended dose constraints. The goal of the present work is to evaluate coverage probabilities for 28 prostate treatment plans to determine DMD sampling parameters that ensure adequate accuracy for postplanning coverage estimates.

Methods: Normally distributed interfraction setup uncertainties were applied to 28 plans for localized prostate cancer, with prescribed dose of 79.2 Gy and 10 mm clinical target volume to planning target volume (CTV-to-PTV) margins. Using angular or isotropic sampling techniques, dosimetric margins were determined for the CTV, bladder and rectum, assuming shift invariance of the dose distribution. For angular sampling, DMDs were sampled at fixed angular intervals ω (e.g., $\omega = 1°, 2°, 5°, 10°, 20°$). Isotropic samples were uniformly distributed on the unit sphere resulting in variable angular increments, but were calculated for the same number of sampling directions as angular DMDs, and accordingly characterized by the effective angular increment ω_{eff}. In each direction, the DM was calculated by moving the structure in radial steps of size $\delta(=0.1, 0.2, 0.5, 1$ mm) until the specified isodose was crossed. Coverage estimation accuracy ΔQ was quantified as a function of the sampling parameters ω or ω_{eff} and δ.

Results: The accuracy of coverage estimates depends on angular and radial DMD sampling parameters ω or ω_{eff} and δ, as well as the employed sampling technique. Target $|\Delta Q| < 1\%$ and OAR $|\Delta Q| < 3\%$ can be achieved with sampling parameters ω or $\omega_{\text{eff}} = 20°$, $\delta = 1$ mm. Better accuracy (target $|\Delta Q| < 0.5\%$ and OAR $|\Delta Q| < \sim 1\%$) can be achieved with ω or $\omega_{\text{eff}} = 10°$, $\delta = 0.5$ mm. As the number of sampling points decreases, the isotropic sampling method maintains better accuracy than fixed angular sampling.

Conclusions: Coverage estimates for post-planning evaluation are essential since coverage values of targets and OARs often differ from the values implied by the static margin-based plans. Finer sampling of the DMD enables more accurate assessment of the effect of geometric uncertainties on coverage estimates prior to treatment. DMD sampling with ω or $\omega_{\text{eff}} = 10°$ and $\delta = 0.5$ mm should be adequate for planning purposes. [DOI: 10.1118/1.3544364]

Key words: sampling parameters, dosimetric margin distribution, coverage estimates, planning evaluation, prostate cancer

I. INTRODUCTION

In conventional radiation therapy treatment planning, geometric uncertainties such as setup uncertainties, delineation uncertainties, and internal motion[1] are accounted for at the start of the planning process by adding clinical target volume to planning target volume (CTV-to-PTV) margins to the CTV and planning organ-at-risk (OAR) volumes (PRVs) to OARs. The CTV, PTV, and PRV concepts, defined in ICRU Report Nos. 50 (Ref. 2) and 62,[3] are familiar to treatment planners. However, the appropriate margin sizes to ensure target coverage or normal tissue sparing for different treatment sites remain an area for physician judgment. There are many publications that determine margin sizes based on planning studies and simulations. (A ScienceDirect search for "radiotherapy margin studies" returned 31 000+results.) Examples include Muren *et al.*,[4] Lee *et al.*,[5] Button *et al.*,[6] and Ekberg *et al.*[7] Additionally, margin sizing formulas (i.e., guidelines) have been developed by Stroom *et al.*,[8] van Herk *et al.*,[9] and others.[10] Regardless of the details, once a margin is defined, planners typically rely on the PTV to act as a surrogate for the CTV. Planned (or static) dose to the PTV is assumed to represent the dose received by the CTV when geometric uncertainties are accounted for.

Margins that are sized according to margin formulas can be expected to be reasonably accurate in ensuring target coverage or normal tissue sparing in many cases. However, margin formulas are approximations based on sampled data[8] or idealized models.[9] For instance, the margin formula of van Herk *et al.* is based on a model that assumes an error-function-shaped beam penumbra. Real beam penumbras do not fit this assumption, leading to possible inaccuracies in the calculated margins.

Furthermore, margin formulas are based on the assumption that relevant isodose surfaces (e.g., the surface for 95% of the prescription dose) will closely conform to the PTV. Proper margin sizing guarantees that the CTV remains within the PTV most of the time, ensuring the corollary that the CTV also remains within the isodose surface. However, planning rarely results in isodose surfaces that exactly conform to the PTV. Instead, isodose surfaces will touch the PTV in some areas but diverge from the surface in others. This implies a mismatch between the coverage provided by the planned dose distribution and the coverage implied by the PTV: Gaps between isodose surfaces and the PTV can cause the CTV to receive a different (sometimes higher) level of coverage than would be implied by the PTV.[11,12]

Additionally, physicians sometimes choose to shave margins where they impinge on an OAR. Here, the PTV no longer matches the assumptions made by the margin formulation and, consequently, the planned PTV dose might no longer accurately represent received CTV dose. Ultimately, coverage is a function of the margin between the CTV and the treated volume (TV) (the volume enclosed by a critical isodose surface) not the PTV. This is not a new observation. It motivates the conformity index defined in ICRU Report No. 62, and earlier studies that have directly evaluated target coverage through simulation.[13]

For the reasons given above, it is desirable to explicitly evaluate the coverage provided to CTVs and OARs and not simply assume that PTVs and PRVs achieve the nominal coverage that is implied by the PTV/PRV margin. The dosimetric margin (DM) and dosimetric margin distribution (DMD) are tools that can be used to evaluate coverage.[12] Clinically, they can be used to evaluate coverage probabilities for postplanning quality assurance (QA), specifically the probabilities that the desired target coverage and normal tissue structure sparing will be achieved. Additionally, although this option is not considered in the present work, the computed probabilities can be incorporated as an objective into a plan optimization to produce coverage optimized treatment plans.[14,15]

Intuitively, the DM is the distance in a specified direction between a structure—CTV or OAR—and a specified isodose surface. More precisely, the DM $M_{v,d}(\varphi,\theta)$ is the maximum distance that the structure can be offset in the specified direction, while still satisfying a dose constraint $D_v \geq d$ for CTV (or $D_v \leq d$ for OAR), where D_v is the dose to $v\%$ of the structure's volume and d is its planned goal. For example, for a CTV dose constraint $D_{\min} \geq 79.2$ Gy (equivalent to $D_{100} \geq 79.2$ Gy), $M_{\min,79.2}(\varphi,\theta)$ [or $M_{100,79.2}(\varphi,\theta)$] denotes the maximum distance the CTV can be offset in the direction (φ,θ) before its $D_{\min}$ falls below 79.2 Gy. Similarly, for an OAR dose constraint $D_{25} \leq 70$ Gy, $M_{25,70}(\varphi,\theta)$ denotes the maximum distance the OAR can be offset in the direction (φ,θ) before its D_{25} exceeds 70 Gy. The DMD is the distribution of DMs $M_{v,d}(\varphi,\theta)$ over all directions.

Once evaluated, the DMD can be used to calculate coverage probabilities for the CTV and OAR. Coverage is defined to be the percentage of uncertainties (e.g., the percentage of organ motions) for which $D_v \geq d$ (see Sec. II). As discussed above, computed coverage is preferred to the nominal coverage implied by PTV and PRV margins. Practical implementation requires the DMD to be sampled over a finite number of directions (φ,θ) using a finite step size along each direction. The accuracy or errors in the calculated dosimetric margins and derived coverage estimates are a function of the angular and distance increments that are utilized in the evaluation. By computing DMDs and coverage values for 28 prostate plans, this paper quantifies the sensitivity of coverage estimates to these sampling parameters. DMD sampling parameters that provide acceptable accuracy for plan evaluation are also determined.

II. METHODS AND MATERIALS

II.A. Prostate plans

This work utilized 28 anonymized prostate anatomies taken from a clinical database with Institutional Review Board approval. Treatment plans were for localized prostate cancer, so targeted only the prostate and proximal seminal vesicles. The CTV consisted of the GTV plus bilateral seminal vesicles lying within 10 mm of the prostate. The PTV was the CTV expanded uniformly by 10 mm. OARs consisted of the rectum, bladder, and femoral heads. Plans utilized seven coplanar (transverse) beams. Optimization criteria for the PTV were $D_{\min} \geq 79.2$ Gy and $D_{\max} \leq 84.7$ Gy. The target minimum dose of 79.2 Gy mirrored the high dose arm of Radiation Therapy Oncology Group (RTOG) Protocol 0126 (www.rtog.org).[16] OAR criteria for the bladder and rectum were taken from RTOG 0126. For the bladder, they were $D_{\max} \leq 84.7$ Gy, $D_{15} \leq 80$ Gy, $D_{25} \leq 75$ Gy, $D_{35} \leq 70$ Gy, and $D_{50} \leq 65$ Gy. For the rectum, they were $D_{\max} \leq 84.7$ Gy, $D_{15} \leq 75$ Gy, $D_{25} \leq 70$ Gy, $D_{35} \leq 65$ Gy, and $D_{50} \leq 60$ Gy. Left and right femoral heads were required to have $D_{\max} \leq 50$ Gy and $D_{50} \leq 35$ Gy. Plans were optimized using a research version (8.1w) of the Pinnacle treatment planning system (Philips Medical Systems, Fitchburg, WI) using standard direct machine parameter optimization and a $2 \times 2 \times 2$ mm dose grid. As the dose distributions in the final optimized plans do not explicitly incorporate setup errors, they are referred to as static plans.

II.B. Geometric uncertainty model

This work considered rigid body interfraction setup uncertainties (i.e., rigid body shifts of the patient anatomy with respect to the treatment beams). Ideally, it would be desirable to use uncertainty models derived from validated deforma-

tion models of the prostate anatomy; however, techniques for generating such models are not yet standardized and remain an active area for research. Given that the goal of the present paper is to find acceptable parameters for DMD evaluation, not to perform DMD evaluation for a specific group of patients, the use of rigid body uncertainties was considered acceptable. Following the model of van Herk *et al.*,[9] uncertainties were assumed to have systematic (preparation) and random (per fraction) components. Both components were assumed to be normally distributed with standard deviations (SDs) Σ (systematic) and σ (random) along each axis [right-left (RL), posterior-anterior (PA), and superior-inferior (SI)]. The 10 mm CTV-to-PTV margin used to design the static plan (above) is consistent with the margin formula of van Herk *et al.*, $M=2.5\Sigma+0.7\sigma\approx 10$ mm for $\Sigma=\sigma=3$ mm.

After optimization, the effect of geometric uncertainties on the planned dose distribution was evaluated. Random errors were incorporated into the plan evaluation by computing the planned dose distribution after convolving beam fluences with a normal kernel having SD $\sigma=3$ mm. This technique is commonly referred to as fluence convolution[17] and is designed to model the blurring effect of many (e.g., 30) random (per fraction) setup uncertainties on the dose distribution. The implementation used here was the same as in the previous studies.[12,14] Fluence convolution was enabled via a script command. When enabled, each beam's 2D fluence after all beam defining elements (jaws and MLC), but before the patient, is convolved with the abovementioned normal kernel using a fast Fourier transform convolution. The patient dose calculation then uses the convolved fluences in place of the original fluences. Fluence convolution has been found to be a reasonable approach to approximate the cumulative effect of random errors for treatment courses consisting of $\sim$30 fractions.[18,19]

Dosimetric margins were calculated for the CTV, bladder, and rectum relative to appropriate isodose surfaces. The dosimetric margins represent the "safety margin" for absorbing systematic uncertainties. That is, as described below, coverage could be calculated from the DMD using the systematic SD Σ only, random uncertainties having been accounted for through fluence convolution.

II.C. Dosimetric margin calculations

The following dosimetric margins were evaluated: CTV $M_{\mathrm{min},79.2}$, bladder $M_{25,75}$ and bladder $M_{25,80}$, and rectum $M_{25,70}$ and rectum $M_{25,75}$. Dosimetric margin calculations assumed shift invariance of the dose distribution, which has been shown to be reasonable for prostate plans.[18] (Shift invariance may not be a valid assumption for other sites, such as head and neck.) The dosimetric margin in each direction was calculated by moving the structure in steps of δ relative to the dose distribution, until the relevant dose metric (e.g., CTV D_{min}) crossed the specified dose value (e.g., until CTV D_{min} fell below 79.2Gy). The results were compared for step sizes $\delta=1$, 0.5, 0.2, and 0.1 mm.

In each direction, structures were stepped out to a maximum distance of 25 mm. The reason of this "cutoff" is that for OAR there are directions (e.g., away from the high dose region) for which the relevant dose metric will never cross the specified isodose. For the range of setup errors (Σ values) considered in this work, there is negligible probability of displacements greater than or equal to 25 mm. Therefore, the 25 mm cutoff was sufficiently large to ensure accurate coverage calculations.

DMDs were sampled over differing numbers of directions using two different techniques: Fixed angular increment (FAI) and isotropic sampling (ISO). The FAI method was applied to the CTV, bladder, and rectum. The ISO method was applied to the CTV only. In the FAI method, dosimetric margins were calculated for directions $\varphi=n\omega$ and $\theta=m\omega$, where φ is the elevation angle and θ is the azimuthal angle ($\varphi=90°$ being the patient's superior-inferior axis and $\theta=0°$ being the patient's right-left axis). The values $\omega=20°$, 10°, 5°, 2°, and 1° (for CTV) or $\omega=20°$, 10°, 4°, and 2° (for bladder and rectum) were used, and integers n and m were incremented from $-90/\omega$ to $90/\omega$ and from $-180/\omega$ to $180/\omega-1$, respectively. The number of FAI directions is a function of ω,

$$N_d = 2 + \left(\frac{180}{\omega} - 1\right)\cdot\frac{360}{\omega}, \tag{1}$$

where N_d ranges from 146 for $\omega=20°$ to 64 442 for $\omega=1°$. A feature of the FAI method is that the steradians covered by each angular segment become smaller as one approaches the poles (i.e., as φ goes to $\pm 90°$).

The ISO method used directions that were approximately isotropically distributed across 4π sr. There is no general analytic solution for sets of N_d isotropic directions covering the unit sphere. Consequently, the spinal point solution was employed.[20] Based on the required number of directions, this method gives a set of angles (φ,θ), which closely approximate an isotropic distribution. In comparing the FAI and ISO results, the same number of directions was employed. ISO results presented below are given as a function of the *effective* angular increment ω_{eff}, which is the fixed angular increment that would give the same number of sampling directions.

In calculating DMDs for a target (i.e., CTV), one is most often concerned with a dose metric (e.g., CTV D_{min}) starting above a dose threshold in the structure's static position and falling below the threshold as the structure moves out of the high dose region (type I target). However, it can sometimes happen that in a particular plan the dose metric starts off already below the desired threshold and can be increased above the threshold for some positional offsets (type II target). Conversely, for an OAR, one is most often concerned with a dose metric starting below a threshold and moving above the threshold as the structure moves into the high dose region (type I OAR). However, for some cases, the metric may start off above the desired threshold and fall below the threshold for some structure offsets (type II OAR).

The presence of type II structures indicates that the static plan does not meet the dose-volume criteria. For example, a CTV, which is of type II for the margin $M_{\mathrm{min},79.2}$, starts off in

the static plan having $D_{\min}<79.2$ Gy but may experience $D_{\min}\geq 79.2$ Gy for some offsets. A bladder, which is of type II for the margin $M_{25,75}$, starts off in the static plan having $D_{25}>75$ Gy but may experience $D_{25}\leq 75$ Gy for some offsets. The distinction between type I and type II cases is emphasized here primarily because it affects the details of the DMD calculation. In the "nonstandard' (type II) cases, dosimetric margins are still calculated as the maximum distance that the structure can move before crossing the dose threshold, but coverage formulas are modified as described below. Also, the type I/II distinction helps one understand why type II targets tend to have low coverage values and type II OAR high coverage values, both of which are undesirable situations. However, regardless of whether a structure is of type I or type II (for a given margin/criterion), the interpretation of coverage Q is unchanged.

II.D. Coverage calculation

For both targets and OAR, coverage $Q_{v,d}$ is defined to be the probability that a dose metric D_v exceeds dose d: $Q_{v,d}=\Pr[D_v\geq d]$. For a target dose constraint $D_v\geq d$, $Q_{v,d}$ is the probability that the constraint will be *satisfied*, so high coverage values (e.g., 95%) are desirable. For an OAR dose constraint $D_v\leq d$, $Q_{v,d}$ is the probability that the constraint will be *violated*, so low coverage values (e.g., 5%) are desirable. In the following, the symbol $Q_{v,d}$ is simplified to Q, where v and d are not explicitly required or are clear from the context.

Because random errors are accounted for via fluence convolution, Q is a function of the DMD and the standard deviation of systematic errors Σ. For normally distributed systematic uncertainties, the cumulative distribution function of 3D displacements r is given by[9]

$$F(r,\Sigma)\equiv \text{normcdf}\ \ 3\text{D}(r,\Sigma)$$
$$=\text{erf}(r/\sqrt{2}\Sigma)-(2/\sqrt{\pi})(r/\sqrt{2}\Sigma)\exp(-r^2/2\Sigma^2), \quad (2)$$

where erf() is the error function. Coverage in the direction (φ,θ) is given by

$$Q(\varphi,\theta,\Sigma)=\begin{cases}F(M_{v,d}(\varphi,\theta),\Sigma) & \text{type I target/type II OAR}\\ 1-F(M_{v,d}(\varphi,\theta),\Sigma) & \text{type II target/type I OAR,}\end{cases} \quad (3)$$

where $M_{v,d}(\varphi,\theta)$ is the dosimetric margin in direction (φ,θ). The overall coverage $Q(\Sigma)$ becomes

$$Q(\Sigma)=\sum_{\varphi,\theta} W_{\varphi,\theta}Q(\varphi,\theta,\Sigma), \quad (4)$$

where $W_{\varphi,\theta}$ is a weighting factor equal to the fraction of 4π sr covered by the ray in the direction (φ,θ). For isotropic sampling with N_d directions, $W_{\varphi,\theta}\equiv(1/N_d)\forall(\varphi,\theta)$. For the FAI method, it is given by

$$W_{\varphi,\theta}=\begin{cases}\dfrac{1-\cos\left(\dfrac{\omega}{2}\right)}{2} & \varphi=90^\circ \quad \text{or} \quad \varphi=-90^\circ\\ \dfrac{\omega}{720}\cdot\left[\sin\left(\varphi+\dfrac{\omega}{2}\right)-\sin\left(\varphi-\dfrac{\omega}{2}\right)\right] & \varphi=90^\circ-\omega,90^\circ-2\omega,\ldots,-90^\circ+\omega.\end{cases} \quad (5)$$

Although planning is done assuming $\Sigma=3$ mm, coverage probabilities are computed for a range of potential systematic errors ($\Sigma=0-7$ mm) to ensure identification of DMD sampling parameters that would provide accurate coverage estimates across a range of Σ values. The use of multiple Σ values also illustrates the types of studies that could be performed as part of plan evaluation. Specifically, it address the question: How do target and OAR coverage values vary if the systematic SD Σ differs from its assumed value?

II.E. Coverage uncertainty analysis

Sampling of the DMD over a finite number of directions and using a finite step size δ will introduce some error or uncertainty into coverage estimates. To quantify this uncertainty, we assume that sufficiently small angular and distance sampling parameters give negligible error. Specifically, we assumed that "baseline" coverage estimates with $\omega=1^\circ$ (or 64 442 isotropic directions) for CTV and $\omega=2^\circ$ (or 16 022 isotropic directions) for bladder and rectum give negligible error due to angular gradients in the dosimetric margin. (When quantifying estimation errors due to ω, baseline DM calculations use $\omega=1^\circ$ and $\delta=1$ mm for the CTV and $\omega=2^\circ$ and $\delta=1$ mm for the bladder and rectum.) Similarly, we assumed that baseline coverage estimates with $\delta=0.1$ mm give negligible error due to radial dose gradients. (When quantifying estimation errors due to finite δ, baseline DM calculations use $\omega=10^\circ$ and $\delta=0.1$ mm for the CTV, bladder, and rectum.) For other values of the sampling parameters ω and δ, coverage uncertainty is estimated by taking the difference with respect to baseline parameters. For example, the uncertainty due to $\omega=5^\circ$ for a certain Σ is determined by $\Delta Q(\Sigma)=Q(\Sigma)|_{\omega=5^\circ,\delta=1\ \text{mm}}-Q(\Sigma)|_{\omega=1^\circ,\delta=1\ \text{mm}}$.

In addition to the sampling parameters ω and δ, Q is a function of Σ. As Σ is varied, coverage will vary, as shown

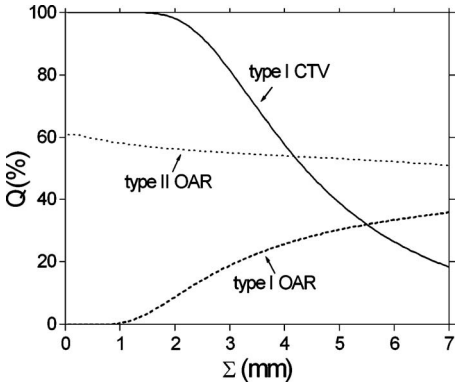

FIG. 1. Variation in the coverage probability Q with systematic standard deviation Σ for type I targets and type I/II OAR. Plots were obtained from patient 28 by calculating CTV $M_{\min,79.2}$ (type I), bladder $M_{25,75}$ (type II), and rectum $M_{25,70}$ (type I) and then computing coverage Q [Eqs. (2)–(4)] for Σ values ranging from 0 to 7 mm. The purpose of these plots is to illustrate the qualitative dependence of Q on Σ, not the specific coverage values obtained for patient 28.

in Fig. 1. Figure 1 was obtained from patient 28 by calculating CTV $M_{\min,79.2}$(type I), bladder $M_{25,75}$ (type II), and rectum $M_{25,70}$ (type I) and then computing the coverage Q [Eqs. (2)–(4)] for Σ values ranging from 0 to 7 mm. The purpose of Fig. 1 is to illustrate the qualitative dependence of Q on Σ for type I targets and type I/II OAR, not the specific coverage values obtained for patient 28. As Σ increases, type I target or type II OAR coverage decreases, while type I OAR coverage increases.

This work calculates $\Delta Q_{\min}$ and $\Delta Q_{\max}$ as the minimum and maximum values of $\Delta Q(\Sigma)$ taken over Σ values in the range [0, 7 mm],

$$\Delta Q_{\min} = \min_{\Sigma}[Q(\Sigma)_{\omega,\delta} - Q(\Sigma)_{\text{baseline}}], \tag{6}$$

$$\Delta Q_{\max} = \max_{\Sigma}[Q(\Sigma)_{\omega,\delta} - Q(\Sigma)_{\text{baseline}}]. \tag{7}$$

Here, $\Delta Q_{\min}$ and $\Delta Q_{\max}$ are sometimes simply referred to as ΔQ for short in the following paragraphs. The results presented below plot the range of 56 ΔQ obtained from the 28 prostate plans as a function of the sampling parameters ω and δ.

DM calculations were performed using a C++ interface to the Pinnacle treatment planning system that allowed the dose distribution to be moved relative to regions of interest (mathematically equivalent to moving structures relative to the dose distribution) and the resulting metric values D_v to be calculated. Additional numerical analysis, including coverage estimation, was performed outside of Pinnacle using MATLAB (MathWorks, Natick, MA).

III. RESULTS

III.A. Coverage values

Figure 2 illustrates the ranges of coverage values obtained for the prostate, bladder, and rectum across the 28 plans. It plots ranges for CTV $Q_{\min,79.2}(\Sigma)$, bladder $Q_{25,75}(\Sigma)$, and rectum $Q_{25,70}(\Sigma)$ for Σ=1, 3, and 5 mm. The results for Σ=1 and 5 mm illustrate what happens if Σ differs from the assumed value Σ=3 mm: Coverage varies substantially. For each box plot, the upper and the lower markers indicate the maximum and minimum values, the box extends from the 10th to 90th percentile, and the triangle is the median value. Coverage values in Fig. 2(a) were calculated using ω=1° and δ=1 mm. Consistent with Fig. 1, CTV coverage decreases as Σ increases. CTV coverage at the prescription dose (79.2 Gy) is ~100% for Σ=1 mm but ranges from approximately 20% to 100% for Σ=5 mm. In the case of the CTV, all plans were of type I (i.e., had CTV static $D_{\min}$ > 79.2 Gy). Figure 2(a) compares coverage values obtained using the FAI (left) versus the ISO (right) methods. The two methods give almost identical results, indicating that for ω =1° and δ=1 mm, the DMD does not depend on the sampling method.

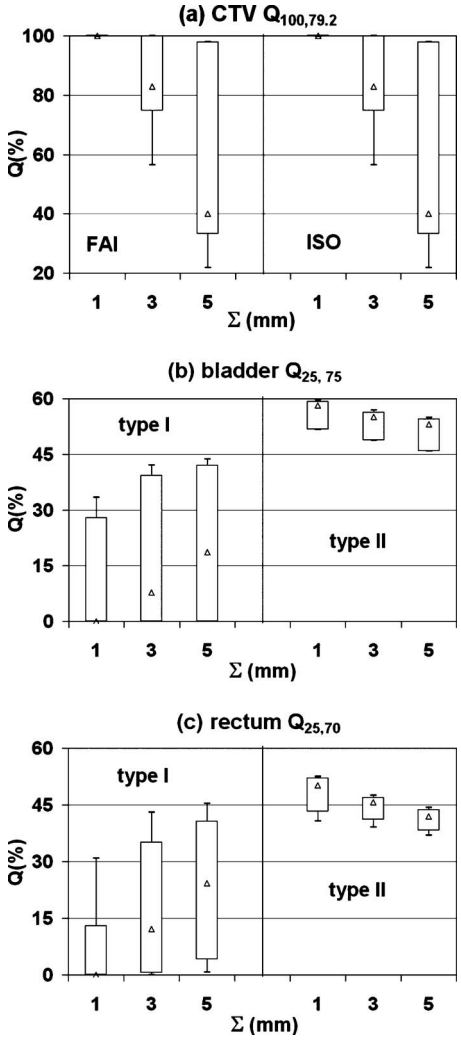

FIG. 2. Box plots of $Q(\Sigma)$: (a) CTV $Q_{\min,79.2}$ obtained using the FAI (left) and ISO (right) methods, (b) bladder $Q_{25,75}$ for type I (left) and type II (right) plans (FAI method), and (c) rectum $Q_{25,70}$ for type I (left) and type II (right) plans (FAI method). The upper and the lower markers are the maximum and minimum values taken over the 28 plans, the box extends from the 10th to 90th percentile, and the triangle is the median value.

Figure 2(b) plots ranges for bladder $Q_{25,75}$. Coverage values in Fig. 2(b) were calculated using $\omega=2°$ and $\delta=1$ mm. For the bladder, 23 plans were of type I (static $D_{25}<75$ Gy) and 5 of type II (static $D_{25}>75$ Gy). Type II plans were for anatomies in which the bladder was in close proximity to the prostate, ensuring that in the planned dose distribution more than 25% of its volume received >75 Gy. Figure 2(b) shows the range of coverage values for the type I (left) and type II (right) plans. Consistent with Fig. 1, as Σ increases, coverage of type I OARs increases, while coverage of type II OARs decreases. Bladder coverage values range from 0% to ~60% across the 28 plans.

Figure 2(c) plots ranges for rectum $Q_{25,70}$. Coverage values in Fig. 2(c) were calculated using $\omega=2°$ and $\delta=1$ mm. In the case of the rectum, 24 plans were of type I (static $D_{25}<70$ Gy) and 4 of type II (static $D_{25}>70$ Gy). Type II plans were for anatomies in which the rectum was in close proximity to the prostate, resulting in a planned dose distribution with more than 25% of its volume receiving >70 Gy. Note that type II rectum plans were mostly distinct from type II bladder plans. Only one of the 28 plans was of type II for both bladder and rectum. Figure 2(c) shows that rectum coverage values range from 0% to ~50% across the 28 plans.

III.B. Coverage uncertainties

III.B.1. CTV

To obtain the coverage uncertainties due to finite sampling of the DMD for 28 patients involved in this study, the minimum and maximum coverage differences $\Delta Q_{\min}$ and $\Delta Q_{\max}$ were calculated as in Eqs. (6) and (7). This produced 28 $\Delta Q_{\min}$ and 28 $\Delta Q_{\max}$ values (and thus a total of 56 ΔQ values) for each group of sampling parameters. Figure 3 presents the range of ΔQ values associated with CTV $Q_{\min,79.2}$. The lower marker in Fig. 3 represents the minimum of the 28 $\Delta Q_{\min}$ values and the upper marker represents the maximum of the 28 $\Delta Q_{\max}$ values. The boxes indicate the 10th and 90th percentiles of the 56 $\Delta Q_{\min}$ and $\Delta Q_{\max}$ values, and the triangle represents the median of the 56 $\Delta Q_{\min}$ and $\Delta Q_{\max}$ values.

Figure 3(a) plots the range of ΔQ values as a function of the angular increment ω. In this case, ΔQ values were calculated relative to baseline coverage values obtained with $\omega=1°$ and $\delta=1$ mm. Figure 3(b) plots the range of ΔQ values as a function of the radial step size δ. In this case, ΔQ values were calculated relative to the baseline coverage values obtained with $\omega=10°$ and $\delta=0.1$ mm. Taken together, Figs. 3(a) and 3(b) show that the uncertainty in the estimated CTV coverage ($|\Delta Q|$) increases from <0.1% for fine sampling parameters ($\omega=2°$ and $\delta=0.2$ mm) up to about 0.8% for coarse sampling parameters ($\omega=20°$ and $\delta=1$ mm).

Also, some differences between the FAI and ISO methods are shown in Figs. 3(a) and 3(b). The ISO method produces smaller ΔQ values. This suggests that more uniform sampling of the DMD over the unit sphere—as in the ISO method—will produce better estimates of the DMD, leading to smaller ΔQ values.

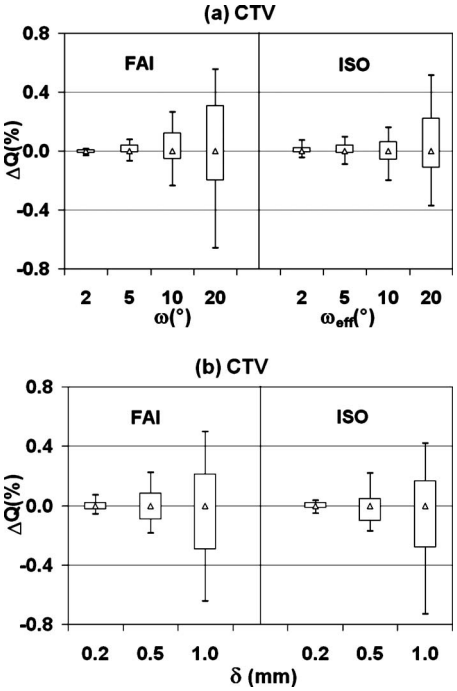

FIG. 3. ΔQ box plots for CTV $Q_{100,79.2}$ obtained using the FAI (left) and ISO (right) calculation methods: (a) ΔQ range versus $\omega=2,5,10,20°$. All Q values were calculated using $\delta=1$ mm. Baseline Q values were obtained using $\omega=1°$. (b) ΔQ range versus $\delta=0.2,0.5,1$ mm. All Q values were calculated using $\omega=10°$. Baseline Q values were obtained using $\delta=0.1$ mm. The upper and the lower markers are the maximum and minimum values taken over the 28 plans, the box extends from the 10th to 90th percentile, and the triangle is the median value.

III.B.2. OAR

Figures 4 and 5 illustrate ΔQ ranges for the bladder and rectum. The minimum, maximum, 10th and 90th percentiles, and median of the 56 ΔQ values are given. The bladder $Q_{25,75}$ corresponds with the bladder optimization criterion $D_{25}<75$ Gy. For this criterion, 23 patient bladders were a type I structure, while the remaining 5 were of type II. The $Q_{25,80}$ plot shows how coverage varies with the isodose surface for which coverage is evaluated. Variability is less for the 80 Gy isodose surface, for which all 28 patient bladders are of type I. Figure 5 presents similar results for the rectum $Q_{25,70}$ and $Q_{25,75}$. There are 24 type I rectums for $Q_{25,70}$ and 28 for $Q_{25,75}$.

IV. DISCUSSION

All IMRT plans used in this study were generated with uniform 10 mm CTV-to-PTV margins. They were evaluated using interfraction setup uncertainties having random SD $\sigma=3$ mm (modeled via fluence convolution) and a range of systematic SDs Σ. According to the margin formula of van Herk *et al.*,[9] for $\sigma=3$ mm and $\Sigma=3$ mm, a 10 mm margin is supposed to ensure that CTV $D_{\min}$ exceeds the prescription dose (strictly, the planned PTV $D_{\min}$) for 90% of motion. That is, the 10 mm margin is supposed to ensure that $Q_{\min,79.2}$ exceeds 90% for all plans.

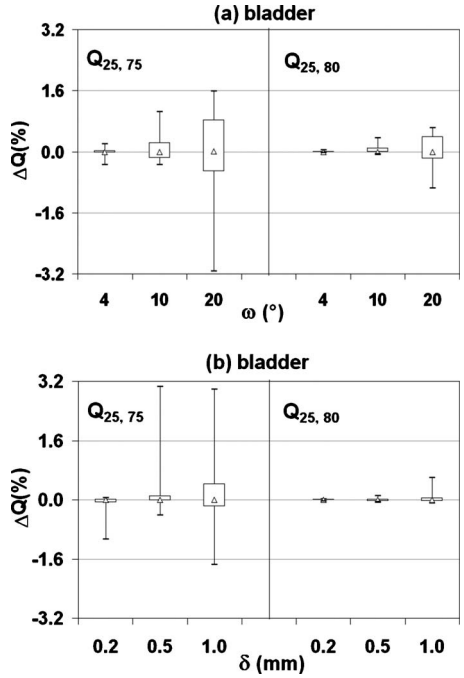

FIG. 4. ΔQ box plots for bladder $Q_{25,75}$ and $Q_{25,80}$ obtained using the FAI calculation method: (a) ΔQ range versus $\omega=4,10,20°$. All Q values were calculated using $\delta=1$ mm. Baseline Q values were obtained using $\omega=2°$. (b) ΔQ range versus $\delta=0.2,0.5,1$ mm. All Q values were calculated using $\omega=10°$. Baseline Q values were obtained using $\delta=0.1$ mm. The upper and the lower markers are the maximum and minimum values taken over the 28 plans, the box extends from the 10th to 90th percentile, and the triangle is the median value.

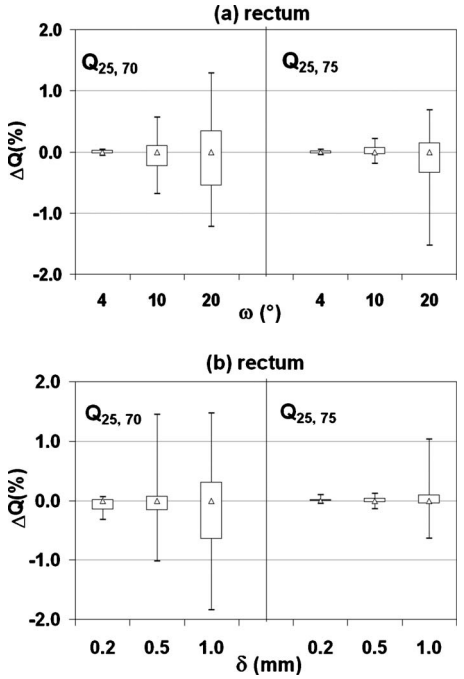

FIG. 5. ΔQ box plot for rectum $Q_{25,70}$ and $Q_{25,75}$ obtained using the FAI calculation method: (a) ΔQ range versus $\omega=4,10,20°$. All Q values were calculated using $\delta=1$ mm. Baseline Q values were obtained using $\omega=2°$. (b) ΔQ range versus $\delta=0.2,0.5,1$ mm. All Q values were calculated using $\omega=10°$. Baseline Q values were obtained using $\delta=0.1$ mm. The upper and the lower markers are the maximum and minimum values taken over the 28 plans, the box extends from the 10th to 90th percentile, and the triangle is the median value.

In clinical plans, this may not occur for two reasons: (a) Competition between target and OAR optimization criteria causes the target dose to fall short of the prescription dose and (b) the treated volume, which is a function of the beam arrangement and individual anatomy, provides a different level of coverage than the PTV was intended to achieve. Both effects contribute to variability of the target coverage. In Fig. 2(a), for $\Sigma=3$ mm, $Q_{\min,79.2}$ varies from ~56% to 100%, with a median value of approximately 83%. These results emphasize the degree to which target coverage depends on the individual patient anatomy and the quality of the IMRT plan, in addition to the PTV margin. In clinical planning, it is common practice to use a standard CTV-to-PTV expansion and simply assume that the resulting PTV ensures an acceptable level of robustness to setup errors and other geometric uncertainties.

Figures 2(b) and 2(c) emphasize the wide range of coverage values that can be experienced by OAR. For $\Sigma=3$ mm, Fig. 2(c) shows that rectum $Q_{25,70}$ varies from 0% to ~50%. That is, for some patients, rectum D_{25} could exceed the 70 Gy constraint (tolerance dose) for as many as 50% of the treatment courses. Most OARs are of type I versus type II. Type I (II) OARs are those that satisfy (violate) the dose constraint (e.g., rectum $D_{25}\leq 70$ Gy) in the static plan. Based on the static plan, it appears that type I OARs are adequately protected—they satisfy their optimization constraints. However, Figs. 2(b) and 2(c) show that even a type I OAR can have a significant probability of exceeding optimization constraints, once probable patient offsets are considered. In general, one would like to ensure that OAR dose constraints (e.g., rectum $D_{25}\leq 70$ Gy) are violated for only a small percent, e.g., 5%, of treatment courses.

Figure 2 emphasizes the desirability of evaluating the actual, as opposed to the nominal, level of the target coverage provided by a PTV expansion. For some patients, if a standard PTV does not provide acceptable coverage, it may be desirable to further expand the PTV until treatment goals are achieved. For others, PTV reduction might continue to yield adequate CTV coverage while affording lower coverage probabilities for OARs. Figure 2 also emphasizes the desirability of postplanning evaluation of OAR coverage. The clinical significance of Fig. 2 is that when geometric uncertainties are considered, patient-specific characteristics can have a large effect on the target and OAR coverages even when the plans are created with the same margin and analyzed using the same assumptions regarding setup errors. Coverage analysis would identify those plans whose CTV coverage falls below some threshold (e.g., 90%, as in the margin formula of van Herk *et al.*), or whose OAR coverage exceeds some threshold (e.g., 10%), allowing replanning to be performed.

In Figs. 2(b) and 2(c), there are differences between the type I and type II results. Type I OARs are from patients with favorable anatomies, in which the rectum and bladder were adequately separated from the prostate. In these patients, OAR coverage values tend to be low. A small number of

patients had unfavorable anatomies, in which the (type II) rectum and bladder were close to the prostate. In these patients, OAR coverage values tend to be high. The gap between type I and type II coverage ranges in Figs. 2(b) and 2(c) is hypothesized to be due to the relatively small number of type II "outlier" patients in the sample. A larger sample might include patients with OAR that are intermediate between the type I and type II samples in Fig. 2. One can speculate that in this case, the coverage ranges for types I and II would join (at Q around 50%) and become contiguous. These hypotheses would need to be proved or rejected by sampling a larger patient population, which is outside the scope of the present paper.

The DMD is a useful tool for evaluating target and OAR coverage. When combined with a geometric uncertainty model, such as the simple normal uncertainty model given in Eq. (2), or more complex models derived from imaging studies, it can be used to quantify coverage as in Eqs. (3) and (4). Practical implementation requires one to sample the DMD using a finite number of directions and finite spatial step size. As a practical issue, it therefore becomes necessary to determine how finely one must sample the DMD in order to derive sufficiently accurate coverage estimates. Figure 3 shows that for a compact target such as the prostate, accuracy of better than 1% ($|\Delta Q|<1\%$) can be achieved with quite coarse sampling parameters, e.g., $\omega=20°$, $\delta=1$ mm, while $\omega=10°$, $\delta=0.5$ mm will ensure $|\Delta Q|<0.5\%$. The use of isotropic sampling gives lower estimation errors due to more uniform sampling of the DMD.

Comparison of Fig. 3 with Figs. 4 and 5 shows that for the same sampling parameters ω and δ, the uncertainty $|\Delta Q|$ is higher for an OAR than it is for the CTV. For example, ΔQ values in Figs. 4 and 5 extend to $\pm 3\%$, whereas in Fig. 3 $|\Delta Q|<1\%$. Figure 6 indicates that this is most likely due to the fact that OAR coverage will be significant only for a subset of directions—specifically, those that move the OAR toward the high dose region. OAR coverage calculations therefore rely more heavily on dosimetric margins for a few directions and are consequently more sensitive to DM estimation errors for those directions.

For CTVs, one will typically require coverage at the prescription dose to be high (e.g., CTV $Q_{\min,79.2}>95\%$). Certainly one would hope that, in most cases, target coverage falls in the range of 90%–100%. In that case, it would be reasonable to require 1% accuracy for coverage estimates: $|\Delta Q|<1\%$. Accuracy of 1% allows one to distinguish with confidence plans that have, e.g., 93% or 97% target coverage from those with 95% coverage. Accuracy of 5% would clearly be too poor, whereas accuracy of much less than 1% would be unnecessary. We therefore propose $|\Delta Q|<1\%$ as a reasonable criterion for target coverage estimates. For an OAR, the goal will typically be to achieve low coverage values ($Q<5\%$) at tolerance doses. However, this goal may be sacrificed if it is in conflict with target coverage. One may end up being satisfied, for example, with rectum $Q_{25,70}<20\%$. In this case, it may not be necessary to estimate OAR coverage values to the same degree of accuracy as CTV coverage values. It may be acceptable, for example, if $|\Delta Q|<3\%$. This can be achieved with sampling parameters: $\omega=20°$, $\delta=1$ mm. More accurate OAR coverage estimates can be achieved by using finer sampling parameters or by

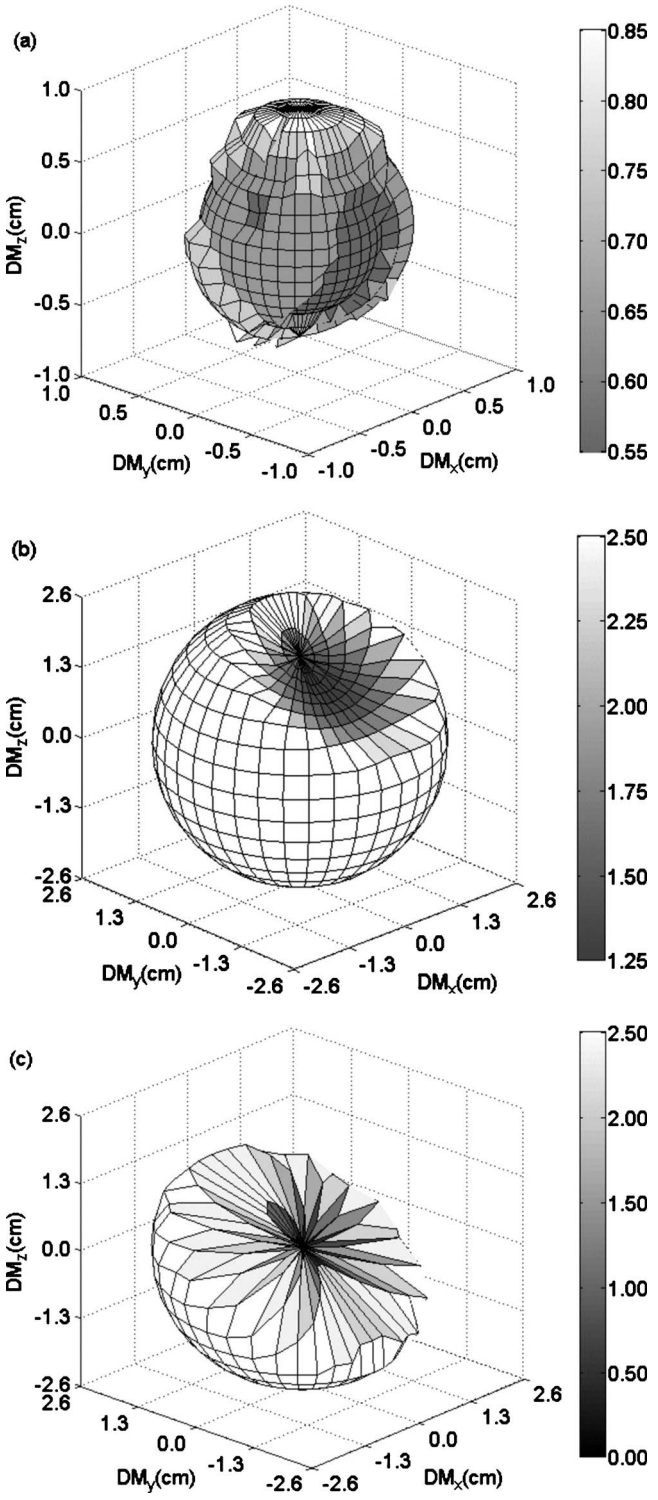

FIG. 6. Example DMDs obtained using the FAI method, plotted on the 3D axes. (a) Patient 4 CTV $D_{\min,79.2}$ (type I), (b) patient 27 bladder $D_{25,75}$ (type I), and (c) patient 21 bladder $D_{25,75}$ (type II). DM_x is the RL component of the dosimetric margin (i.e., the dosimetric margin projected onto the patients RL axis). DM_y and DM_z are the PA and SI components.

using an adaptive strategy that samples the DMD more finely in the directions that move the OAR toward the high dose region. The latter remains an area for further research. We note that our suggested accuracy criteria—$|\Delta Q|<1\%$ for targets and $|\Delta Q|<3\%$ for OAR—are a reasonable but subjective choice. The purpose of this paper is to quantify the dependence of $|\Delta Q|$ on sampling parameters. Using the results of this paper, users can set DMD sampling parameters to appropriate values to give more accurate coverage estimates, if desired.

This work assumed a simple rigid body uncertainty model, in which interfraction setup uncertainties are normally distributed. Intrafraction motion was not modeled. The assumed baseline parameter values $\Sigma=\sigma=3$ mm are broadly in line with the values reported in literature for setup to external markers.[21] For image-guided setup, interfraction setup errors will be smaller, but in order to properly account for geometric uncertainties, one should also account for intrafraction motion[22–24] and delineation errors.[25] The present work did not go to this level of detail. One could also contemplate adaptive replanning throughout the treatment course. This approach complements image-guided setup and has the potential to further reduce geometric uncertainties by generating plans that are tailored to the patient's anatomy on each treatment day. However, even with adaptive planning, geometric uncertainties are likely to persist due to factors such as image-registration errors and intrafraction motion.

It is important to note that the results in Fig. 2 are strongly dependent on the assumed uncertainty model and on the margin that was employed. Different setup and adaptive planning strategies will produce different uncertainty models and therefore will produce quantitatively different results from Fig. 2. Nevertheless, the qualitative features in Fig. 2 are expected to occur for other uncertainty models: It is expected that target and OAR coverage will vary significantly between patients even when the same motion model and CTV-to-PTV margin are employed. Thus, our conclusions regarding the desirability of performing postplan coverage evaluation are likely to remain valid, regardless of the uncertainty model specifics.

Figures 3–5 quantify the relative errors that can occur in coverage estimates based on the sampling parameters ω and δ. These results are chiefly a function of the DMD algorithm, not the uncertainty/motion model. In the future, it will be desirable to perform postplan coverage evaluation using validated deformable motion models for prostate, bladder, and rectum. These types of deformable models are being actively researched and are likely to emerge from principle components analysis once sufficient data and experience have been accumulated. Our conclusion regarding adequate values for ω and δ are expected to remain valid for these models.

Postplan coverage evaluation will add some time to the overall therapy process. The nonoptimized DMD algorithm used for this work took on the order of 30 min/patient for parameter values $\omega=10°$, $\delta=0.5$ mm on a 1.6 GHz SunFire V445. Run times on modern hardware and with optimized code would likely be substantially less. The algorithm was completely scripted within the Pinnacle treatment planning system and could be executed with a single (script) button push. Clinically, we envision that coverage evaluation could be performed as a standard part of postplanning QA. Coverage that is judged to be unacceptable could require replanning. This decision would need to be made on-balance with respect to other plan evaluation parameters such as DVHs, isodoses, and the physician's knowledge/experience. As in other cases where a plan fails a QA step, replanning would need to be performed with alternative planning parameters—most likely, a change in margin, either for an entire structure, or in a specific location. However, replanning might also entail a change in treatment angles or other parameters that affect the shape of the periphery of the dose distribution. As experienced is gained, we speculate that efficient ways to adjust plans will likely be discovered by the dosimetrists/treatment planners. This provides a model for how postplan coverage evaluation could be performed as part of the treatment workflow without imposing an undue burden.

V. CONCLUSIONS

Postplanning evaluation shows that realized target coverage can fall short of the nominal value implied by standard CTV-to-PTV margin expansions. The evaluation of OAR coverage shows a wide range of realized coverage values. In most cases, OARs are of type I (i.e., satisfy their dose constraints in the static plan) but can still have a significant probability of exceeding the dose constraints when patient offsets are accounted for. These results emphasize the desirability of performing postplanning coverage evaluation. Target and OAR coverage vary with individual patient anatomy and with plan quality. Postplan evaluation enables the effect of geometric uncertainties to be accurately assessed prior to treatment.

The accuracy of coverage estimates depends on angular and radial DMD sampling parameters ω and δ. The results suggest that coarse sampling parameters ($\omega=20°$, $\delta=1$ mm) will achieve accuracy $|\Delta Q|<1\%$ in target coverage estimates and accuracy of $|\Delta Q|<\sim 3\%$ in OAR coverage estimates. Finer sampling parameters ($\omega=10°$, $\delta=0.5$ mm) will achieve accuracy $|\Delta Q|<0.5\%$ in target coverage estimates and accuracy of $|\Delta Q|<\sim 1\%$ in OAR coverage estimates, which is judged to be sufficiently accurate for planning purposes. Even finer DMD sampling ($\omega<10°$, $\delta<0.5$ mm), which would require significant processing time by the treatment planning system, is unlikely to be required.

ACKNOWLEDGMENTS

This work was supported, in part, by NIH Grant No. P01CA116602 and by a research contract with Philips Medical Systems.

[a]Electronic mail: xuh2@vcu.edu

[1]J. M. Wilkinson, "Geometric uncertainties in radiotherapy," Br. J. Radiol. **77**(914), 86–87 (2004).
[2]ICRU, "Prescribing, recording and reporting photon beam therapy," Report No. 50 (ICRU, Bethesda, MD, 1994).
[3]ICRU, "Prescribing, recording and reporting photon beam therapy (Supplement to ICRU Report 50)," Report No. 62 (ICRU, Bethesda, MD, 2000).

[4]L. P. Muren, R. Smaaland, and O. Dahl, "Organ motion, set-up variation and treatment margins in radical radiotherapy of urinary bladder cancer," Radiother. Oncol. **69**(3), 291–304 (2003).
[5]S. W. Lee *et al.*, "Patterns of failure following high-dose 3-D conformal radiotherapy for high-grade astrocytomas: A quantitative dosimetric study," Int. J. Radiat. Oncol., Biol., Phys. **43**(1), 79–88 (1999).
[6]M. R. Button, C. A. Morgan, E. S. Croydon, S. A. Roberts, and T. D. Crosby, "Study to determine adequate margins in radiotherapy planning for esophageal carcinoma by detailing patterns of recurrence after definitive chemoradiotherapy," Int. J. Radiat. Oncol., Biol., Phys. **73**(3), 818–823 (2009).
[7]L. Ekberg, O. Holmberg, L. Wittgren, G. Bjelkengren, and T. Landberg, "What margins should be added to the clinical target volume in radiotherapy treatment planning for lung cancer?," Radiother. Oncol. **48**(1), 71–77 (1998).
[8]J. C. Stroom, H. C. de Boer, H. Huizenga, and A. G. Visser, "Inclusion of geometrical uncertainties in radiotherapy treatment planning by means of coverage probability," Int. J. Radiat. Oncol., Biol., Phys. **43**(4), 905–919 (1999).
[9]M. 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**(4), 1121–1135 (2000).
[10]M. van Herk, "Errors and margins in radiotherapy," Semin. Radiat. Oncol. **14**(1), 52–64 (2004).
[11]J. 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**(1), 202–214 (2007).
[12]J. J. Gordon and J. V. Siebers, "Evaluation of dosimetric margins in prostate IMRT treatment plans," Med. Phys. **35**(2), 569–575 (2008).
[13]B. C. Cho, M. van Herk, B. J. Mijnheer, and H. Bartelink, "The effect of set-up uncertainties, contour changes, and tissue inhomogeneities on target dose-volume histograms," Med. Phys. **29**(10), 2305–2318 (2002).
[14]J. J. Gordon and J. V. Siebers, "Coverage-based treatment planning: Optimizing the IMRT PTV to meet a CTV coverage criterion," Med. Phys. **36**(3), 961–973 (2009).
[15]J. 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**(2), 550–563 (2010).
[16]Radiation Therapy Oncology Group, RTOG Protocol 0126: A phase III randomized study of high dose 3D-CRT/IMRT versus standard dose 3D-CRT/IMRT in patients treated for localized prostate cancer.
[17]W. A. Beckham, P. J. Keall, and J. V. Siebers, "A fluence-convolution method to calculate radiation therapy dose distributions that incorporate random set-up error," Phys. Med. Biol. **47**(19), 3465–73 (2002).
[18]T. Craig, J. Battista, and J. Van Dyk, "Limitations of a convolution method for modeling geometric uncertainties in radiation therapy. I. The effect of shift invariance," Med. Phys. **30**(8), 2001–2011 (2003).
[19]M. van Herk, M. Witte, J. van der Geer, C. Schneider, and J. V. Lebesque, "Biologic and physical fractionation effects of random geometric errors," Int. J. Radiat. Oncol., Biol., Phys. **57**(5), 1460–1471 (2003).
[20]E. Saff and A. Kuijlaars, "Distributing many points on a sphere," Math. Intell. **19**(1), 5–11 (1997).
[21]B. Schaly, G. S. Bauman, W. Song, J. J. Battista, and J. V. Dyk, "Dosimetric impact of image-guided 3D conformal radiation therapy of prostate cancer," Phys. Med. Biol. **50**(13), 3083–3101 (2005).
[22]K. Langen *et al.*, "Observations on real-time prostate gland motion using electromagnetic tracking," Int. J. Radiat. Oncol., Biol., Phys. **71**(4), 1084–1090 (2008).
[23]H. Li *et al.*, "Dosimetric consequences of intrafraction prostate motion," Int. J. Radiat. Oncol., Biol., Phys. **71**(3), 801–812 (2008).
[24]K. T. Malinowski *et al.*, "Efficient use of continuous, real-time prostate localization," Phys. Med. Biol. **53**(18), 4959–4970 (2008).
[25]C. Njeh, "Tumor delineation: The weakest link in the search for accuracy in radiotherapy," Journal of Medical Physics **33**(4), 136 (2008).