Int. J. Radiation Oncology Biol. Phys., Vol. 52, No. 5, pp. 1407–1422, 2002

0360-3016/02/$–see front matter

PII S0360-3016(01)02805-X

PHYSICS CONTRIBUTION

INCLUSION OF GEOMETRIC UNCERTAINTIES IN TREATMENT PLAN EVALUATION

MARCEL VAN HERK, PH.D., PETER REMEIJER, PH.D., AND JOOS V. LEBESQUE, M.D., PH.D.

Radiotherapy Department, The Netherlands Cancer Institute/Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands

Purpose: To correctly evaluate realistic treatment plans in terms of absorbed dose to the clinical target volume (CTV), equivalent uniform dose (EUD), and tumor control probability (TCP) in the presence of execution (random) and preparation (systematic) geometric errors.

Materials and Methods: The dose matrix is blurred with all execution errors to estimate the total dose distribution of all fractions. To include preparation errors, the CTV is randomly displaced (and optionally rotated) many times with respect to its planned position while computing the dose, EUD, and TCP for the CTV using the blurred dose matrix. Probability distributions of these parameters are computed by combining the results with the probability of each particular preparation error. We verified the method by comparing it with an analytic solution. Next, idealized and realistic prostate plans were tested with varying margins and varying execution and preparation error levels.

Results: Probability levels for the minimum dose, computed with the new method, are within 1% of the analytic solution. The impact of rotations depends strongly on the CTV shape. A margin of 10 mm between the CTV and planning target volume is adequate for three-field prostate treatments given the accuracy level in our department; i.e., the TCP in a population of patients, TCP_{pop}, is reduced by less than 1% due to geometric errors. When reducing the margin to 6 mm, the dose must be increased from 80 to 87 Gy to maintain the same TCP_{pop}. Only in regions with a high-dose gradient does such a margin reduction lead to a decrease in normal tissue dose for the same TCP_{pop}. Based on a rough correspondence of 84% minimum dose with 98% EUD, a margin recipe was defined. To give 90% of patients at least 98% EUD, the planning target volume margin must be approximately $2.5\ \Sigma + 0.7\ \sigma - 3$ mm, where Σ and σ are the combined standard deviations of the preparation and execution errors. This recipe corresponds accurately with 1% TCP_{pop} loss for prostate plans with clinically reasonable values of Σ and σ.

Conclusion: The new method computes in a few minutes the influence of geometric errors on the statistics of target dose and TCP_{pop} in clinical treatment plans. Too small margins lead to a significant loss of TCP_{pop} that is difficult to compensate for by dose escalation. © **2002 Elsevier Science Inc.**

Radiotherapy, Geometric uncertainties, Margins, Tumor control probability, Equivalent uniform dose, Clinical target volume.

INTRODUCTION

In conformal radiotherapy, margins are often reduced to increase normal-tissue sparing. It is, however, unclear whether the potential clinical benefit outweighs the risk of missing the target. At present, most treatment planning systems evaluate and optimize the planned dose distribution using static volumes. This approach grossly simplifies geometric errors and gives rise to a classical dilemma (Fig. 1a). When a planned dose distribution is evaluated on the clinical target volume (CTV), i.e., the volume that does not include a margin for geometric uncertainties, all geometric errors are explicitly ignored. This leads to an estimated tumor control probability (TCP) that will, in general, be higher than can be expected when geometric errors are included. However, it is also not correct to use the planning target volume (PTV), which does include a margin for geometric uncertainties, to evaluate the dose or compute the TCP. The PTV is generally chosen in such a way that it will include the CTV with a certain probability. However, in reality the CTV will have an unknown position within the PTV, and the CTV also has a smaller volume than the PTV. Because of the volume difference, the TCP derived from the PTV will underestimate the real TCP. In addition, the effect of dose inhomogeneities in the PTV will be overestimated. We will describe another, more correct, way of evaluating a treatment plan; this method does not use the PTV but is based on detailed knowledge of the distribution of geometric errors in the patient population (Fig. 1b). Of course, the

Reprint requests to: Marcel van Herk, Radiotherapy Department, The Netherlands Cancer Institute/Antoni van Leeuwenhoek Hospital, Plesmanlaan 121, 1066 CX, Amsterdam, The Netherlands. Tel: +31-20-5122212; Fax: +31-20-6691101; E-mail: portal@nki.nl

Part of this material was presented at the 2000 ASTRO meeting in Boston, MA.

Received Jan 18, 2001, and in revised form Jun 25, 2001. Accepted for publication Nov 21, 2001.

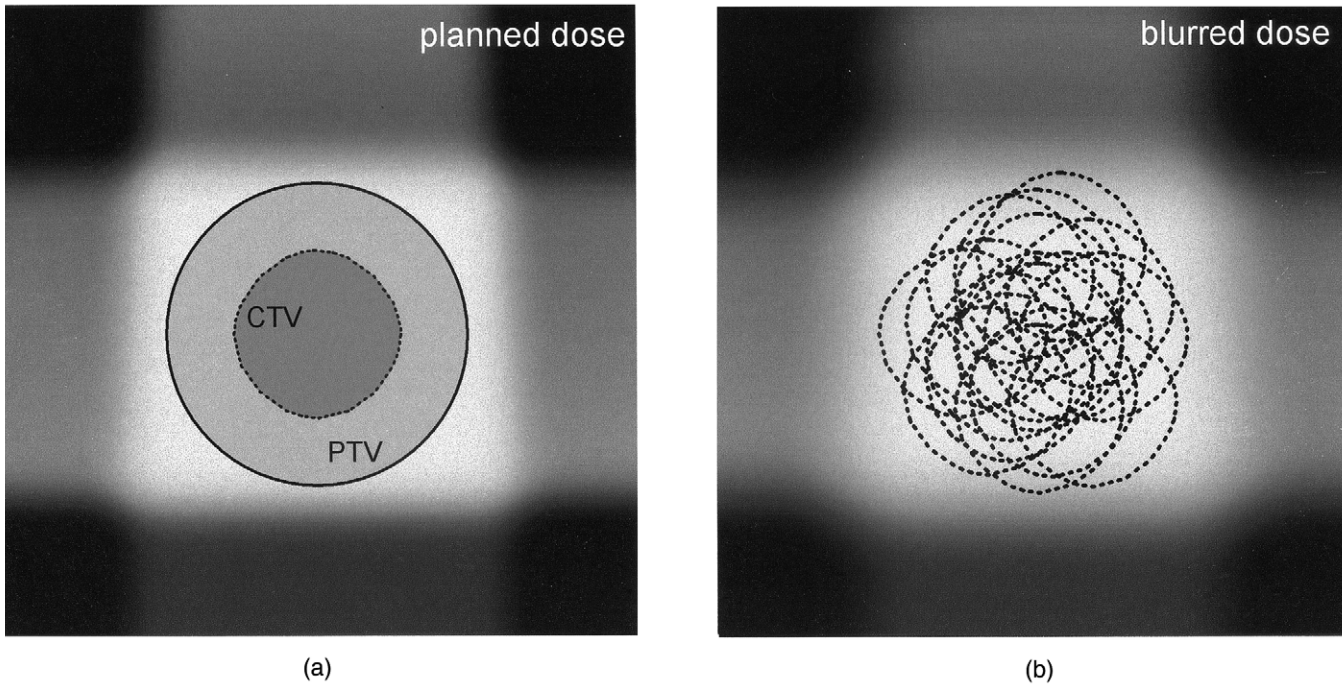

Fig. 1. (a) When considering the CTV as a static structure, evaluation of its dose distribution is too optimistic, because it is known that in reality, the CTV will be in a different place for each patient and for every fraction. Dose evaluation over the PTV is in general too pessimistic, because the PTV is much larger than the target. (b) A more correct evaluation of a treatment plan is obtained by blurring the dose distribution for the day-to-day variation and testing a large number of possible preparation (systematic) shifts of the CTV.

same knowledge of geometric errors should have been the basis for determining the margin used for expanding the CTV to PTV.

In this paper, it is assumed that information about the error distributions of geometric uncertainties, such as organ (target) movement, setup error, and delineation uncertainty, is available. It should be stressed that, in general, these uncertainties differ from institution to institution and that determination of these uncertainties is an essential first step in implementing any form of conformal radiotherapy.

We need to differentiate between two types of errors (e.g., Refs. 1–3). First, one must consider those errors that arise in the preparation stage of therapy, in general called systematic errors, but which we will refer to as "preparation errors." Second, one must consider those errors that arise during treatment execution that are often called random errors, but which we will refer to as "treatment execution errors." Preparation (systematic) errors include setup error and organ motion on the CT scanner, delineation errors, and equipment calibration errors, whereas execution (random) errors include target movement and day-to-day variation in the patient setup or equipment. Systematic errors that occur during treatment execution (e.g., because of laser misalignment) are called preparation errors in our nomenclature, because these types of errors are caused by the preparation of the equipment.

Correction procedures based on on-line electronic portal imaging (e.g., Ref. 4), implanted markers (e.g., Ref. 5), or other forms of image guidance may be used to reduce the level of execution and preparation errors. Off-line correction procedures, e.g., those based on electronic portal imaging (6) or repeat CT (7), are less time-consuming, but these only reduce preparation errors. Even though it is very important to implement correction procedures, there will always be residual preparation and execution errors. For instance, electronic portal imaging will not correct organ motion and target volume delineation inaccuracies; moreover, each correction procedure will have its own inaccuracies, for instance because of the limited accuracy with which measurements and corrections can be performed. For this reason, a margin around the CTV will always be required, and treatment plan evaluation should include the residual uncertainties.

This paper does not deal with the gross tumor volume (GTV) to CTV expansion and with nonuniform tumor cell density. These topics will be studied in the future.

In this paper, a method will be described that allows a correct evaluation of a treatment plan in the presence of geometric errors with a known probability distribution. As input, a volume of interest (e.g., the CTV) and a planned dose distribution are used, which will be combined to compute dose–volume histograms (DVHs) (8). The new method explicitly takes treatment preparation errors and treatment execution errors into account according to a model that we previously published (3). The analysis is performed on realistic treatment plans (e.g., Ref. 9). The novelty of this work is that biologic parameters such as the equivalent uniform dose (EUD) (10, 11) and TCP (12) are included. In

this paper we will consider only the target volume. However, the described methods are equally applicable for studying the dose to organs at risk and for performing normal tissue complication probability computations.

In our study, a distribution of both preparation and execution errors is taken into account. A limitation of many previous studies was that only execution (random) errors were evaluated (13–19), whereas others tested only a few preparation (systematic) errors (20–22). Because preparation errors have a much larger impact on the target dose than execution errors (See, e.g., Refs. 3 and 22), the effect of excluding preparation errors in TCP evaluation is that very small margins might seem adequate. In reality, preparation errors play the most important role, so such small margins will lead to frequent geometric misses.

The aim of this study is to develop a method for correct statistical evaluation of realistic treatment plans in terms of absorbed dose, EUD, and TCP in the presence of execution (random) and preparation (systematic) errors. Its purposes are to better understand the effect of geometric errors and to help design treatment plans that are robust under "normal" geometric variation. In addition, we will use the algorithm to test the effect of margin choice on TCP and to define a novel margin recipe for the CTV to PTV expansion (based on the probability of delivering a given EUD).

METHODS AND MATERIALS

Theory

In this paper, the margin is defined as the distance between CTV and PTV. During the planning process, the beams will be shaped such that the planned minimum dose to the PTV will be 95% of the dose to the reference point (i.e., the isocenter). In this paper, uniform margins in three dimensions (3D) will be used in all tests, for simplicity. The proposed method of analysis, however, uses only the CTV and the planned dose distribution and is therefore valid for any type of margin or dose distribution.

When considering the target as static, execution (random) errors cause a blurring of the dose distribution, and they can simply be modeled in this way (e.g., Ref. 13). The execution error distribution gives the shape of the kernel with which the planned distribution must be convolved to obtain a blurred dose distribution. This blurred distribution is an estimate of the cumulative dose distribution delivered to the CTV, i.e., the total dose delivered to the CTV over all treatment fractions when no systematic errors are present. The actual cumulative CTV dose will deviate somewhat from this estimate, in particular when a limited number of fractions is given or when biologic fractionation effects are important. This paper deals with standard fractionated radiotherapy. We therefore tested the impact of fractionation effects for a typical number of fractions, 36, on the position of isodose lines in the normalized total dose (23). The errors from biologic fractionation effects seem to be smaller than 1 mm, even when using very small α/β ratios in the linear-quadratic model. These tests will be the topic of a future paper. In this work, fractionation effects are further ignored.

A preparation error causes a shift of the cumulative dose distribution relative to the volume of interest. However, the preparation error is generally unknown and different for each patient. The analysis therefore has to determine the effect of every possible preparation error within a reasonable range. By combining these effects with the probability distribution of the preparation errors, a population-based analysis of dose and TCP will be achieved. Note that shifting the dose distribution is only an approximation of the effect of a systematic error. In reality, a shift of the patient will also affect the shape of the dose distribution. In addition, beams might move relative to each other. For simplicity, these effects will be ignored in this paper.

We first explored a grid-based analysis. We blurred the planned dose distribution and performed a grid-based test of all preparation errors. However, a problem of this grid-based approach is that the required amount of computer time increases sharply with the required resolution. For instance, if we decrease the spacing of the three-dimensional computation grid by a factor of 2, both the number of dose points and the number of error points increase by a factor of 8, leading to a factor 64 increase in computation time. If the number of dimensions of the error space is increased further, for instance to include rotational errors, the computational complexity becomes prohibitively large.

To solve this problem, we chose to use a Monte Carlo technique, which is often used for numeric computation of integrals over a large number of dimensions. In our situation, this Monte Carlo technique is equivalent to simulating the treatment of a large number of patients. Figure 2 illustrates the procedure. By random sampling, a large number of preparation errors is simulated.

In this study, the errors are limited to translations and rotations. However, tumor or normal tissue shape variations can be taken into account by increasing the number of preparation error parameters (24, 25). Each of the preparation errors describes one possible treatment scenario or "case." The range of these errors is chosen such that the tails of the probability distribution are fully included. In practice, we use an ellipsoidal error space with a radius of four times the standard deviation (SD) of the preparation errors. Using the probability distribution of the preparation errors, each of the cases is assigned a probability. The total probability of all cases is normalized in such a way that the sum is equal to 1. Next, DVHs are computed for each of the cases. For this purpose, a large number of points is generated within the CTV using random sampling, and each point is assigned a volume such that the sum of the volumes of all points is equal to the original volume of the CTV. These points are transformed using the preparation error under consideration and are then used to sample the blurred dose distribution using trilinear interpolation. The resulting dose values are used to compute DVHs (26).

Our Monte Carlo simulation method is nonstandard in the sense that often samples are drawn from the given distribu-

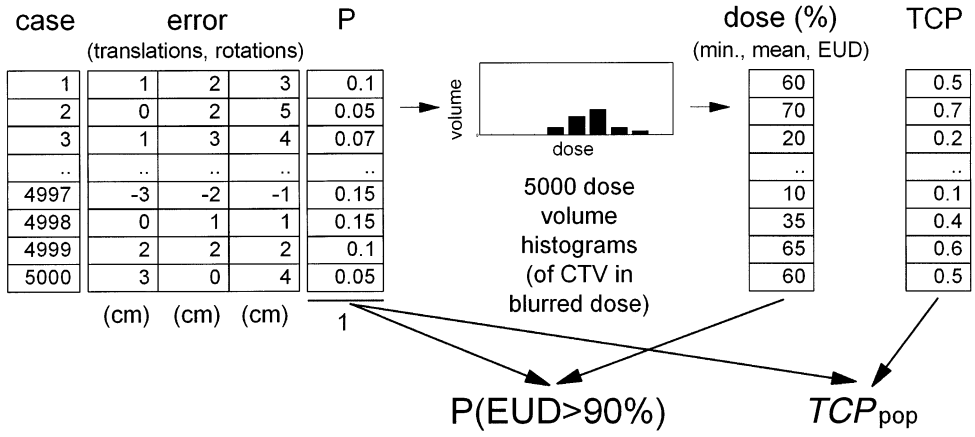

Fig. 2. Monte Carlo–based implementation of the evaluation of Fig. 1. First, preparation (systematic) errors (e.g., translations and rotations) are generated for a large number of cases by a random number generator. Using the probability distribution of the preparation errors, each of the cases is assigned a probability (P). Next, DVHs are computed for each of the preparation errors from a dose distribution that has been blurred to take execution (random) errors into account. From each of these DVHs, physical dose parameters and biologic parameters are computed. Finally, the computed parameters are combined with the probability of each case to compute the overall tumor control probability TCP_{pop} and the probability that a dose parameter (such as the EUD) exceeds a certain level.

tions, and no probability weighting is performed. We chose this particular method of Monte Carlo simulation, with probability weighting, because of the possibility of modifying the distribution of errors without recomputation. Instead, each case would just be assigned a different probability. Another advantage of the current method is that it samples a large number of low-probability, high-impact events, causing the estimated probability distribution plots to run more smoothly at lower probabilities.

The histograms were then analyzed to provide the following data for the CTV: minimum and mean dose, dose to 99% of the volume, EUD, and TCP. For the TCP computations, we used the Poisson model of Webb and Nahum (12). Probability distributions of these data values are computed by combining the results with the probability of each particular preparation error. The overall TCP, TCP_{pop}, is found as the sum of the products of TCP and probability for each simulated case. The TCP values are not only averaged over the simulated cases, but population averaging is also applied to take a spread in the tumor cell sensitivity into account. The computed data values were stored on computer disk for later analysis. Using these data, the relation between various physical CTV dose parameters and EUD was studied.

Experiments

In this study, we focus on the treatment of prostate patients. We first tested which parameters should be used in the numeric code to achieve a good precision. For this purpose, the grid size for the dose distribution, the number of points used to sample the CTV for DVH computation, the number of dose bins in the DVH analysis, and the number of tested preparation errors was varied. In succeeding experiments, the SD of execution and preparation errors and the margin between CTV and PTV (a uniform expansion in 3D) were varied to establish the relation between margin, error level, and biologic and dose parameters. Finally, rotational errors were switched on and off to test their impact. For all experiments, the biologic measures were computed for prescribed isocenter dose values ranging from 0 to 100 Gy in steps of 1 Gy. Most data will, however, be presented for a prescribed dose of 80 Gy.

Patient group–specific parameters

Table 1 lists the measured SDs of execution and preparation errors for the major error sources in radiotherapy of the prostate: target volume delineation (27), target (organ) movement (28), and setup errors (29, 30). The data include both translation and rotation errors. In our organ motion study, the statistics of random and systematic prostate motion were not analyzed separately, so that these values are assumed to be the same. The delineation variability was described in terms of a localized SD (31). For instance, the delineation variability is larger at the apex and the seminal vesicles. For this study we will assume, however, that the delineation error can be described by a simple translation. An average of the position-dependent errors is used to obtain error estimates for the left-right, anterior-posterior, and superior-inferior direction.

The rotational errors in the prostate position were recomputed using the apex as center of rotation (28, 30). The rotational errors were added quadratically, but because the rotation errors in setup were much smaller than the rotation of the prostate, the center of rotation was not moved: It was kept at the prostate apex.

For the TCP computations, we assumed that the CTV (prostate plus seminal vesicles) contained tumor with a uniform cell density of 10^7 cm^{-3}, and we used an *effective* value of α in the TCP model of 0.35 Gy^{-1} (32). The quadratic term (β) in the linear-quadratic model was set to

Table 1. Overview of prostate treatment uncertainties (SD of translations and rotations) as obtained from different studies from our group

	Treatment execution (random) errors (translation along 3 axes, rotation around 3 axes)			Treatment preparation (systematic) errors (translation along 3 axes, rotation around 3 axes)		
	LR	SI	AP	LR	SI	AP
Target volume delineation (27)				1.7 mm	2.8 mm*	2 mm
Organ motion (translation) (28)	0.9 mm	1.7 mm	2.7 mm	0.9 mm	1.7 mm	2.7 mm
Organ motion (rotation) (28)	4.0°	2.1°	1.3°	4.0°	2.1°	1.3°
Setup error (translation) (29)	2.0 mm	1.8 mm	1.7 mm	2.6 mm†	2.4 mm†	2.4 mm†
Setup error (rotation) (30)	1.1°	0.6° dg	0.5°	0.9°	0.6°	0.8°

Abbreviations: LR = left-right; SI = superior-inferior; AP = anterior-posterior.

* This value is an average of the uncertainty in target volume delineation near the apex and the seminal vesicles (3.5-mm SD) and for the rest of the prostate (about 2-mm SD).

† These values are estimates of the preparation error without the use of a correction protocol.

zero; i.e., fractionation effects were ignored. To obtain a realistic dose–response curve, we assumed that α had an SD of 0.08 Gy^{-1} over the population of patients (32). The biologic parameters were varied in some experiments to test the robustness of the results. The EUD was computed according to a definition that is similar to the ones used by Marks *et al.* (10) and Niemierko *et al.* (11): EUD is a dose that, delivered homogeneously over the target volume, has the same TCP as the nonhomogeneous dose distribution under consideration. The EUD values were computed for a prescribed dose of 80 Gy to the isocenter. The EUD is used as an effective dose that has a one-to-one relation with the TCP for a particular target. The advantage of using EUD instead of TCP is that the EUD is a relative measure that is less sensitive to biologic parameters such as tumor volume and tumor cell density (11).

The simulations were performed at different levels of conformity between the dose distribution and the CTV. First, we used a spherical CTV with a spherical symmetric dose distribution with a Gaussian penumbra. The CTV has a radius of 2.5 cm. To create the dose distribution, first the 3D image was created of a solid sphere with a radius of 2.5 cm, plus the required margin to the 50% isodose surface. Then this 3D image was blurred with a Gaussian filter described by an SD of $\sqrt{\sigma_p^2 + \sigma^2}$, where σ_p describes the penumbra width (We used 3.2 mm), and σ gives the SD of the random errors. In reality, such a perfectly conforming dose distribution (i.e., where the isodose surfaces have a constant distance to the CTV) can never be realized. However, this situation gives the highest achievable conformity at a given margin. Our previously published mathematic analysis was based on such a dose distribution (3), and the mathematic solution for the probability distribution of the minimum dose to the CTV is known for this case. Next, we analyzed a simple conformal prostate cancer irradiation technique that is based on three beam's-eye–view-shaped fields, using a multileaf collimator or blocks. This technique is applied to the majority of our prostate cancer patients, using a uniform margin between CTV and PTV of 10 mm (in 3D space). In this study, a range of margins was tested.

The dose distributions were obtained from various sources. The spherical dose distribution was computed analytically, whereas the realistic plans were obtained from our 3D treatment planning system (U-MPlan, University of Michigan, Ann Arbor, MI) and from a simple approximate "treatment planning system." The latter system, which was used for practical reasons (high speed and integration with the evaluation software), accurately models the geometry of the treatment beams and the dose gradients in the high-dose region.

RESULTS

A full analysis of a single treatment plan takes about 3 min on an 800-MHz PC; this includes TCP computation for 5,000 systematic errors at 20 values for α, and for prescribed dose values of 0 to 100 Gy in steps of 1 Gy. Values for TCP_{pop} as a function of dose were obtained by weighted summation over the 20 α levels used and the 5,000 systematic errors.

Algorithm validation

In the first experiments, we tried to reproduce our mathematic analysis (3) of the probability distribution of the minimum cumulative dose in a population of patients. In this experiment, the CTV was spherical, and a spherical symmetric dose distribution with a Gaussian penumbra was used. Figure 3a shows the results obtained after determining the proper sampling parameters. The analytic and numerically obtained curves overlap (within 1%) when 5,000 or more preparation errors and a dose computation grid of 2 mm or smaller are used. For the best agreement, the CTV needs to be sampled with 10,000 points or more. When the dose grid is coarser than 2 mm, the curves deviate from each other (Fig. 3b). Using too few preparation error samples has a similar effect. When the CTV is sampled with 1,000 or less points, the results become slightly too optimistic (in particular for small margins), implying that the CTV is undersampled (Fig. 3c).

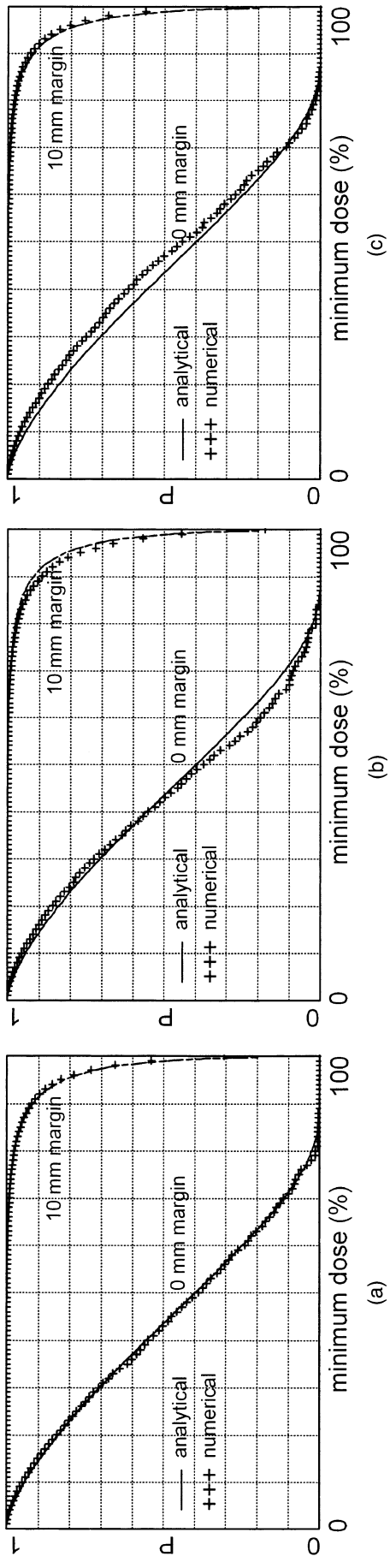

Fig. 3. This figure shows a comparison between the analytic and numeric computation of the probability (P) of the minimum cumulative CTV dose for a population of patients. The CTV was a sphere of 2.5 cm diameter, and the dose distribution had a spherical symmetry with Gaussian penumbra (σ_p = 3.2 mm). The error levels used are given in Table 1, but rotations were excluded. (a) By using a dose grid with a spacing of 2 mm and sampling the CTV with 10,000 points, an agreement within 1% is achieved. (b) Here a dose grid size of 4 mm was used, which is too coarse. (c) Here the CTV was sampled with too few points (1,000), resulting in a slightly too optimistic analysis, because the edge of the CTV is undersampled.

Analysis of realistic plans and the effect of rotational errors

Next, we analyzed several three-field prostate plans and tested the effect of rotations on the probability distributions of minimum dose and EUD of the CTV (prostate and seminal vesicles). In these experiments, the error levels of Table 1 were used (including or excluding the rotations). Figure 4 shows typical examples of the probability distribution of the minimum CTV dose and EUD for a uniform CTV to PTV margin of 10 mm. The EUD curves run higher than the minimum dose curves, implying that a certain reduction of minimum dose corresponds to a much lower reduction in EUD. When including rotations, the probability of reaching a 95% minimum dose is reduced from 90% to 60%. The probability of reaching a 98% EUD is decreased from 99% to 90%. However, these differences have only a small impact on TCP_{pop} (a 0.5% reduction due to rotations), indicating that a PTV margin of 10 mm is adequate even when rotational errors are included. When the margin is reduced, the effect of rotational errors increases. Table 2 lists, for the three tested prostate cases at a PTV margin of 8 mm, TCP_{pop} and population statistics of the minimum CTV dose and the EUD. For this margin, the effect of rotations is significant (up to 2% loss of TCP_{pop}), indicating that an 8-mm margin might be too small in the presence of rotations. The difference in TCP_{pop} values without rotations is mainly because of differences in CTV volume; e.g., Case 2 has the smallest prostate volume and therefore the highest TCP_{pop}. The differences in the effect of rotations are caused by differences in CTV shape. For instance, the prostate for Case 3 is much more spherical than the other two cases, and the corresponding plan is therefore much less sensitive to rotational errors.

Relation between systematic errors and margin for a spherical CTV

In a first analysis of the effect of margin choice on tumor control probability, we varied both margin and preparation error level for the spherical case (perfect conformation). The execution error SD was kept constant (2.6-mm SD), and rotations were excluded. Figure 5a shows that at a certain margin choice, the TCP_{pop} remains constant for small preparation error levels, until a threshold value is reached above which the TCP_{pop} starts to decrease. However, note that with a margin of zero, the TCP_{pop} is already lower at a preparation error SD of zero. This effect is the result of a combination of the penumbra and the execution errors. The plan is made in such a way that the 95% dose level of the unblurred dose distribution is placed around the PTV (which equals the CTV for a margin of zero). The execution errors cause a further blurring of the dose distribution, which widens the penumbra. At zero margin, this penumbra enters the CTV and lowers the TCP, even when the SD of the preparation error is zero. The decrease in TCP_{pop} for increasing SD of the systematic error indicates the presence of geometric misses: failure to control the tumor for some of the cases because of a geometric error. The dotted line at 85.2% indicates a TCP_{pop} that is 1% lower than the maximum

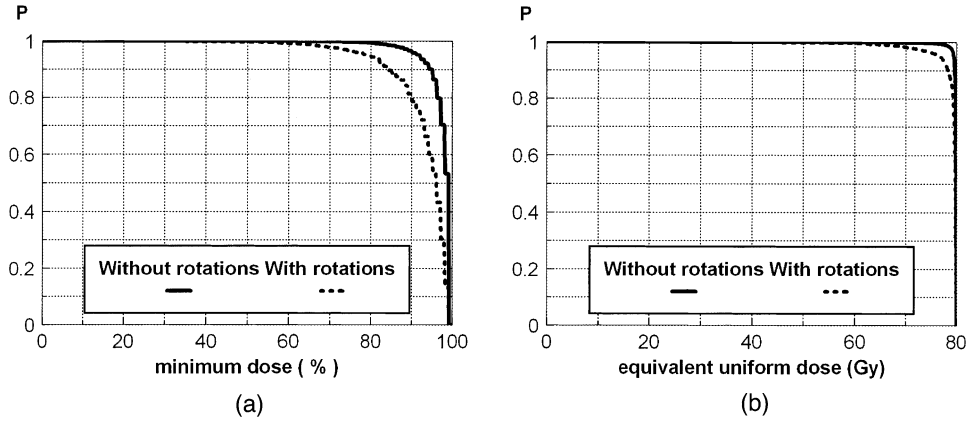

Fig. 4. This is an evaluation of a three-field prostate plan (The prescribed dose in the isocenter was 80 Gy) in which the CTV includes the prostate plus the seminal vesicles, and a uniform margin of 10 mm was used between the CTV and PTV. Dose population histograms, which show cumulative dose probability (P) distributions over a population of patients, are shown for the (a) minimum CTV dose and (b) equivalent uniform CTV dose. The solid curve excludes rotational errors, whereas the dotted curve includes rotations. The error levels used are given in Table 1. The impact of the rotations is mainly caused by the seminal vesicles, which rotate out of the field, because they are far from the rotation point located at the apex of the prostate.

TCP_{pop} (i.e., the TCP for a large margin and zero error at the prescribed dose of 80 Gy). In Fig. 5b, the error level is plotted at which the TCP_{pop} intersects this 1% loss line as a function of the margin choice. The straight line is a least-squares fit that is given by $2.55\ \Sigma - 1.6$ mm, where Σ is the SD of the systematic errors (The SD of the random errors was kept constant in this experiment).

We repeated the analysis of Fig. 5 with different biologic parameters. The curves for 1% TCP_{pop} loss were determined for four different situations (Fig. 6). Even though the maximum value of TCP_{pop} (i.e., the TCP for a large margin, zero error, and a prescribed dose of 80 Gy) was very different (between 0.145 and 1.000), the 1% TCP_{pop} loss curves are similar, indicating that the analysis is not very sensitive to the precise values of the biologic parameters.

Relation between systematic errors, random errors, and margin for a spherical CTV

In the following experiment on the perfectly conformal spherical case, we varied the SDs of both preparation and execution errors in steps of 1-mm SD from 0 to 10 mm. TCP_{pop} values were computed for PTV margins from 0 to 14 mm in steps of 2 mm. To summarize this huge amount of data, curves were constructed at TCP_{pop} loss levels of 1%, 2%, and 5% (Fig. 7). Again the reference situation (i.e., 0% TCP_{pop} loss) was the maximum value of TCP_{pop}, i.e., at a large margin and with zero error for the prescribed dose of 80 Gy. The curves follow more or less perfect ellipses, with the exception of the curves for a margin of 0 mm. This is the same effect that occurs in Fig. 5a, i.e., a TCP reduction at zero margin, because the penumbra enters the CTV. We found that the measured curves could be well fitted by a single collection of elliptical

Table 2. Effect of rotational errors on three prostate cases, planned with a margin of 8 mm between the CTV and the PTV (The isocenter dose was 80 Gy) on the population-averaged TCP_{pop}, and on the statistics of the minimum dose and EUD delivered to the CTV

			Minimum CTV dose (%)			EUD to CTV (Gy)		
	Rotations	TCP_{pop}	Mean	SD	90% CI	Mean	SD	90% CI†
Case 1	Yes*	0.829	88.4	11.6	73.2	78.1	4.9	75.2
Case 2	Yes*	0.854	86.6	11.3	71.1	77.4	4.9	72.4
Case 3	Yes*	0.838	92.0	7.6	82.1	79.2	2.2	78.5
Case 1	No	0.850	94.1	6.5	87.0	79.4	1.6	79.2
Case 2	No	0.873	91.9	7.7	82.2	79.0	2.5	77.9
Case 3	No	0.840	93.7	6.8	86.0	79.4	1.9	79.0

* These results include both random and systematic rotations. All statistics were computed over 5,000 treatments.

† CI stands for confidence interval, i.e., 90% of treatments would result in a minimum CTV dose or EUD larger than the given value. Because of the distinct asymmetry of the EUD distributions, mean and SD cannot be used to derive the CI.

Abbreviations: CTV = clinical target volume; PTV = planning target volume; TCP = tumor control probability; EUD = equivalent uniform dose; SD = standard deviation.

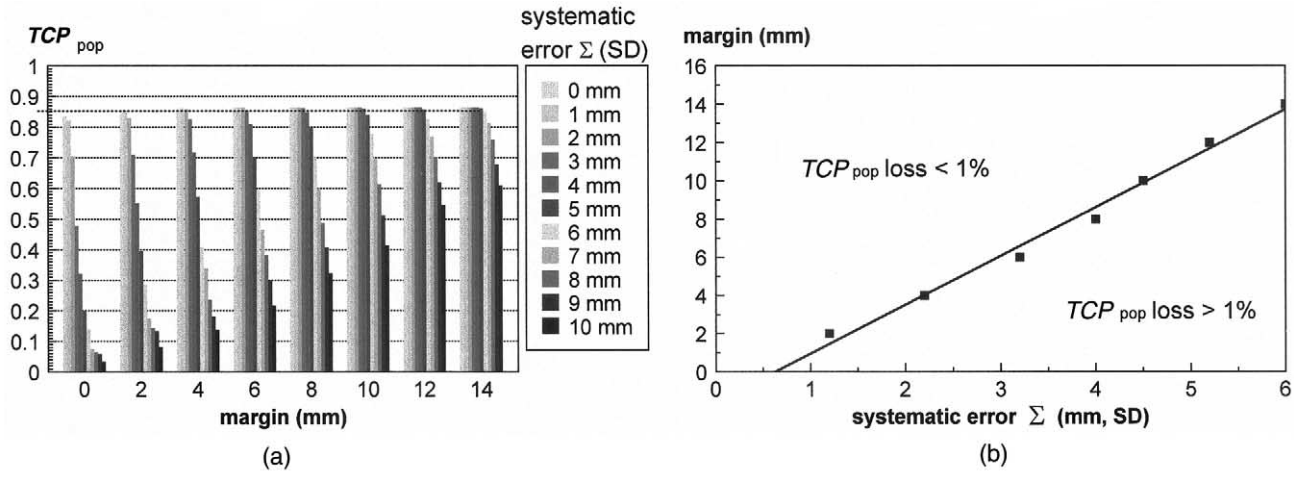

Fig. 5. (a) Effect of the level of preparation (systematic) errors on the population-averaged tumor control probability (TCP_{pop}) for a spherical CTV of 2.5 cm radius treated by a perfectly conformal dose distribution with varying margins (The isocenter dose is 80 Gy). The execution (random) error levels used are given in Table 1,, whereas rotations were excluded. The dotted line is a TCP_{pop} level of 85.2%, which is 1% lower than the maximum level of TCP_{pop}. (b) By determining the margin at which the TCP_{pop} decreases by 1%, a margin recipe is found. The straight line is a least-square fit, which is described by 2.55 Σ − 1.6 mm, where Σ is the SD of the preparation errors.

curves (Fig. 8 shows an example for 1% TCP loss). The best fit for Fig. 7a (1% TCP_{pop} loss) is given by margin = $\sqrt{2.7^2\Sigma^2 + 1.6^2\sigma^2} - 2.8$ mm. The best fit for Fig. 7b (2% TCP_{pop} loss) is given by margin = $\sqrt{2.8^2\Sigma^2 + 1.4^2\sigma^2} - 3.8$ mm. The best fit for Fig. 7c (5% TCP_{pop} loss) is given by margin = $\sqrt{2.6^2\Sigma^2 + 1.2^2\sigma^2} - 4.4$ mm.

Relation between margin, TCP, and dose for a realistic plan

To test the impact of more complicated shapes and dose distributions, we next computed population-averaged dose–response curves for various margin choices for the three-field prostate plans. These plans were obtained with the approximate treatment planning system, which allowed easier variation of the margin. The error levels of Table 1 were used (including rotations). Again, the TCP_{pop} values were computed for a CTV that consisted of the prostate plus the seminal vesicles. Figure 9 shows a typical example, where the TCP_{pop} is plotted as a function of the prescribed dose to the isocenter. The overall structure of these results is the same as in the previous experiments. Above a certain mar-

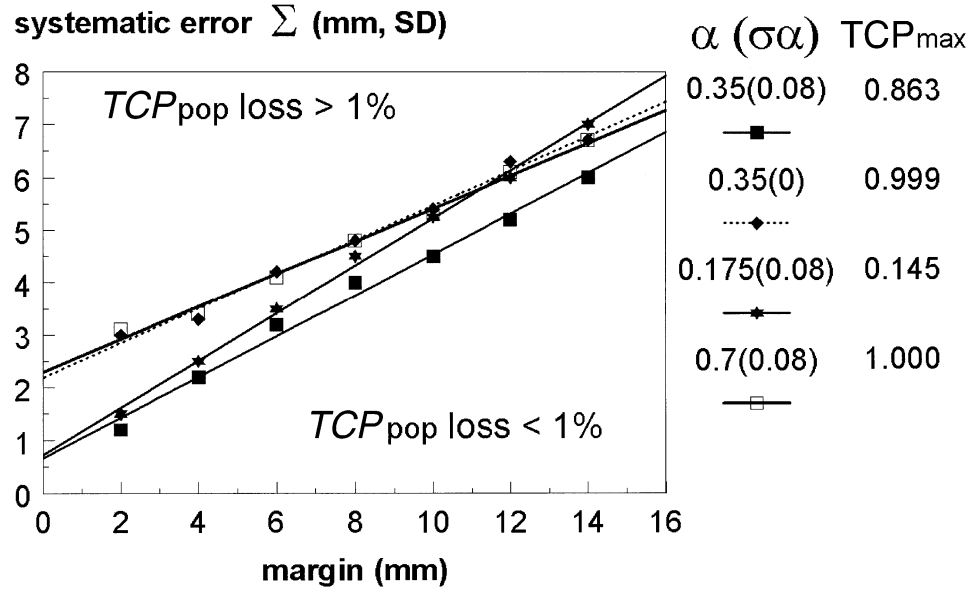

Fig. 6. Margin recipe determination for different biologic parameters. TCP_{max} gives the highest TCP value achieved for a given curve, i.e., for a large margin and zero error. The axes for these curves are reversed compared with Fig. 5b. The relation between systematic error and margin for a given TCP_{pop} loss of 1% differs only slightly, even though the computations were performed for completely different biologic parameters (α and σ_α of the Nahum-Webb model), resulting in completely different TCP_{max} levels between 0.145 and 1.000.

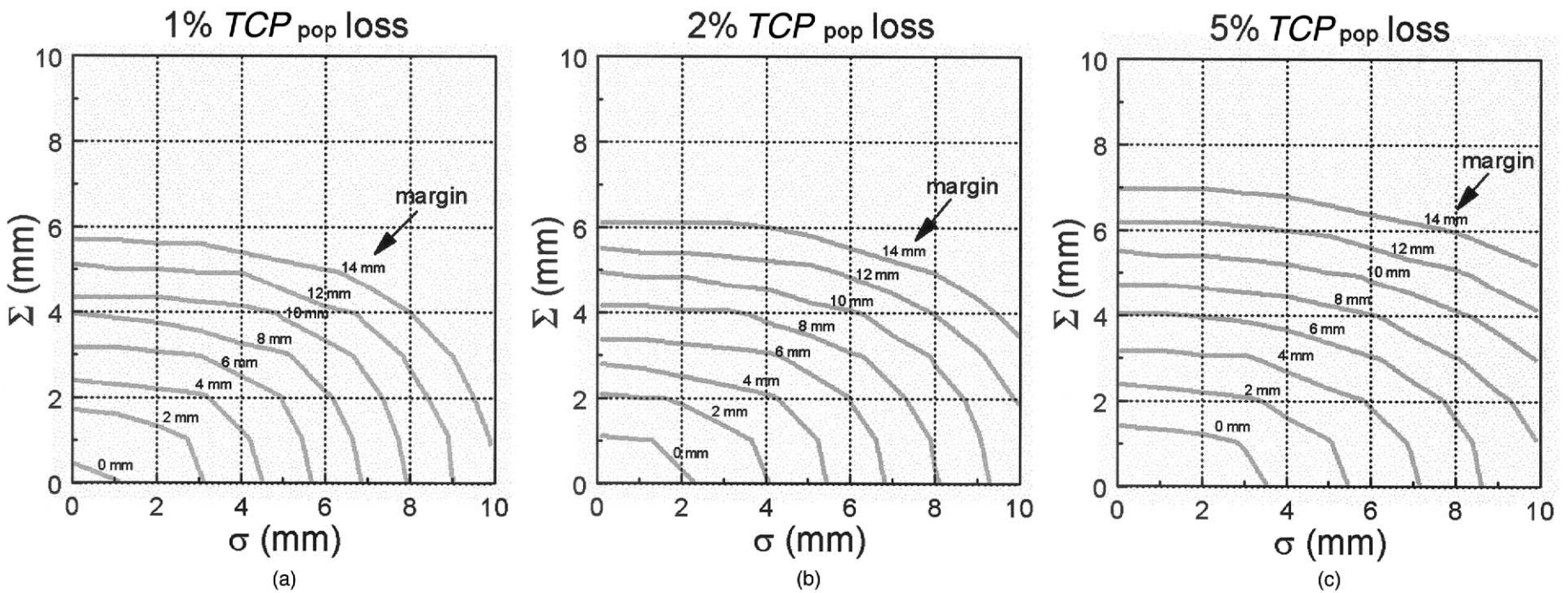

Fig. 7. In this experiment, preparation (Σ) and execution error levels (σ) were varied in steps of 1 mm, whereas the margin was varied in steps of 2 mm. A spherical CTV of 2.5-cm radius treated by a perfectly conformal dose distribution was used. The curves show the margins at which the TCP_{pop} is reduced by (a) 1%, (b) 2%, or (c) 5% compared with the maximum value of TCP_{pop}, i.e., with a large margin and zero errors.

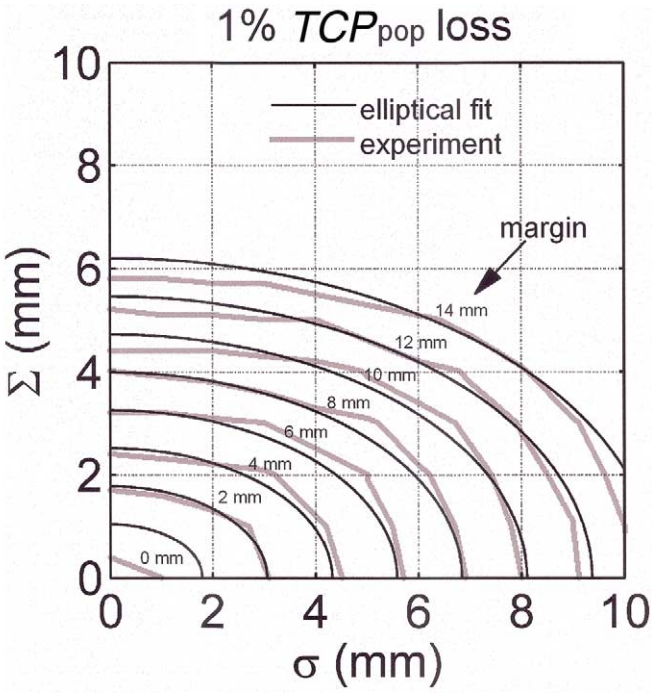

Fig. 8. The TCP_{pop} loss curves shown in Fig. 6 can be well fitted by a single set of elliptical curves, with the exception of the curve for 0 mm margin. At 0 mm margin, the penumbra region overlaps the CTV, causing a larger loss of TCP_{pop}.

gin value, the TCP_{pop} curves overlap, whereas the curves decrease sharply when smaller margins are used. These plots allow analysis of the possibility of compensation for the geometric misses caused by (too small) margins by an increase in dose. For instance, at a dose of 80 Gy and a margin of 10 mm, a TCP_{pop} level of 87% is achieved. To maintain the same TCP_{pop} level when using a margin of 6 mm, a dose level of 87 Gy is required. At a margin of 4 mm, the required dose level is 97 Gy. For the other tested prostate cases, the absolute TCP_{pop} values differed by a few percent (mainly because of different prostate volumes), but the relation between margin and dose was the same.

DISCUSSION

General points

This paper offers a framework for the inclusion of geometric uncertainties in treatment plan evaluation, which is based on the simulation of thousands of possible treatments and the computation of probability distributions of dose-related and biologic parameters. With sufficient sampling of the preparation (systematic) probability distribution, dose grid, and CTV, the numeric solution results in the same probability distribution (within 1%) for the minimum cumulative CTV dose in a population of patients as our previously published analytic solution (3). One notable finding is that a dose grid of 2 mm or finer is required to correctly describe the dose gradients at the edge of the PTV, even though trilinear interpolation was used (Compare Fig. 3a and 3b). For the best agreement, the CTV needs to be sampled with 10,000 points. This finding is consistent with the dose grid requirement of 2 mm. Because the volume of our test CTV was 65 cm^3, each of the 10,000 points represents a volume of 0.0065 cm^3. A cube with sides of about 2 mm has the same volume. Because the minimum dose is the most sensitive to geometric errors, the accuracy of other physical dose parameters and TCP and EUD is better. This observation was verified by repeating the Monte-Carlo simulation for several cases (data not shown). However, to remain on the safe side, we did not reduce the sampling.

The idea that execution (random) variations can be included in treatment planning evaluation is very old (e.g., Ref. 13). Also, a few groups have tested the impact of preparation (systematic) errors by recomputing DVHs using a few possible shifts (20–22). In this study, a large number of preparation errors are simulated such that the probability distributions of dose and biologic parameters can be computed in great detail for a population of patients. This approach has the advantage that large errors are also simulated and scored according to their probability and that errors in all possible directions are tested.

One limitation of this study is that the use of a simple blurring operator to take execution variations into account is not fully correct. First, this approach does not take fractionation effects into account. In future, we will account for this effect by blurring the normalized total dose instead of the physical dose (23). Second, if the number of fractions is small, the concept of dose blurring is invalid, and the computation of the cumulative dose should be included in the Monte Carlo simulations. Third, the displacement of the computed dose distribution assumes that any inhomogeneities have only a small impact. Because the effect of execution errors is minor anyway (3), these assumptions are reasonable. Similarly, the use of a shift operator assumes that the delivered dose does not vary significantly with geometric variations. In the pelvis or the head, which are fairly homogeneous, this is a valid assumption (16, 18, 19). For lung tumors, more extensive calculations or measurements are required to test the effect of geometric errors (22). In addition, a shift in the beam direction has some impact on the dose because of the inverse square law (i.e., 2% dose change per cm shift). Because in most cases beams will be given from other directions as well, this effect is reduced accordingly and can be ignored.

A more complete simulation of the effect of geometric errors in radiotherapy has been performed by Killoran *et al.* (9). They replanned the dose distributions in a simulation system that included both execution (random) and preparation (systematic) errors. These simulations were, however, extremely time-consuming. Also, no probability distributions were computed. Our approach is somewhat less accurate in its dose derivation (Inhomogeneities were not taken into account, which is reasonable for the prostate [18]), but it simulates a larger range of errors and is fast enough to be applicable as a routine evaluation tool for planned dose distributions.

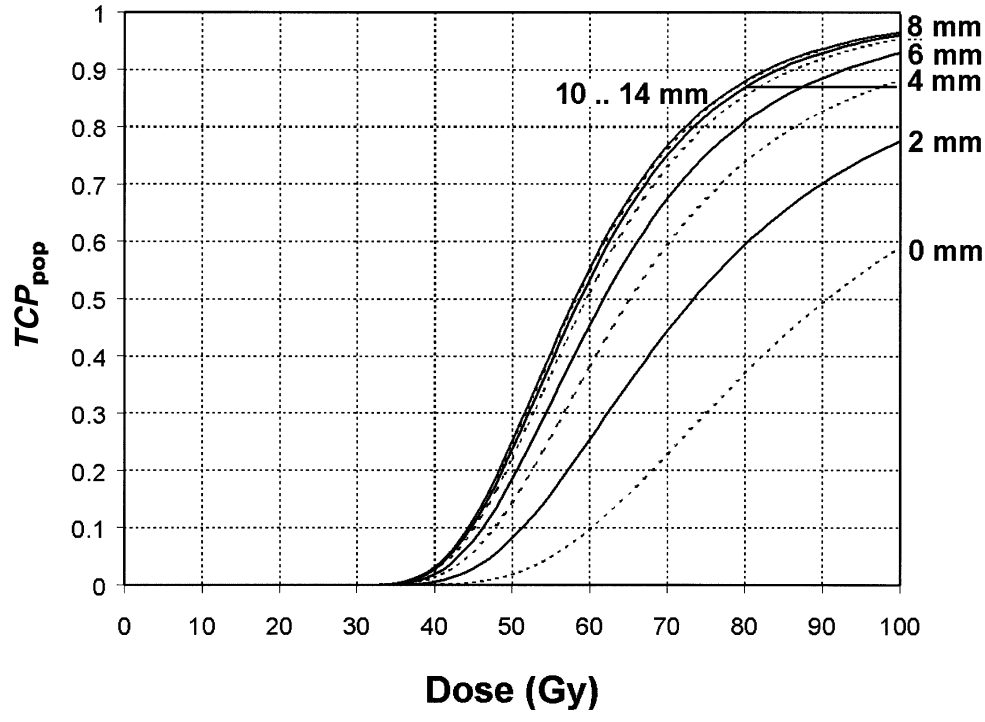

Fig. 9. Population-averaged tumor control probability (TCP_{pop}) as a function of the dose level for a prostate case (including seminal vesicles) treated with three multileaf collimator–shaped fields with varying margins. The error levels used are given in Table 1 (rotations were included). The curves are almost the same for margins of 10, 12, and 14 mm. The TCP_{pop} levels start to decrease when smaller margins are used, indicating the presence of geometric misses. Increasing the dose can partly counteract the loss in TCP_{pop}. For instance, to maintain the TCP_{pop} level corresponding to a margin of 10 mm at 80 Gy (See the horizontal line) while decreasing the margin to 6 mm, a dose increase of 7 Gy is required.

We found that rotational errors have some impact on the treatment of prostate and seminal vesicles (Fig. 4 and Table 2). This finding is obvious if one considers that rotational errors were not included in the margin generated by the treatment planning system. Remeijer *et al.* describe an algorithm to compute the required nonuniform margins for preparation (systematic) errors (34). In future work, their algorithm will be tested using the method described in this paper. We also evaluated the effect of off-center rotations on a spherical CTV treated with a spherical symmetric dose distribution. Here the effect was smaller, but it was still not zero (data not shown).

Table 2 gives, besides values for TCP_{pop}, some population statistics, i.e., the mean, SD, and 90% confidence interval, for the minimum CTV dose and the EUD over a population of patients. There is no simple relation between the population statistics of the dose and TCP_{pop}. This effect is caused by the distinct asymmetry of the distributions of dose values (See, e.g., Fig. 4). Although the majority of patients receive an adequate dose to the CTV, there are a few patients in whom large reductions in dose occur, i.e., when the preparation error is larger than the applied margin. These few patients have a large effect on the population statistics. In our opinion, a confidence interval (or a probability distribution) of a dose variable is therefore a better clinical measure of the robustness of a plan than a population mean of such a variable.

The dose–effect curves in Fig. 9 show the relation between TCP, margin, and dose. One might be tempted to reduce the margin between CTV and PTV and use the decrease in normal tissue dose to allow for dose escalation. However, there is a steep increase in dose required to maintain the same TCP_{pop} when the margin is reduced too much. For instance, to maintain a TCP_{pop} of 88%, the dose must be increased from 80 Gy to 81.5 Gy (2%), 87 Gy (9%), and 97 Gy (21%) for margins of 10, 8, 6, and 4 mm, respectively. Depending on the dose gradients, this increase in dose counteracts the gain achieved by reducing the margin. Assuming a dose gradient of 5% per mm (which would typically occur for a part of a normal structure adjacent to the PTV), the dose at the edge of the normal structure will be reduced by 10% for a change in margin of 2 mm. By combining this dose reduction with the required dose increase to maintain the same TCP_{pop}, the dose at the edge of the normal structure will change by −8%, −11%, and −9% for a margin change from 10 to 8, 6, and 4 mm, respectively. Interestingly, there seems to be an optimal margin (of 8 mm) that gives the lowest dose at the edge of the normal tissue for the same TCP_{pop}. However, in general, the dose gradient will be smaller for most parts of the normal structure. Therefore, this optimum will not be valid for points inside the normal tissue structure. In other words, reducing the margin below 10 mm for prostate cases given the error levels used in this paper while maintaining the same TCP_{pop}

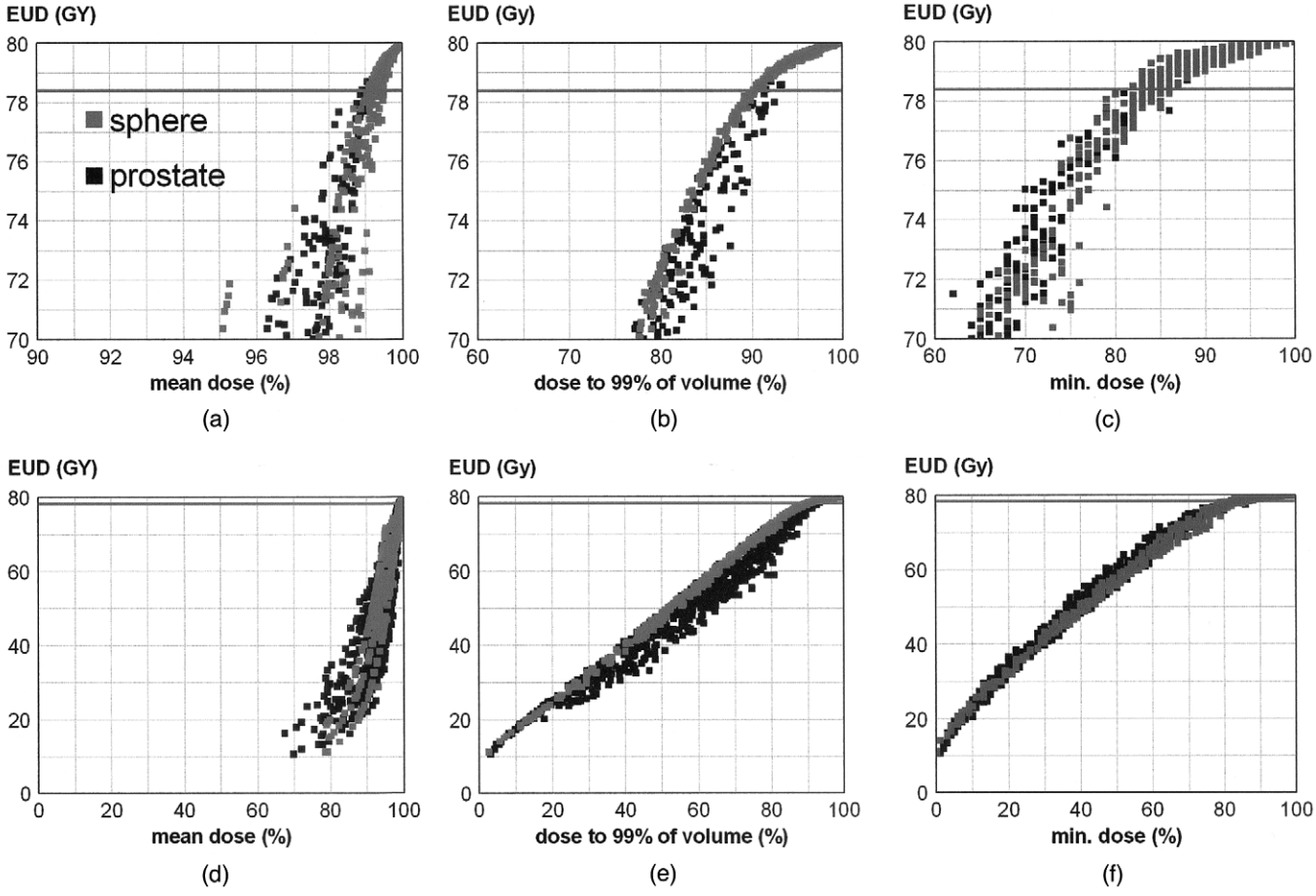

Fig. 10. Relation between several physical dose parameters and EUD as determined from the Monte Carlo experiments for a spherical tumor and three prostate cases (pooled data). In these plots, the SD of the preparation error was set to 10 mm to generate a sufficiently wide distribution. The SD of execution errors was varied in three steps: 0 mm, 2.6 mm, and 4 mm. The margin between CTV and PTV varied between 0 and 14 mm. (a–c) Relation of EUD with mean CTV dose, minimum dose delivered to 99% of the CTV volume, and minimum CTV dose. These plots have an expanded scale in the high-dose region. (d–f) Same plots, but now showing the full dose range.

level will in general increase the dose to the normal structure and will be counterproductive.

The effect of different magnitudes of execution and preparation geometric uncertainties on TCP_{pop} was shown in Figs. 5–9. Both the linear fit in Fig. 5 and the elliptical fit in Fig. 8 suggest that a simple margin recipe for TCP_{pop} loss can be defined. Previously, we computed the probability of the minimum dose in a spherical target treated with a spherical dose distribution analytically (3). Based on these computations, a margin recipe was defined to assure that a given minimum cumulative dose level (e.g., 95%) is delivered to the CTV with a certain confidence (e.g., 90%). To obtain a link between this physical margin recipe and the analysis in this paper, we tested the relation between EUD and the minimum CTV dose for several simulation experiments for the sphere CTV and for three prostate cases (Fig. 10). The graphs in Fig. 10 contain pooled data points of three prostate cases and a sphere. The tests included margins between 0 and 14 mm, execution errors of 0 to 10 mm SD, and a large systematic error SD of 10 mm to obtain a large spread of CTV dose values. In the same experiments, we tested the relation between two other physical dose parameters (mean dose and minimum dose delivered to 99% of the volume) and the EUD. It turns out that, over the whole dose range, the minimum CTV dose has a fairly consistent relation with the EUD, which does not differ much for the three tested prostate cases and the tested sphere case. The spread in the sphere data in Fig. 10 is caused by the use of different levels of execution error, causing the magnitude of the dose gradient in the blurred dose to vary. The spread in the patient data in Fig. 10 is caused by local differences in shape of prostate and seminal vesicles and by local differences in dose gradient because of the three-field conformal technique.

Apparently, both the minimum cumulative CTV dose and the dose to 99% of the CTV volume can be used as a gauge to describe the magnitude of a geometric error and the associated biologic response. A minimum dose of 84% corresponds, on average, to an EUD of 98% of 80 Gy (78.4 Gy). At higher minimum doses, the EUD saturates. This finding will depend on the biologic parameters and on the magnitude of the dose gradient (As the dose gradient gets

lower, the volume of the target that receives a dose between the minimum dose and 100% increases). The mean dose is less suitable as a gauge, especially at lower EUD values, as shown by the large spread in the EUD–mean dose plot. One should note that one could trivially produce distributions that differ in EUD, although they have the same minimum dose or mean dose. However, we demonstrated that for the tested prostate plans, the minimum dose can be used as a predictor for EUD, because the shape of the high-dose part of the dose distributions is very similar for the tested cases. However, this finding might not be valid for, for instance, intensity-modulated radiation therapy plans with nonhomogeneous dose distributions.

Our findings seem to contradict the findings of Levegrun *et al.*, who found that the minimum dose of the PTV did not correlate with local control as established by core biopsies (35), whereas the mean dose did. Similarly, Nahum and Sanchez-Nieto found that the EUD of the PTV is much closer to the mean than to the minimum PTV dose (36). In these studies, geometric errors were not taken into account, and the PTV dose was reasonably constant. When outcome is correlated with the planned dose distribution without taking possible geometric errors into account, the variability in the minimum dose delivered to the CTV is not taken into account. In addition, a poor homogeneity of the dose distributions over the PTV will complicate the relation between planned dose and delivered dose to the CTV. For instance, the EUD for the CTV might differ significantly as a function of the direction of the preparation (systematic) error, even when the error is small compared with the margin size. In our study, the minimum dose value acts as an indicator of the magnitude of the preparation error resulting in a nonhomogeneous CTV dose with reduced TCP. This is a completely different situation from that in the previous studies, where the mean and the minimum PTV dose (without geometric errors and with a reasonable homogeneous PTV dose) were tested for their usability as an indicator for the prescription dose (35, 36).

Craig *et al.* (21) describe the relation between the geometric measure of CTV coverage and TCP for a distribution of random errors and a few selected systematic errors. They find that errors should be added in quadrature (with which we agree) and that there exists a threshold level for systematic errors. However, we feel that it is very difficult to guarantee a threshold for the systematic error. For instance, if portal imaging is used, the magnitude of the setup error can be controlled. However, there will still be a distribution of organ motion and delineation uncertainties that must be taken into account in a stochastic analysis. They find a range of TCP values for a particular coverage; they find also that the relation between coverage and TCP depends on the use of random or systematic errors. In our opinion, the minimum of the cumulative CTV dose (i.e., the sum of the dose of all fractions) is a better gauge of the biologic effect of geometric errors than the geometric concept of coverage. The minimum dose takes into account the shape and the gradient of the dose distribution and therefore has a better one-to-one relation with the TCP and the EUD.

When inspecting Fig. 5, it is apparent that a few-millimeter reduction in the margin may lead to a significant drop in TCP_{pop} once the margin is below a certain threshold level. It is our advice, therefore, to be always extremely careful about margin reduction, even when new equipment is used to improve the accuracy of one or more aspects of the treatment. Because there are many sources of error in radiotherapy (See, for example, Table 1, though there are probably other error sources, such as machine tolerances, as well), there will always be a residual error level, and it is therefore dangerous to use very small margins.

Margin recipes

In the final part of this paper, we would like to discuss margin recipes. In this paper we simulated the effect of geometric errors on biologic response predictors. The purpose of this part of the paper is to try to define a simple analytic margin recipe for TCP_{pop} loss. However, because no analytic expression has been derived to express TCP_{pop} loss, we will try to find a surrogate margin recipe that comes close, using an EUD-based analysis. Many authors defined margin recipes without separating execution (random) and preparation (systematic) errors (37). Bel *et al.* determined a margin recipe for random errors only (38). Recently, our group has established a margin recipe that ensures a minimum dose to the CTV of 95% for 90% of the patients (3). In that case, the margin between CTV and PTV is required to be 2.5 times the total SD of preparation (systematic) errors (Σ), plus 1.64 times the total SD (σ) of execution (random) errors combined with the penumbra minus 1.64 times the SD describing the Gaussian penumbra (σ_p). The 95% isodose surface of the nonblurred dose distribution must cover the PTV. For variable dose and probability levels, this margin recipe can be written as follows (Eq. 1):

$$m_{ptv}=a\Sigma+b\sigma-b\sigma_p, \tag{1}$$

where a is obtained from a chi-square distribution for three dimensions, and b is derived from a one-sided cumulative normal distribution. Van Herk *et al.* (3) give tables for the parameters in this equation (We use here a and b instead of α and β in the previous work, because α is used as a biologic parameter in this paper). Several authors (39, 40) give equations similar to those used in our derivation. Note that m_{ptv}, Σ, σ, and σ_p are vectors, thus allowing this equation to be used for defining nonisotropic margins. For a σ_p of 3.2 mm, a 90% probability level, and a 95% dose level, this recipe can be approximated by (Eq. 2):

$$m_{ptv}=2.5\Sigma+0.7\sigma', \tag{2}$$

where $\sigma' = \sqrt{{\sigma_m}^2 + {\sigma_s}^2}$, i.e., the combination of the SD of all treatment execution (random) variations (organ movement σ_m and setup error σ_s). Note that this approximation is

only valid over a limited range of uncertainties and for a given penumbra shape (That is, the factor given as 0.7 depends on the penumbra width and on the range of uncertainties, but it is identical to results found in earlier simulations [38]).

In this paper, we correlated the minimum dose with the EUD and found that 84% minimum CTV dose corresponds roughly with 98% EUD, both for the spherical cases and for the three-field prostate cases (Fig. 10). A small modification of Eq. 1 allows the computation of the PTV margin (where the 95% isodose line must cover the PTV) for an arbitrary minimum dose level (e.g., corresponding with a required EUD level), as follows (Eq. 3):

$$m_{ptv}=a\Sigma+c\sigma-b\sigma_p, \tag{3}$$

where c must be derived from a one-sided cumulative normal distribution and equals 1.0 for the 84% dose level. This margin recipe assures 84% minimum CTV dose (which corresponds to 98% EUD) for 90% of the patients. For a σ_p of 3.2 mm, this recipe is approximated in Eq. 4:

$$m_{ptv}=2.5\Sigma+0.7\sigma'-3 \text{ mm}, \tag{4}$$

with m_{ptv}, Σ, and σ' in mm. Strikingly, this recipe is almost identical to the 95% minimum dose recipe, except for a reduction of the margin by 3 mm. This reduction of 3 mm corresponds to the distance between the 95% and 84% dose level in the dose distribution for small values of σ. When the minimum dose is 84%, there will be a shell (of about 3-mm thickness) on one edge of the CTV with dose levels between 84% and 95%.

Apparently, the biologic effect (described by an EUD of 98%) of this 3-mm shell with reduced dose is similar for the four tested cases in this paper. This effect is related to the volume of the reduced-dose region. Because the tumor cell density is assumed constant, and the volume of the shell is reasonably constant, the biologic effect of the shell will be the same, irrespective of the direction of the geometric error. However, the margin recipe will depend on the biologic tumor parameters and the tumor volume.

The plot in Fig. 5b corresponds quite well to Eq. 4. Substituting an execution (random) error level, σ', of 2.6 mm, Eq. 4 would read 2.5 Σ − 1.2 mm, which is almost identical to the least-squares fit in Fig. 5b. Remember that the data points in Fig. 5b were derived from a TCP_{pop} loss of 1% for a spherical case. Apparently, 2% loss of EUD for 10% of the patients corresponds well with 1% TCP_{pop} loss (This finding would depend on the shape of the distribution of EUD values). Eq. 4 shows that the systematic errors require 3.5 times (2.5/0.7) more margin than a random error of the same size. The reason for this finding is that execution errors blur a dose distribution that is unsharp to begin with, because of the penumbra. For small values of σ, therefore, the cumulative dose hardly changes, and only a very small extra margin is required. It is therefore most efficient to address preparation (systematic) errors first when working to improve the quality of radiotherapy.

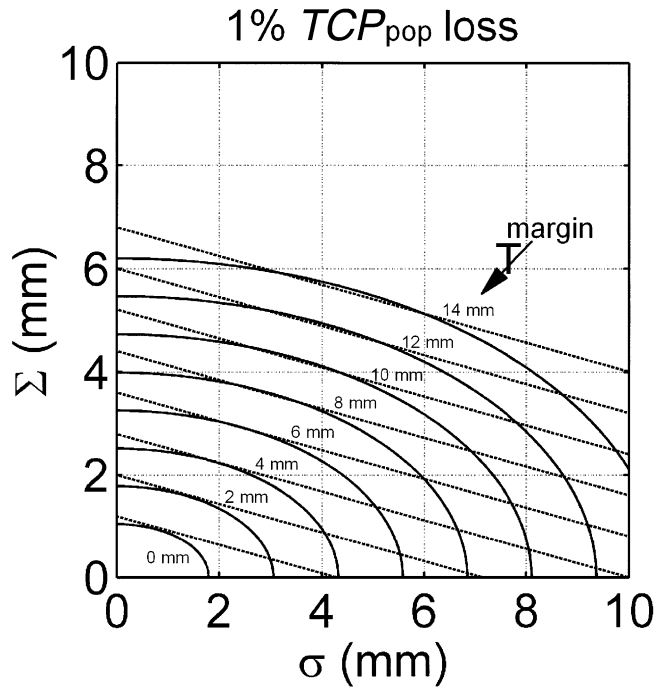

Fig. 11. Comparison between the simple linear margin recipe developed in this paper based on the 90% probability level of delivering 98% EUD and the elliptical fit for the 1% TCP_{pop} loss. For lower values of the random error, σ, both margin recipes are in close agreement.

It might seem that the linear margin recipe defined in Eq. 4 is in contradiction to the elliptical TCP_{loss} curves shown in Figs. 7 and 8. To compare these recipes, curves for Eq. 4 and the elliptical fit from Fig. 8 are shown in Fig. 11. It seems that both margin recipes are fairly consistent for values of σ between 0.5 Σ and Σ. If σ is larger than σ_p, the dose gradients will start to decrease, causing the volume between the 84% and 95% isodose surfaces to increase. This effect invalidates a margin recipe based on the position of a certain isodose surface; i.e., a larger margin is required for TCP_{pop} than for the minimum dose for large values of σ.

By adding the uncertainties listed in Table 1 quadratically, one can show that the combined execution (random) errors are 2.2 mm, 2.5 mm, and 3.2 mm in the left-right, superior-inferior, and anterior-posterior directions, respectively, and that the combined execution errors are 3.2 mm, 4.0 mm, and 4.1 mm, respectively. Based on the recipe of Eq. 4, a sufficient margin would be 6.5 mm, 8.7 mm, and 9.5 mm. However, this analysis did not include rotations. In an analysis that did include rotations, we found that the TCP_{pop} starts to decrease at margins of 10 mm or less (Fig. 9). We can therefore conclude that for our institute, a margin of 10 mm is safe for prostate treatments given the error levels of Table 1.

Another numeric approach is to use coverage probabilities to derive margins (40, 41). Stroom *et al.* defined a

margin recipe based on average DVHs that is similar to the recipe developed in this paper, as follows (Eq. 5):

$$m_{ptv}=2.0\Sigma+0.7\sigma' \quad (5)$$

This recipe assures that a high percentage of the CTV (on average >99% over the population of patients) receives a high dose (>95%) (40). A fundamental problem of coverage probabilities is that they describe probabilities for individual voxels in space. The problem of deriving statistics concerning the whole CTV from these individual probabilities has not been solved. However, the similarity between Eq. 5 and Eq. 2 is probably related to the interaction between the error distributions and the dose gradients. Apparently, a drop of 1% in average DVH at 95% dose occurs when the minimum dose to the CTV is less than 95% for about 10% of the patients.

In future work, we will include more biologic parameters in the evaluation. We plan to test the effect of tumor cell density variations between the GTV and CTV (e.g., Ref. 42) and to include the probability of tumor cell presence for certain regions (43). It is to be expected that the reduced tumor cell density between the edge of the GTV and the CTV will reduce the required margin for some tumor types, i.e., that our specified margins are too generous. However, for prostate cancer, the situation is different. Generally, the entire prostate is defined as the CTV, although the most likely location of tumor foci is in the peripheral zone, at the posterior edge of the prostate (44). This means that the most critical margin lies at the posterior edge of the prostate (CTV) near the rectum, the location at which margin reduction is generally applied. As a result, the margins computed in this paper are reasonable for the prostate.

CONCLUSIONS

We have presented a practical method to quantitatively evaluate the effect of preparation (systematic) and execution (random) errors on target dose and TCP_{pop}. Preparation errors have a much larger impact on target dose and TCP_{pop} than execution errors do. It was demonstrated that for three-field conformal prostate treatments, rotational errors (in addition to translational errors) are important in the sense that they decrease the probability of delivering a high EUD and thereby also decrease the TCP_{pop}. Given the uncertainty levels for prostate treatments at our institution, it was found that a margin of 10 mm is adequate in the sense that it results in less than 1% TCP_{pop} loss compared with larger margins. A margin reduction from 10 to 6 mm requires a 7-Gy increase in dose to the reference point to maintain the same TCP_{pop} level. This dose increase is such that the reduction of margins below 10 mm without improving the accuracy will most likely increase the normal tissue dose when the same TCP_{pop} is maintained.

In the presence of geometric errors that cause a large dose inhomogeneity over the CTV, the minimum cumulative dose can be used as a gauge for EUD and TCP loss. Both for the spherical CTV and the prostate cases, a minimum CTV dose of 84% corresponds roughly to 98% EUD. Using this relation, a simple margin recipe was derived to guarantee a certain EUD with a certain confidence. It turns out that to give 90% of patients an EUD of at least 98%, the PTV margin must be approximately $2.5\,\Sigma + 0.7\,\sigma - 3$ mm. Here Σ and σ are the combined SD of the preparation and execution errors, respectively. The margin defined by this recipe is such that the TCP_{pop} loss due to geometric errors is about 1%. Based on our accuracy level for prostate treatments (excluding rotations), a nonuniform margin of 6.5 mm, 8.7 mm, and 9.5 mm in left-right, superior-inferior, and anterior-posterior directions, respectively, would therefore be appropriate. When including rotations, a uniform margin of 10 mm is more appropriate.

REFERENCES

1. Rabinowitz I, Broomberg J, Goitein M, McCarthy K, Leong J. Accuracy of radiation field alignment in clinical practice. *Int J Radiat Oncol Biol Phys* 1985;11:1857–1867.
2. Bijhold J, Lebesque JV, Hart AA, Vijlbrief RE. Maximizing setup accuracy using portal images as applied to a conformal boost technique for prostatic cancer. *Radiother Oncol* 1992; 24:261–271.
3. van Herk M, Remeijer P, Rasch C, Lebesque JV. The probability of correct target dosage: Dose-population histograms for deriving treatment margins in radiotherapy. *Int J Radiat Oncol Biol Phys* 2000;47:1121–1135.
4. Van de Steene J, Van den Heuvel F, Bel A, *et al*. Electronic portal imaging with on-line correction of setup error in thoracic irradiation: Clinical evaluation. *Int J Radiat Oncol Biol Phys* 1998;40:967–976.
5. Balter JM, Lam KL, Sandler HM, Littles JF, Bree RL, Ten Haken RK. Automated localization of the prostate at the time of treatment using implanted radiopaque markers: Technical feasibility. *Int J Radiat Oncol Biol Phys* 1995;33:1281–1286.
6. Bel A, van Herk M, Bartelink H, Lebesque JV. A verification procedure to improve patient set-up accuracy using portal images. *Radiother Oncol* 1993;29:253–260.
7. Yan D, Lockman D, Brabbins D, Tyburski L, Martinez A. An off-line strategy for constructing a patient-specific planning target volume in adaptive treatment process for prostate cancer. *Int J Radiat Oncol Biol Phys* 2000;48:289–302.
8. Drzymala RE, Mohan R, Brewster L, *et al*. Dose-volume histograms. *Int J Radiat Oncol Biol Phys* 1991;21:71–78.
9. Killoran JH, Kooy HM, Gladstone DJ, Welte FJ, Beard CJ. A numerical simulation of organ motion and daily setup uncertainties: Implications for radiation therapy. *Int J Radiat Oncol Biol Phys* 1997;37:213–221.
10. Marks LB. Radiosurgery dose distributions: Theoretical impact of inhomogeneities on lesion control. *Acta Neurochir Suppl (Wien)* 1994;62:13–17.
11. Niemierko A. Reporting and analyzing dose distributions: A concept of equivalent uniform dose. *Med Phys* 1997;24:103–110.
12. Webb S, Nahum AE. A model for calculating tumour control probability in radiotherapy including the effects of inhomoge-

neous distributions of dose and clonogenic cell density. *Phys Med Biol* 1993;38:653–666.
13. Leong J. Implementation of random positioning error in computerised radiation treatment planning systems as a result of fractionation. *Phys Med Biol* 1987;32:327–334.
14. Stavrev PV, Stavreva NA, Round WH. A new method for optimum dose distribution determination taking tumour mobility into account. *Phys Med Biol* 1996;41:1679–1689.
15. Ragazzi G, Mangili P, Fiorino C, *et al*. Variations of tumor control and rectum complication probabilities due to random set-up errors during conformal radiation therapy of prostate cancer. *Radiother Oncol* 1997;44:259–263.
16. Keall PJ, Beckham WA, Booth JT, Zavgorodni SF, Oppelaar M. A method to predict the effect of organ motion and set-up variations on treatment plans. *Australas Phys Eng Sci Med* 1999;22:48–52.
17. Lujan AE, Larsen EW, Balter JM, Ten Haken RK. A method for incorporating organ motion due to breathing into 3D dose calculations. *Med Phys* 1999;26:715–720.
18. McCarter SD, Beckham WA. Evaluation of the validity of a convolution method for incorporating tumour movement and set-up variations into the radiotherapy treatment planning system. *Phys Med Biol* 2000;45:923–931.
19. Zavgorodni SF. Treatment planning algorithm corrections accounting for random setup uncertainties in fractionated stereotactic radiotherapy. *Med Phys* 2000;27:685–690.
20. MacKay RI, Graham PA, Moore CJ, Logue JP, Sharrock PJ. Animation and radiobiological analysis of 3D motion in conformal radiotherapy. *Radiother Oncol* 1999;52:43–49.
21. Craig T, Battista J, Moiseenko V, van Dyk J. Considerations for the implementation of target volume protocols in radiation therapy. *Int J Radiat Oncol Biol Phys* 2001;49:241–250.
22. Engelsman M, Damen EMF, De Jaeger K, van Ingen KM, Mijnheer BJ. The effect of breathing and set-up errors on the cumulative dose to a lung tumour. *Rad Oncol* 2001;60:95–105.
23. Lebesque JV, Keus RB. The simultaneous boost technique: The concept of relative normalized total dose. *Radiother Oncol* 1991;22:45–55.
24. Mageras GS, Kutcher GJ, Leibel SA, *et al*. A method of incorporating organ motion uncertainties into three-dimensional conformal treatment plans. *Int J Radiat Oncol Biol Phys* 1996;35:333–342.
25. Yan D, Jaffray DA, Wong JW. A model to accumulate fractionated dose in a deforming organ. *Int J Radiat Oncol Biol Phys* 1999;44:665–675.
26. Niemierko A, Goitein M. Random sampling for evaluating treatment plans. *Med Phys* 1990;17:753–762.
27. Rasch C, Barillot I, Remeijer P, Touw A, van Herk M, Lebesque JV. Definition of the prostate in CT and MRI: A multi-observer study. *Int J Radiat Oncol Biol Phys* 1999;43:57–66.
28. van Herk M, Bruce A, Kroes AP, Shouman T, Touw A, Lebesque JV. Quantification of organ motion during conformal radiotherapy of the prostate by three dimensional image registration. *Int J Radiat Oncol Biol Phys* 1995;33:1311–1320.
29. Bel A, Vos PH, Rodrigus PT, *et al*. High-precision prostate cancer irradiation by clinical application of an offline patient setup verification procedure, using portal imaging. *Int J Radiat Oncol Biol Phys* 1996;35:321–332.
30. Remeijer P, Geerlof E, Ploeger L, Gilhuijs K, van Herk M, Lebesque JV. 3-D portal image analysis in clinical practice: An evaluation of 2-D and 3-D analysis techniques as applied to 30 prostate cancer patients. *Int J Radiat Oncol Biol Phys* 2000;46:1281–1290.
31. Remeijer P, Rasch C, Lebesque JV, van Herk M. A general methodology for three-dimensional analysis of variation in target volume delineation. *Med Phys* 1999;26:931–940.
32. Sanchez-Nieto B, Nahum AE. The delta-TCP concept: A clinically useful measure of tumor control probability. *Int J Radiat Oncol Biol Phys* 1999;44:369–380.
33. McKenzie AL, van Herk M, Mijnheer B. The width of margins in radiotherapy treatment plans. *Phys Med Biol* 2000;45:3331–3342.
34. Remeijer P, van Herk M, Lebesque JV. Margins for translational and rotational uncertainties: a probabilistic approach. *Int J Radiat Oncol Biol Phys.* Accepted for publication, 2002.
35. Levegrun S, Jackson A, Zelefsky MJ, *et al*. Analysis of biopsy outcome after three-dimensional conformal radiation therapy of prostate cancer using dose-distribution variables and tumor control probability models. *Int J Radiat Oncol Biol Phys* 2000;47:1245–1260.
36. Nahum AE, Sanchez-Nieto B. Tumor control probability modelling: Basic principles and applications in treatment planning. *Physica Medica* 2001;17(Suppl. 2):13–23.
37. Austin-Seymour M, Kalet I, McDonald J, *et al*. Three dimensional planning target volumes: A model and a software tool. *Int J Radiat Oncol Biol Phys* 1995;33:1073–1080.
38. Bel A, van Herk M, Lebesque JV. Target margins for random geometrical treatment uncertainties in conformal radiotherapy. *Med Phys* 1996;23:1537–1545.
39. Antolak JA, Rosen II. Planning target volumes for radiotherapy: How much margin is needed? *Int J Radiat Oncol Biol Phys* 1999;44:1165–1170.
40. Stroom JC, de Boer HC, Huizenga H, Visser AG. Inclusion of geometrical uncertainties in radiotherapy treatment planning by means of coverage probability. *Int J Radiat Oncol Biol Phys* 1999;43:905–919.
41. Mageras GS, Fuks Z, Leibel SA, *et al*. Computerized design of target margins for treatment uncertainties in conformal radiotherapy. *Int J Radiat Oncol Biol Phys* 1999;43:437–445.
42. Brahme A, Ågren AK. Optimal dose distribution for eradication of heterogeneous tumours. *Acta Oncol* 1987;26:377–385.
43. Goitein M, Schultheiss TE. Strategies for treating possible tumor extension: Some theoretical considerations. *Int J Radiat Oncol Biol Phys* 1985;11:1519–1528.
44. Chen ME, Johnston DA, Tang K, Babaian RJ, Troncose P. Detailed mapping of prostate carcinoma foci: Biopsy strategy implications. *Cancer* 2000;89:1800–1809.