

## Assessment of radiobiological metrics applied to patient-specific QA process of VMAT prostate treatments

Francisco Clemente-Gutiérrez,<sup>1a</sup> Consuelo Pérez-Vara,<sup>1</sup>

María H. Clavo-Herranz,<sup>2</sup> Concepción López-Carrizosa,<sup>2</sup>

José Pérez-Regadera,<sup>3</sup> Carmen Ibáñez-Villoslada<sup>2</sup>

*Sección de Radiofísica, Servicio de Oncología Radioterápica,<sup>1</sup> Hospital Central de*

*la Defensa "Gómez Ulla", Madrid, Spain; Servicio de Oncología Radioterápica,<sup>2</sup>*

*Hospital Central de la Defensa "Gómez Ulla", Madrid, Spain; Servicio de Oncología*

*Radioterápica,<sup>3</sup> Hospital Universitario Doce de Octubre, Madrid, Spain*

*p.clementegutierrez@gmail.com*

Received 4 April, 2015; accepted 19 November, 2015

VMAT is a powerful technique to deliver hypofractionated prostate treatments. The lack of correlations between usual 2D pretreatment QA results and the clinical impact of possible mistakes has allowed the development of 3D verification systems. Dose determination on patient anatomy has provided clinical predictive capability to patient-specific QA process. Dose-volume metrics, as evaluation criteria, should be replaced or complemented by radiobiological indices. These metrics can be incorporated into individualized QA extracting the information for response parameters (gEUD, TCP, NTCP) from DVHs. The aim of this study is to assess the role of two 3D verification systems dealing with radiobiological metrics applied to a prostate VMAT QA program. Radiobiological calculations were performed for AAPM TG-166 test cases. Maximum differences were 9.3% for gEUD, -1.3% for TCP, and 5.3% for NTCP calculations. Gamma tests and DVH-based comparisons were carried out for both systems in order to assess their performance in 3D dose determination for prostate treatments (high-, intermediate-, and low-risk, as well as prostate bed patients). Mean gamma passing rates for all structures were better than 92.0% and 99.1% for both 2%/2 mm and 3%/3 mm criteria. Maximum discrepancies were  $(2.4\% \pm 0.8\%)$  and  $(6.2\% \pm 1.3\%)$  for targets and normal tissues, respectively. Values for gEUD, TCP, and NTCP were extracted from TPS and compared to the results obtained with the two systems. Three models were used for TCP calculations (Poisson, sigmoidal, and Niemierko) and two models for NTCP determinations (LKB and Niemierko). The maximum mean difference for gEUD calculations was  $(4.7\% \pm 1.3\%)$ ; for TCP, the maximum discrepancy was  $(-2.4\% \pm 1.1\%)$ ; and NTCP comparisons led to a maximum deviation of  $(1.5\% \pm 0.5\%)$ . The potential usefulness of biological metrics in patient-specific QA has been explored. Both systems have been successfully assessed as potential tools for evaluating the clinical outcome of a radiotherapy treatment in the scope of pretreatment QA.

PACS number(s): 87.56.Fc, 87.55.Qr, 87.55.dk, 87.55.dh, 87.10.Vg, 87.55.km, 87.53.Bn, 87.55.-x, 87.56.-v

Key words: VMAT QA, pretreatment verifications, 3D dose reconstruction, radiobiological metrics

<sup>a</sup> Corresponding author: Francisco Clemente-Gutiérrez, Sección de Radiofísica, Servicio de Oncología Radioterápica, Hospital Central de la Defensa "Gómez Ulla", Glorieta del Ejército, S/N, 28047, Madrid, Spain; phone: (0034) 914 222 666; fax: (0034) 914 222 655; email: p.clementegutierrez@gmail.com

## I. INTRODUCTION

Radiation therapy (RT) for prostate cancer has substantially evolved during the last years. Dose escalation improves disease control, at the expense of an increment in toxicity.<sup>(1)</sup> Modulated techniques can reduce toxicity by optimizing treatment conformation.<sup>(2)</sup> Hypofractionated schemes are suitable in prostate treatment because of the low  $\alpha/\beta$  ratio for the prostate gland.<sup>(3)</sup> Hypofractionated plans delivered with intensity-modulated radiation therapy (IMRT) techniques lead to extended treatment times compared to traditional techniques. Volumetric-modulated arc therapy (VMAT) has been developed due to rotational capabilities recently implemented in conventional linacs.<sup>(4)</sup> Treatment times are noticeably reduced within this new paradigm, making VMAT a powerful tool for hypofractionated prostate treatments.<sup>(5-9)</sup>

VMAT, as one kind of IMRT technique, requires a detailed patient-specific quality assurance (QA) program.<sup>(9)</sup> This independent pretreatment QA process usually consists of comparing dose measurements acquired with phantoms/detectors of regular geometries with treatment planning system (TPS) calculations made under the same conditions.<sup>(10,11)</sup> Ion chambers are used to perform point measurements. Two-dimensional (2D) (plane) dose distributions are measured with several systems: electronic portal imaging devices, films or 2D detector arrays. Tests involving gamma index passing rates are common in these comparisons.<sup>(12,13)</sup> Three-dimensional (3D) verifications start with specifically developed solutions for volumetric techniques.<sup>(14,15)</sup> The lack of correlations between usual 2D pretreatment QA results and the clinical impact of possible mistakes has been established.<sup>(16,17)</sup> In a second step, based on previous conclusions, 3D verification systems are developed under the scope of determining dose on patient anatomy, providing clinical predictive capability to these systems. Solutions for redundant calculations on patient CT information or 3D dose reconstruction from measurements have already been developed.<sup>(18)</sup> Three-dimensional dose calculation and reconstruction have introduced DVH-based metrics in QA process, allowing for dose-volume information comparisons.

The quality of an RT plan has been traditionally judged by dose-volume parameters rather than biological ones. However, dose-volume criteria should be complemented by biological indices.<sup>(19)</sup> Eventually, biological models should be routinely introduced and validated because these models have demonstrated their predictive ability in the evaluation of the treatment outcome. Although a whole replacement of standard DVH-based metrics should not be recommended, the efforts may be addressed in order to validate outcome prediction models, overcoming the traditional evaluation metrics.<sup>(20-23)</sup> Radiobiological data and response parameters — such as generalized equivalent uniform dose (gEUD),<sup>(24,25)</sup> tumor control probability (TCP),<sup>(26-28)</sup> or normal tissue complication probability (NTCP)<sup>(29-35)</sup> — can be obtained from DVH information. Hence, radiobiological metrics can be incorporated into individualized pretreatment QA process. This paper assesses the role of two 3D dose verification systems dealing with radiobiological metrics applied to a VMAT prostate treatment QA program.

## II. MATERIALS AND METHODS

### A. Treatment unit and TPS

VMAT plans were generated with Monaco 3.1 (Elekta; Stockholm, Sweden). Treatments were delivered with a monoenergetic (6 MV) Synergy (Elekta) machine.

### B. 3D dose verification systems

Two 3D verification systems were assessed. Mobius3D software (Mobius Medical Systems, Houston, TX) provides an independent dose calculation engine for treatments generated by TPS. COMPASS (v. 3.1) (IBA Dosimetry, Schwarzenbruck, Germany) is capable of reconstructing dose on patient CT from measurements taken with an associated detector. In addition, it provides an independent and redundant dose verification of TPS calculations, as does Mobius3D.

Both require, as initial information, DICOM treatment plan information (CT images, RTPlan, RTStruct, and RTDose).

### B.1 Mobius3D system description

The software uses stock reference values for common linear accelerators to model the beams. Mobius3D works with a collapsed cone convolution/superposition algorithm independently developed and updated from its original conception.<sup>(36-41)</sup> The algorithm is accelerated through-out graphic processing units (GPUs), increasing the calculation speed significantly compared to CPUs.

### B.2 COMPASS system description

COMPASS consists of two different devices: the detector and associated software. The detector device is a 2D ion chamber array (MatriXX<sup>Evolution</sup>, IBA Dosimetry). It has 1020 ion chambers (0.08 cm<sup>3</sup>) covering an active area of 24.4 × 24.4 cm<sup>2</sup>. This detector has already been evaluated for VMAT pretreatment QA.<sup>(42)</sup> The MatriXX device must be placed in a holder attached to the head treatment unit in order to ensure a rigid rotation of the detector with the gantry. Source-to-detector distance is 100 cm. A buildup thickness of 2.5 cm was used with the previous arrangement. An angle sensor is attached to the gantry in order to associate each measured fluence with its detection angle; the sensor has an angular tolerance of ± 0.6°. COMPASS software requires a beam modeling process fitting basic parameters, as expected for TPS. The model connects with a collapsed cone convolution/superposition algorithm that allows both calculating and reconstructing (from measurements) dose on patient CT. A commissioning process is required for the MatriXX device in the software. It consists of background (20 s) and pre-irradiation (5 Gy or higher) measurements together with a square field (10 × 10 cm<sup>2</sup>) acquisition that automatically corrects detector shifting and rotation. An absolute dose calibration with a known-dose reference field is also required. Sampling time for each measurement was 250 ms.

## C. Dose-response models

### C.1 gEUD

For a nonuniform tumor dose distribution, equivalent uniform dose (EUD) is defined as the uniform dose that yields the same biological effect, if treatment is delivered over the same number of fractions as the nonuniform original dose distribution.<sup>(24)</sup> Niemierko<sup>(25)</sup> proposed a phenomenological expression to extend the previous concept to normal tissues, referred to as the generalized EUD (gEUD):

$$gEUD = \left( \sum_j v_j d_j^a \right)^{1/a} \quad (1)$$

where  $v_j$  is the fractional tumor volume receiving a dose  $d_j$ , and  $a$  is a tissue-specific parameter describing volume effect. For tumors,  $a$  takes negative values; for serial-like structures,  $a$  takes large positive values; and for parallel-like structures,  $a$  takes values close to 1.

### C.2 TCP

Tumor control probability (TCP) can be modeled as a Poisson distribution.<sup>(26)</sup> If the number of initial clonogen cells is  $N_c$  and the clonogen surviving fraction after irradiation with a single fraction is denoted by  $S$ , TCP can be written for a course of  $n$  fractions as:

$$TCP = e^{-N_c S^n} \quad (2)$$

Assuming that the average number of surviving clonogenic cells is an exponential function of the dose, the characteristic sigmoid dose-response curve is obtained. This simple assumption has been replaced by introducing the linear-quadratic (LQ) model to obtain the surviving fraction<sup>(27)</sup> as:

$$S = e^{-\alpha d \left(1 + \frac{d}{\alpha/\beta}\right)} \quad (3)$$

where  $d$  is the dose per fraction and  $\alpha/\beta$  are usual parameters in the LQ model. For an inhomogeneous irradiation of dose  $d_j$  in a fractional volume  $v_j$ , TCP can be calculated as<sup>(28)</sup> (Poisson model for TCP):

$$TCP = \prod_j \exp \left\{ -N_c v_j \exp \left[ -n \alpha d_j \left(1 + \frac{d_j}{\alpha/\beta}\right) \right] \right\} \quad (4)$$

The number of clonogenic cells can be determined using some of the following relations:<sup>(43-45)</sup>

$$N_c = \ln 2 \exp \left[ -D_{50} \alpha d_j \left(1 + \frac{d_j}{\alpha/\beta}\right) \right] \quad (5)$$

$$N_c = \frac{-\ln 2}{\ln \left\{ 1 - \exp \left[ -D_{50} \alpha d_j \left(1 + \frac{d_j}{\alpha/\beta}\right) \right] \right\}} \quad (6)$$

where  $D_{50}$  is the tumor dose required to obtain a TCP of 50%. In addition, an expression for TCP may be obtained if data of clonogenic cells are not available, in terms of sigmoidal dose response parameters as<sup>(43,44,46)</sup> (sigmoidal model for TCP):

$$TCP = (1/2)^{\sum_j v_j \exp \left[ \frac{2\gamma_{50}}{\ln 2} \left(1 - \frac{d_j}{D_{50}}\right) \right]} \quad (7)$$

The  $\gamma_{50}$  parameter is the slope of dose response at TCP of 50%. Another possibility consists of using gEUD concept previously introduced, obtaining TCP as<sup>(28)</sup> (Niemierko model for TCP):

$$TCP = \frac{1}{1 + (D_{50} / gEUD)^{4\gamma_{50}}} \quad (8)$$

### C.3 NTCP

The Lyman model<sup>(29)</sup> describes complication probabilities for uniformly irradiated organ volume. The characteristic sigmoid dose-response curve is described by three parameters. Sigmoid curve dependence on dose is described by  $TD_{50}$  and curve steepness by  $m$ . The magnitude of the volume effect is described by  $n$  parameter in a power-law relationship between the tolerance dose and irradiated volume:

$$TD(v) = TD(1)v^{-n} \quad (9)$$

where  $TD(v)$  is the tolerance dose for a given partial volume fraction  $v$ , and  $TD(1)$  is the tolerance dose for the full volume. For an inhomogeneous irradiation, the Lyman model can be completed with an algorithm to convert a heterogeneous dose distribution into a uniform organ irradiation resulting in the same NTCP. The effective volume method<sup>(31)</sup> is most commonly used to complement the Lyman model, resulting in a combined formalism named the Lyman-Kutcher-Burman (LKB) model. NTCP can be calculated, for an inhomogeneous irradiation of dose  $d_j$  in a fractional volume  $v_j$ , as:<sup>(31,32)</sup>

$$NTCP = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-u^2/2} du \quad (10)$$

where  $t$  is

$$t = \frac{D_{eff} - TD_{50}}{mTD_{50}} \quad (11)$$

with  $D_{eff}$

$$D_{eff} = \left( \sum_j v_j d_j^{1/n} \right)^n \quad (12)$$

Another possibility entails using the concept of gEUD concept previously introduced and obtaining NTCP in the Niemierko model as:<sup>(32,34)</sup>

$$NTCP = \frac{1}{1 + (TD_{50} / gEUD)^{4\gamma_{50}}} \quad (13)$$

## D. Prostate treatments

Prostate treatments were planned and delivered with a single arc. VMAT technique was applied to all prostate cases treated with external beam therapy: treatment for usual staging (high-, intermediate- and low-risk) and radiotherapy after prostatectomy. Prostate gland, seminal vesicles, and pelvic lymph nodes were treated in high-risk patients; prostate gland and seminal vesicles were the targets for intermediate-risk cases; and prostate volume was the single target for low-risk staging. A simultaneous integrated boost (SIB) technique was used with two or three target volumes. A moderate hypofractionation scheme was applied for the previous three staging levels. The prescription doses were 70 Gy to prostate gland, 56 Gy to seminal vesicles, and 50.4 Gy to pelvic lymph nodes, delivered in 28 fractions. For prostate bed treatments after

prostatectomy, the prescription dose was 74 Gy delivered in 37 fractions. Contoured organs at risk (OARs) were usually rectum, bladder, and femoral heads. For each usual staging, 25 prostate cases were analyzed, together with 25 cases of prostate bed treatments, for a total of 100 analyzed treatments.

### E. Verifications of biological metrics for both systems with TG-166 benchmark phantom and test cases

In order to test the capabilities of both systems for radiobiological calculations, tests taken from the AAPM TG-16621 report were performed.

#### E.1 Benchmark phantom test

Benchmark phantom consists of a large cubical phantom with four simple structures (three rectangular, one triangular) created inside the phantom<sup>(21)</sup> (Fig. 1). A single 6 MV, 100 cm source-to-surface distance,  $20 \times 20$  cm<sup>2</sup> photon beam was calculated in the TPS, with a prescribed dose of 72 Gy in 40 fractions to a point at 6 cm depth along the central axis. Dose imparted over benchmark phantom structures was determined by Mobius3D (M3D), COMPASS calculation (CC), and reconstruction (CR) modules; then, gEUD, TCP, and NTCP were calculated and compared among TPS, M3D, CC, and CR from the previous models. An in-house developed software was used to perform previous calculations. The software reads the DVH information from the different systems and applies Eqs. (1), (4), (5), (6), (7), (8), (10), (11), (12), and (13) in order to obtain gEUD, TCP, and NTCP values.

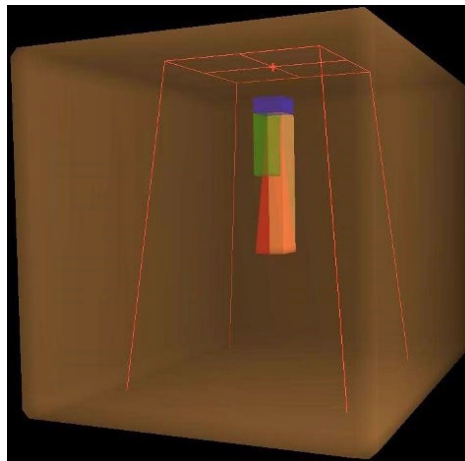


FIG. 1. Benchmark phantom test case from the AAPM TG-166 report,<sup>(21)</sup> with four simple structures (three rectangular, one triangular) and the edge of a single 6 MV, 100 cm source-to-surface distance,  $20 \times 20$  cm<sup>2</sup> photon beam.

#### E.2 Representative test cases

Treatment plans for three representative test cases (head and neck [H&N], prostate, and brain) were also calculated in the TPS according to the volumes and dose prescriptions defined in the report. gEUD, TCP, and NTCP values were determined for TPS plans and compared to those extracted by M3D, CC, and CR for the same treatments.

## F. Evaluation of prostate treatments with classical (gamma) and DVH-based dose-volume metrics

In order to assess correct performance in 3D dose calculation and reconstruction processes, traditional gamma tests and DVH-based comparisons were carried out for both systems. Resorting to classical metrics, comparisons with TPS by means of global gamma passing rates for all structures were reported with two criteria (2%/2 mm and 3%/3 mm, global normalization to maximum, with a low-dose threshold at 10% of global maximum). In addition, 3D dose evaluation was performed by comparing all prostate plans generated by the TPS and those determined by M3D, CC, and CR. Representative dosimetric parameters were obtained from DVHs. ICRU recommendations for recording and reporting IMRT treatments<sup>(20)</sup> were used to extract evaluation parameters for PTVs ( $D_{98}$ ,  $D_2$ ,  $D_{50}$ ,  $D_{\text{mean}}$ ). Maximum and mean doses were extracted for OARs. For normal tissue, depending on the case, classical<sup>(47)</sup> and recently reviewed dose constraints (QUANTEC)<sup>(48)</sup> were also reported.

## G. Introducing biological metrics in patient-specific QA for prostate treatments

gEUD, TCP, and NTCP were extracted from TPS and compared to those values obtained with M3D, CC, and CR for all the models discussed above. These calculations were performed with the same in-house software previously defined. In order to support the introduction of radiobiological metrics, the correlation between biological indices and differences in DVH parameters was studied.

## H. Radiobiological parameters

Parameters used for tumor calculations (gEUD and TCP) for TG-166 test cases and prostate treatments are summarized in Table 1. Values for a parameter were extracted from the TG-166 report<sup>(21)</sup> (value for benchmark phantom case was also taken as -10). Selected  $\alpha$  value was  $0.1 \text{ Gy}^{-1}$ .<sup>(49)</sup> Selected  $\alpha/\beta$  values were 10 Gy for TG-116 cases, with the exception of the prostate case, where 3 Gy was selected,<sup>(45)</sup> therefore taking values recommended in the report. For analyzed prostate treatments,  $\alpha/\beta$  was 1.5 Gy, which is customary in our institution.<sup>(50-53)</sup>  $D_{50}$  and  $\gamma_{50}$  values for benchmark phantom case were taken from the TG-166 report. Values for the remaining TG-116 cases were also extracted from the study by Okunieff et al.<sup>(45)</sup> For

TABLE 1. Selected radiobiological parameters for tumor calculations (gEUD and TCP). Values were obtained from AAPM TG-166 report, Cheung,<sup>(54,55)</sup> King,<sup>(56)</sup> Okunieff,<sup>(45)</sup> and Levegrun<sup>(57)</sup> studies.

			Parameters				
			<i>a</i>	$\alpha$ ( $\text{Gy}^{-1}$ )	$\alpha/\beta$ (Gy)	$D_{50}$ (Gy)	$\gamma_{50}$ (%/%)
TG-166	BPhant	PTVRect	-10	0.1	10	63.3	5
		Rect1	-10	0.1	10	44.2	1.6
	H&N	PTV 70	-10	0.1	10	51.77	2.28
		PTV54	-10	0.1	10	51.77	2.28
		PTV50	-10	0.1	10	51.77	2.28
	Prostate	PTV70.2	-10	0.1	3	46.29	0.95
		Brain	GTV54	-10	0.1	10	22.17
	PTV50.4		-10	0.1	10	22.17	0.70
Analyzed prostate cases	Cheung and King studies	HR PTV	-10	0.1	1.5	75.5	1.7
		IR PTV	-10	0.1	1.5	67.5	2.2
		LR PTV	-10	0.1	1.5	57.3	1.4
		Bed PTV	-10	0.1	1.5	66.8	2.54
	Okunieff Levegrun	All staging	-10	0.1	1.5	46.29	0.95
		PTV	-10	0.1	1.5	70.5	2.9

BPhant = benchmark phantom; PTVRect = PTV Rectangle; Rect1 = Rectangle 1; H&N = head and neck; HR = high-risk; IR = intermediate-risk; LR = low-risk.

the analyses of high-, intermediate-, and low-risk prostate treatments,  $D_{50}$  and  $\gamma_{50}$  values were extracted from studies by Cheung et al.<sup>(54,55)</sup> For prostate bed treatments, parameters were taken from the study by King et al.,<sup>(56)</sup> using the relationship between absolute and relative slope at  $D_{50}$ .<sup>(45)</sup> In addition, TCP values for the prostate cases under analysis were calculated with the values reported by Okunieff et al.<sup>(45)</sup> and Levegrun et al.<sup>(57)</sup>

The parameters used for normal tissue calculations (gEUD and NTCP) for TG-166 test cases and prostate treatments are summarized in Table 2. Values for “a” parameter were extracted from TG-166 report.<sup>(21)</sup> The selected “a” values for benchmark phantom case were taken as 12, 1, 4, and 2 for the PTV Rectangle, Rectangle 1, Rectangle 2, and Triangle 1 structures, respectively. These values were selected in order to fit the results from both LKB and Niemierko models (“a” parameter was not reported for the Niemierko model for NTCP calculations in TG-166 report). Selected  $\alpha/\beta$  values were 3 Gy for all the cases. Values for  $TD_{50}$ ,  $\gamma_{50}$ , m, and n parameters were shown in the previous report, taken from the study by Burman et al.<sup>(58)</sup> Additional  $\gamma_{50}$  parameters were taken from studies by Stavrev et al.<sup>(59)</sup> (cord and mandible), Huang et al.<sup>(60)</sup> (inner ear), and Lee et al.<sup>(61)</sup> (parotid gland). The  $\gamma_{50}$  parameter for the pubic bone was taken as 4, similar to the value for the femoral head. As in the Niemierko model case, m and n parameters from LKB were not reported in the AAPM report; standard values for bone (m = 0.12 and n = 0.25) were taken in order to perform the calculations.

TABLE 2. Selected radiobiological parameters for normal tissue calculations (gEUD and NTCP). Values were obtained from AAPM TG-166 report and the studies by Burman et al.,<sup>(58)</sup> Stavrev et al.,<sup>(59)</sup> Huang et al.,<sup>(60)</sup> and Lee et al.<sup>(61)</sup>

			Parameters					
			a	$\alpha/\beta$ (Gy)	$TD_{50}$ (Gy)	$\gamma_{50}$ (%/%)	m	n
TG-166	BPhant	PTVRect	12	3	80	3	0.12	0.25
		Rect1	1	3	75.1	2.8	0.12	0.25
		Rect2	4	3	55.3	3.1	0.12	0.25
		Triang1	2	3	46	1.8	0.12	0.25
	H&N	Cord	20	3	66.5	2.6	0.175	0.05
		Parotid	1	3	46	2.2	0.18	0.7
		Mandible	10	3	72	3.1	0.1	0.07
	Prostate	Rectum	8	3	80	4	0.15	0.12
		Bladder	8	3	80	4	0.11	0.5
		Fem head	12	3	65	4	0.12	0.25
		PBone	12	3	65	4	0.12	0.25
	Brain	BStem	16	3	65	3	0.14	0.16
		OptCh	16	3	65	3	0.14	0.25
		Eye	16	3	65	2	0.19	0.2
		OptN	16	3	65	3	0.14	0.25
		Inner ear	16	3	65	1.74	0.095	0.01
Prostate cases	Rectum	8.33	3	80	4	0.15	0.12	
	Bladder	2	3	80	4	0.11	0.5	
	Fem head	4	3	65	4	0.12	0.25	

BPhant = benchmark phantom; PTVRect = PTV Rectangle; Rect = Rectangle; Triang1 = Triangle 1; H&N = head and neck; Fem head = femoral head; PBone = pubic bone; BStem = brain stem; OptCh = optic chiasm; OptN = optic nerve.

## I. Statistical analysis

Results were described as mean  $\pm$  standard deviation (SD). Data were compared using a paired and two-tailed Student’s *t*-test. The difference was considered statistically significant for *p*-values  $< 0.05$ . Possible correlation between variables was studied by means of Pearson’s correlation coefficient (*r*).



### III. RESULTS

#### A. Verifications of biological metrics for both systems with TG-166 benchmark phantom and test cases

Differences and comparisons between gEUD, TCP, and NTCP calculations from TPS and those obtained from Mobius3D and COMPASS for TG-166 test cases are shown in Tables 3, 4, and 5, respectively. DVHs from previous cases are plotted in Figs. 2 and 3. Maximum deviations for gEUD were found for left and right inner ear in M3D calculations (8.8% and 9.3%, respectively). CR results were better than M3D and CC results for gEUD discrepancies. The maximum difference for TCP evaluations was found for M3D calculations using the Poisson model (-1.3%). M3D and CR results were better than those from CC. For NTCP, the maximum discrepancy was found in the COMPASS reconstructed dose of Triangle 1 structure from benchmark phantom test (5.3%). There were no statistically significant differences for NTCP comparisons.

TABLE 3. Comparisons of gEUD calculations (target volumes and normal tissues) for TG-166 test cases between TPS results and those obtained from Mobius3D and COMPASS.

	TPS Values	Differences (%)		
		M3D	CC	CR
<i>Benchmark Phantom</i>				
PTV Rectangle (as target)	73.6	-0.8	-0.1	-0.3
PTV Rectangle (as normal tissue)	56.1	0.7	1.0	0.7
Rectangle 1 (as target)	72.1	-1.2	-0.2	-0.1
Rectangle 1 (as normal tissue)	55.7	1.1	-2.2	-2.4
Rectangle 2	50.1	-0.4	1.4	1.5
Triangle 1	40.9	2	2.6	2.7
<i>Head &amp; Neck</i>				
PTV 70	72.2	0.0	0.0	-1.0
PTV 54	59.4	-1.0	-1.1	-1.7
PTV 50	52.9	-4.2	-0.4	-3.3
Cord	35.5	-0.9	-0.7	-1.1
Left Parotid	20.4	0.8	5.2	-2.6
Right Parotid	19.9	-3.4	1.6	-5.0
Mandible	45.3	6.9	2.4	0.5
<i>Prostate</i>				
PTV	69.7	0.3	0.4	0.7
Rectum	42.7	4.1	0.9	4.3
Bladder	43.1	-1.3	-0.2	-2.6
Left Femoral Head	13.9	0.1	3.6	1.0
Right Femoral Head	14.2	-0.5	2.4	1.7
Pubic Bone	39.6	2.8	3.4	-1.9
<i>Brain</i>				
GTV 54	56.3	-0.3	0.2	0.0
PTV 50.4	54.5	0.9	0.1	-0.2
Brain Stem	41.5	0.3	-0.8	1.1
Optic Chiasm	46.5	-0.8	0.0	-2.5
Left Eye	11.0	0.0	1.6	-7.7
Right Eye	11.1	-0.8	1.6	-2.6
Left Optic Nerve	29.5	7.4	1.5	-8.3
Right Optic Nerve	33.8	3.2	2.6	-6.0
Left Inner Ear	35.2	8.8	3.2	5.3
Right Inner Ear	43.5	9.3	4.0	5.5
		M3D vs. CC	M3D vs. CR	CC vs. CR
	p-values	0.93	< 0.05	< 0.05

M3D = Mobius3D; CC = COMPASS dose calculation; CR = COMPASS dose reconstruction.

TABLE 4. Comparisons of TCP calculations for TG-166 test cases between TPS results and those obtained from Mobius3D and COMPASS, using the three described models (Poisson, sigmoidal, and Niemierko).

		TPS Values	Differences (%)		
			M3D	CC	CR
<i>Benchmark Phantom</i>					
PTV Rectangle	TCP Poisson	82.6	-1.3	-0.1	-0.4
	TCP Sigmoidal	93.3	-0.9	-0.2	-0.3
	TCP Niemierko	95.3	-0.7	-0.1	-0.2
Rectangle 1	TCP Poisson	83.7	0.6	1.2	0.9
	TCP Sigmoidal	84.4	0.4	1.0	0.8
	TCP Niemierko	82.2	0.6	0.9	0.7
<i>Head &amp; Neck</i>					
PTV 70	TCP Poisson	94.4	0.0	0.0	-0.6
	TCP Sigmoidal	95.0	0.0	0.0	-0.5
	TCP Niemierko	95.4	0.0	0.0	-0.4
<i>Prostate</i>					
PTV 70.2	TCP Poisson	98.7	0.1	0.1	0.1
	TCP Sigmoidal	84.6	0.2	0.3	0.7
	TCP Niemierko	82.6	0.2	0.2	0.4
<i>Brain</i>					
GTV 54	TCP Poisson	98.9	0.0	0.0	0.0
	TCP Sigmoidal	97.0	-0.1	0.1	0.0
	TCP Niemierko	93.2	0.0	0.0	0.0
		p-values	M3D vs. CC < 0.05	M3D vs. CR 0.15	CC vs. CR < 0.05

M3D = Mobius3D; CC = COMPASS dose calculation; CR = COMPASS dose reconstruction.



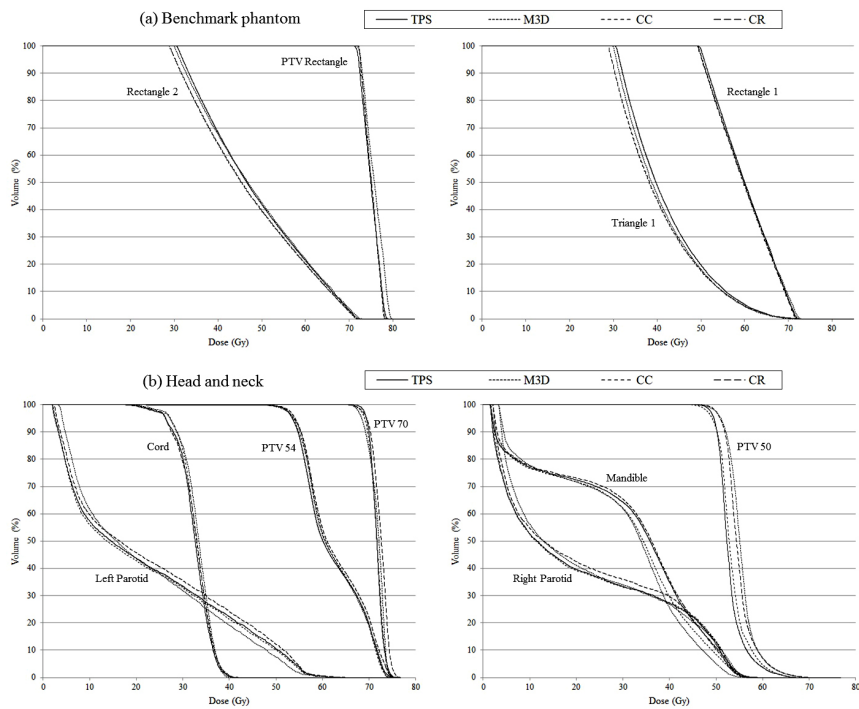


FIG. 2. Cumulative dose-volume histograms for TPS, Mobius3D (M3D), and COMPASS dose calculation (CC) and reconstruction (CR) for two AAPM TG-166 test cases: (a) benchmark phantom and (b) head and neck case.

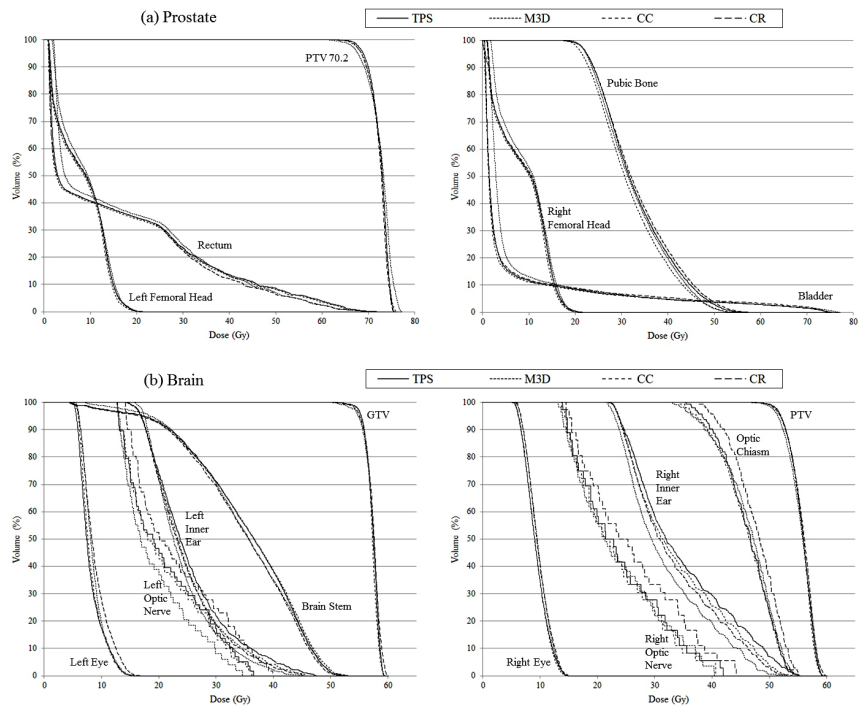


FIG. 3. Cumulative dose-volume histograms for TPS, Mobius3D (M3D), and COMPASS dose calculation (CC) and reconstruction (CR) for two AAPM TG-166 test cases: (a) prostate and (b) brain cases.

## B. Evaluation of prostate treatments with classical (gamma) and DVH-based dose-volume metrics

Mean global gamma passing rates for all structures are shown in Table 6. Considering mean values for all structures, M3D passing rates were worse than those from COMPASS in all cases, with the exception of values for 3%/3 mm criterion in low-risk treatments (mean passing rates for all structures were 99.8% for both M3D and CR results).

TPS mean values of the dosimetric parameters analyzed for each prostate case are shown in Table 7. Mean differences and comparisons between previous values and results for the same parameters determined by M3D, CC, and CR are also presented in this table. Maximum discrepancies for PTVs were found in COMPASS reconstructed low-risk and prostate bed patients ( $2.4\% \pm 0.8\%$ ). The largest differences observed for rectum were found for COMPASS reconstructed mean dose in bed patients ( $6.2\% \pm 1.3\%$ ). For bladder, the maximum differences were also located in mean dose Mobius3D calculated values for bed treatments ( $-3.3\% \pm 1.6\%$ ). The worst mean differences for femoral heads were found for CR results for the right head in low-risk patients ( $3.1\% \pm 1.3\%$ ). CR results were worse ( $p < 0.05$ ) than M3D and CC results for a huge number of cases. Statistically significant differences for comparisons between M3D and CC were also observed. Deviations were better for M3D than CC in some cases, and vice

TABLE 6. Mean gamma passing rates (2%/2 mm and 3%/3 mm, global normalization to maximum with a low-dose threshold at 10% of global maximum) for 100 analyzed prostate treatments, 25 from each group (high-, intermediate-, low-risk, and prostate bed patients).

		Mean Global Gamma Passing Rates (%)					
		Gamma 2%/2 mm			Gamma 3%/3 mm		
		M3D	CC	CR	M3D	CC	CR
HR	Prostate PTV	85.4±8.5	93.7±3.4	86±13	98.9±2.9	99.5±0.7	98.3±4.2
	SV PTV	95.3±4.7	99.4±1.0	96.2±4.2	99.2±1.9	100.0±0.0	99.8±1.0
	LN PTV	55±13	86.8±4.1	90.9±6.6	88.9±7.8	97.3±1.9	99.4±0.9
	Rectum	91.2±5.5	99.3±1.1	96.3±4.0	99.2±1.2	99.9±0.3	99.9±0.2
	Bladder	86.7±7.0	98.5±1.8	97.4±5.7	98.1±2.2	99.9±0.2	99.8±0.8
	L Fem Head	98.5±1.7	98.2±1.6	97.2±2.4	100.0±0.1	100.0±0.0	100.0±0.1
	R Fem Head	97.8±2.5	98.2±1.8	94.7±3.7	100.0±0.0	100.0±0.0	99.9±0.2
	Whole Vol	93.5±1.6	97.8±0.8	98.1±1.2	99.3±0.4	99.7±0.2	99.9±0.2
IR	Prostate PTV	92.9±6.9	99.3±1.0	96.3±2.8	99.1±1.5	99.9±0.1	99.7±0.8
	SV PTV	91.8±7.5	99.8±0.4	97.2±3.0	99.2±1.9	100.0±0.0	99.9±0.2
	Rectum	90.1±5.0	99.9±0.1	93.0±4.1	99.7±0.4	100.0±0.0	99.8±0.6
	Bladder	92.2±5.7	99.9±0.1	99.3±1.0	100.0±0.1	100.0±0.0	100.0±0.0
	L Fem Head	97.7±4.4	99.7±0.8	95.8±3.8	99.9±0.2	100.0±0.0	99.8±0.6
	R Fem Head	97.4±3.9	99.6±1.1	92.9±5.8	99.5±2.1	100.0±0.0	99.5±0.9
	Whole Vol	98.7±0.3	99.8±0.9	99.1±1.7	99.9±0.1	99.8±0.9	99.6±1.7
	LR	Prostate PTV	92.6±5.9	99.5±0.4	96.0±2.8	99.5±1.6	99.9±0.1
Rectum		92.6±3.8	99.8±0.6	93.7±3.8	99.9±0.4	100.0±0.0	99.9±0.2
Bladder		86.4±8.7	99.1±1.0	99.4±0.7	100.0±0.0	100.0±0.1	99.9±0.3
L Fem Head		99.0±1.1	99.8±0.5	98.2±2.3	100.0±0.1	100.0±0.0	99.9±0.3
R Fem Head		98.5±1.9	99.6±0.7	96.0±4.1	99.9±0.2	100.0±0.1	99.7±1.2
Whole Vol		98.7±0.4	99.9±0.0	99.5±0.6	100.0±0.1	100.0±0.0	99.9±0.2
Bed		Bed PTV	87.0±7.8	98.0±0.9	96.2±2.3	98.0±2.1	99.9±0.1
	Rectum	86.9±4.6	99.6±0.9	91.0±5.6	99.2±1.4	100.0±0.0	99.6±0.7
	Bladder	83.5±7.1	99.6±0.3	99.3±1.0	99.8±0.4	100.0±0.0	100.0±0.1
	L Fem Head	98.4±1.6	99.7±0.6	96.3±3.1	99.9±0.3	100.0±0.0	99.9±0.3
	R Fem Head	98.6±1.3	99.7±0.3	93.7±3.6	100.0±0.1	100.0±0.0	99.6±0.7
	Whole Vol	97.7±0.6	99.9±0.0	99.3±0.3	99.9±0.1	100.0±0.0	100.0±0.1

M3D = Mobius3D; CC = COMPASS dose calculation; CR = COMPASS dose reconstruction; HR = high-risk; IR = intermediate-risk; LR = low-risk; SV = seminal vesicles; LN = pelvic lymph nodes; Fem Head = femoral heads; Whole Vol = whole volume.

versa. For all the parameters, mean values were  $0.5\% \pm 1.9\%$ ,  $0.4\% \pm 1.0\%$ , and  $1.1\% \pm 2.6\%$  for M3D, CC, and CR, respectively. Considering all parameters, M3D results were better than CR results, and CC values were better than those obtained by M3D and CR.

TABLE 7. Differences and comparisons for dosimetric parameters (target volumes and normal tissues) for 100 analyzed prostate treatments, 25 from each group (high-, intermediate-, low-risk, and prostate bed patients).

		TPS Mean Values	Mean Differences (%)			p-values		
			M3D	CC	CR	M3D vs. CC	M3D vs. CR	CC vs. CR
<i>High-risk</i>								
Prost PTV	D <sub>98</sub>	78.3±0.6	1.2±1.1	0.2±0.4	0.8±0.8	< 0.05	0.16	< 0.05
	D <sub>2</sub>	85.2±0.5	-0.9±0.8	-1.3±0.6	-1.3±0.6	< 0.05	< 0.05	0.70
	D <sub>50</sub>	82.6±0.4	-0.6±0.6	-0.6±0.3	-1.1±0.6	0.92	< 0.05	< 0.05
SV PTV	D <sub>m</sub>	82.4±0.4	-0.4±0.6	-0.6±0.3	1.5±12.2	0.13	0.44	0.40
	D <sub>98</sub>	63.1±1.9	0.1 ± 1.6	-0.2 ± 0.6	1.9 ± 1.4	0.28	< 0.05	< 0.05
	D <sub>2</sub>	84.0±0.7	-0.1 ± 0.8	-0.3 ± 0.6	-0.7 ± 0.9	0.16	< 0.05	< 0.05
LN PTV	D <sub>50</sub>	75.2±3.9	0.4 ± 1.1	0.0 ± 0.5	0.8 ± 1.2	0.06	0.18	< 0.05
	D <sub>m</sub>	74.8±2.4	0.1 ± 0.9	-0.1 ± 0.4	0.4 ± 0.9	0.39	0.17	< 0.05
	D <sub>98</sub>	54.6±0.5	-1.1±1.2	2.3±0.9	-0.1±0.8	< 0.05	< 0.05	< 0.05
Rectum	D <sub>2</sub>	67.5±5.6	-0.7±0.9	-1±0.4	-0.9±0.8	0.06	0.30	0.61
	D <sub>50</sub>	58.7±0.4	-2.7±0.7	-0.5±0.3	-1.2±0.5	< 0.05	< 0.05	< 0.05
	D <sub>m</sub>	59.2±1.2	-2.3±0.7	-0.2±0.2	-1.0±0.5	< 0.05	< 0.05	< 0.05
Bladder	V <sub>50</sub>	25.5±5.9	1.3±1.9	0.6±0.8	3.2±1.8	< 0.05	< 0.05	< 0.05
	V <sub>60</sub>	14.5±5.2	1.9±1.3	0.5±0.4	2.6±1.1	< 0.05	< 0.05	< 0.05
	V <sub>65</sub>	10.9±4.3	1.8±1.2	0.3±0.3	2.2±0.9	< 0.05	< 0.05	< 0.05
	V <sub>70</sub>	7.9±3.6	1.8±1.2	0.3±0.3	2.0±0.9	< 0.05	0.24	< 0.05
	V <sub>75</sub>	4.4±2.8	1.5±1.4	0.1±0.4	1.5±0.9	< 0.05	0.71	< 0.05
L FemH	D <sub>M</sub>	41.2±2.9	0.6±2.1	1.1±0.9	3.9±1.8	0.16	< 0.05	< 0.05
	V <sub>65</sub>	17.9±6.4	-0.1±0.9	0.3±0.3	-1.0±0.7	< 0.05	< 0.05	< 0.05
	V <sub>70</sub>	13.3±5.1	0.1±0.7	0.2±0.3	-1.0±0.7	0.27	< 0.05	< 0.05
	V <sub>75</sub>	8.3±3.6	0.1±0.8	0.2±0.4	-1.1±0.8	0.69	< 0.05	< 0.05
	V <sub>80</sub>	1.5±1.3	-0.7±1.1	-0.3±0.5	-1.8±1.8	< 0.05	< 0.05	< 0.05
R FemH	D <sub>M</sub>	51±3.2	-1.6±0.9	0.4±0.4	-1.2±0.8	< 0.05	0.13	< 0.05
	D <sub>max</sub>	46.2±1.5	2.6±2.1	2.9±0.9	2.2±1.2	0.47	0.40	< 0.05
	D <sub>max</sub>	46.9±1.2	2.3±1.7	2.7±1.0	2.8±1.1	0.19	0.25	0.75
<i>Intermediate-risk</i>								
Prost PTV	D <sub>98</sub>	78.5±0.9	1.9±1.3	0.3±0.4	2.3±0.7	< 0.05	0.10	< 0.05
	D <sub>2</sub>	85.0±0.4	0.0±0.6	-0.3±0.3	0.0±0.4	< 0.05	0.59	< 0.05
	D <sub>50</sub>	82.8±0.5	0.0±0.6	0.0±0.3	0.1±0.4	0.66	0.66	0.26
SV PTV	D <sub>m</sub>	82.5±0.5	0.4±0.9	0.0±0.3	0.2±0.4	< 0.05	0.46	< 0.05
	D <sub>98</sub>	62.4±1.3	0.0±1.7	-0.3±0.4	1.8±1.2	0.33	< 0.05	< 0.05
	D <sub>2</sub>	83.7±0.7	-0.1±0.8	-0.2±0.3	0.1±0.7	0.69	0.16	0.06
Rectum	D <sub>50</sub>	73.7±3.6	-0.2±1.1	-0.1±0.4	0.9±1.1	0.42	< 0.05	< 0.05
	D <sub>m</sub>	73.9±2.2	-0.3±0.9	-0.2±0.4	0.7±0.8	0.27	< 0.05	< 0.05
	V <sub>50</sub>	24.6±6.4	1.5±0.9	0.6±0.4	3.8±0.9	< 0.05	< 0.05	< 0.05
Bladder	V <sub>60</sub>	15.4±4.8	1.9±0.9	0.4±0.4	3.3±0.9	< 0.05	< 0.05	< 0.05
	V <sub>65</sub>	11.3±4	1.9±0.8	0.4±0.3	2.8±0.9	< 0.05	< 0.05	< 0.05
	V <sub>70</sub>	8.1±3.3	1.9±0.9	0.3±0.3	2.6±1.0	< 0.05	< 0.05	< 0.05
	V <sub>75</sub>	4.4±2.4	1.6±1.0	0.2±0.4	2.2±1.0	< 0.05	< 0.05	< 0.05
	D <sub>M</sub>	36.4±3.6	-0.3±1.4	1.1±0.8	5.5±1.5	< 0.05	< 0.05	< 0.05
L FemH	V <sub>65</sub>	13.1±5.8	0.3±0.7	0.3±0.2	-1.2±0.6	0.71	< 0.05	< 0.05
	V <sub>70</sub>	10.2±4.6	0.4±0.7	0.2±0.2	-1.1±0.6	0.08	< 0.05	< 0.05
	V <sub>75</sub>	6.8±3.1	0.3±0.7	0.1±0.3	-1.3±0.7	0.08	< 0.05	< 0.05
	V <sub>80</sub>	0.8±0.7	-0.8±0.8	-0.3±0.4	-1.1±0.7	< 0.05	< 0.05	< 0.05
	D <sub>M</sub>	30.1±8.2	-2.7±2.1	2.1±0.8	-2.0±1.2	< 0.05	0.23	< 0.05
R FemH	D <sub>max</sub>	36.2±4.8	2.0±1.7	2.2±0.8	3.1±1.6	0.50	< 0.05	< 0.05
	D <sub>max</sub>	36.3±4.9	1.3±1.6	1.9±0.7	2.9±1.3	< 0.05	< 0.05	< 0.05

TABLE 7. (cont'd.)

		TPS Mean Values	Mean Differences (%)			p-values		
			M3D	CC	CR	M3D vs. CC	M3D vs. CR	CC vs. CR
<i>Low-risk</i>								
Prost PTV	D <sub>98</sub>	78.6±0.9	2.1±1.5	0.2±0.4	2.4±0.8	< 0.05	0.33	< 0.05
	D <sub>2</sub>	84.9±0.6	-0.2±0.7	-0.3±0.2	0.1±0.5	0.18	0.06	< 0.05
	D <sub>50</sub>	82.8±0.6	-0.2±0.5	-0.2±0.2	0.1±0.4	0.53	< 0.05	< 0.05
Rectum	D <sub>m</sub>	82.5±0.6	0.1±0.6	-0.2±0.2	0.3±0.4	< 0.05	0.17	< 0.05
	V <sub>50</sub>	22.4±6.4	1.2±0.8	0.5±0.3	3.1±0.7	< 0.05	< 0.05	< 0.05
	V <sub>60</sub>	15.8±5.1	1.7±0.8	0.4±0.2	2.8±0.7	< 0.05	< 0.05	< 0.05
	V <sub>65</sub>	12.9±4.6	1.9±0.9	0.4±0.3	2.9±0.9	< 0.05	< 0.05	< 0.05
	V <sub>70</sub>	9.9±3.9	2.0±0.9	0.4±0.3	2.7±0.9	< 0.05	< 0.05	< 0.05
	V <sub>75</sub>	6.2±3.0	1.9±1.1	0.2±0.4	2.6±1	< 0.05	< 0.05	< 0.05
	D <sub>M</sub>	32.6±5.4	-0.5±1.8	1.4±0.6	6.0±1.1	< 0.05	< 0.05	< 0.05
Bladder	V <sub>65</sub>	14.6±6.9	0.8±0.7	0.4±0.2	-0.9±0.7	< 0.05	< 0.05	< 0.05
	V <sub>70</sub>	11.8±5.7	0.8±0.8	0.3±0.2	-0.9±0.6	< 0.05	< 0.05	< 0.05
	V <sub>75</sub>	8.1±4.2	0.6±0.8	0.2±0.3	-0.9±0.6	< 0.05	< 0.05	< 0.05
	V <sub>80</sub>	1.5±1.2	-0.8±0.9	-0.4±0.4	-0.7±0.9	< 0.05	0.56	0.18
	D <sub>M</sub>	28.7±9.8	-3.7±3.1	2.5±0.7	-1.5±1.5	< 0.05	< 0.05	< 0.05
L FemH	D <sub>max</sub>	35.2±4.3	1.1±1.6	1.7±0.5	2.2±2.6	0.07	0.07	0.33
R FemH	D <sub>max</sub>	34.7±4.1	0.5±1.4	1.7±0.7	3.1±1.3	< 0.05	< 0.05	< 0.05
<i>Prostate Bed</i>								
Bed PTV	D <sub>98</sub>	71.0±0.9	2.1±1.5	0.0±0.5	2.4±0.8	< 0.05	0.13	< 0.05
	D <sub>2</sub>	78.1±0.4	0.3±0.7	-0.4±0.4	0.2±0.5	< 0.05	0.47	< 0.05
	D <sub>50</sub>	75.4±0.4	-0.1±0.8	-0.2±0.3	0.0±0.4	0.17	0.53	< 0.05
Rectum	D <sub>m</sub>	75.2±0.4	0.2±0.8	-0.2±0.3	0.2±0.4	< 0.05	0.91	< 0.05
	V <sub>50</sub>	32.0±7.0	2.3±1.2	0.7±0.3	4.3±1.0	< 0.05	< 0.05	< 0.05
	V <sub>60</sub>	22.9±5.6	3.2±1.4	0.6±0.5	4.3±1.2	< 0.05	< 0.05	< 0.05
	V <sub>65</sub>	17.8±4.9	3.4±1.6	0.5±0.6	4.2±1.4	< 0.05	< 0.05	< 0.05
	V <sub>70</sub>	11.4±3.9	3.3±2.0	0.2±0.9	3.7±1.6	< 0.05	0.08	< 0.05
	V <sub>75</sub>	2.4±1.9	1.4±1.4	0.0±1.0	1.4±1.1	< 0.05	0.96	< 0.05
	D <sub>M</sub>	37.5±5	0.8±1.7	1.3±0.8	6.2±1.3	0.09	< 0.05	< 0.05
Bladder	V <sub>65</sub>	25.6±7.2	-0.1±0.4	0.1±0.1	-1.0±0.7	< 0.05	< 0.05	< 0.05
	V <sub>70</sub>	22.1±6.4	-0.2±0.5	0.0±0.2	-1.2±0.7	< 0.05	< 0.05	< 0.05
	V <sub>75</sub>	9.9±4.6	-2.8±2.8	-0.7±1.8	-2.1±2.4	< 0.05	0.11	< 0.05
	V <sub>80</sub>	0.0±0.0	-0.2±0.8	-0.1±0.5	0.0±0.0	0.78	0.32	0.31
	D <sub>M</sub>	35.4±8.0	-3.3±1.6	1.2±0.6	-0.6±1.3	< 0.05	< 0.05	< 0.05
L FemH	D <sub>max</sub>	41.5±4.1	2.4±1.7	2.2±1.0	2.6±1.2	0.67	0.37	0.05
R FemH	D <sub>max</sub>	41.2±3.4	2.7±1.7	2.4±0.6	2.8±1.4	0.27	0.66	0.09

M3D = Mobius3D; CC = COMPASS dose calculation; CR = COMPASS dose reconstruction; Prost = prostate; SV = seminal vesicles; LN = pelvic lymph nodes; FemH = femoral heads.

### C. Biological metrics applied prostate treatments

Discrepancies in gEUD calculations between TPS, dose calculation and reconstruction for all prostate cases are shown in Table 8, taking previous reported data from different studies. Maximum mean differences were found for rectum COMPASS reconstructed gEUD for intermediate- and low-risk patients ( $4.7\% \pm 1.3\%$ ). Prostate gland absolute mean differences were lower than 1.3% in all cases; the maximum mean discrepancy for bladder was found in intermediate-risk COMPASS reconstructed cases ( $-2.5\% \pm 1.3\%$ ); the worst femoral head deviation was also found in the intermediate-risk COMPASS reconstructed (right femoral head) cases ( $4.2\% \pm 1.4\%$ ). CC results were better than both CR and M3D results for prostate gland values in low-risk treatments, for all rectum values, and for bladder values in high-risk and prostate bed treatments. In addition, CC results were better than CR results for prostate gland values in high-risk treatments, for bladder values in low- and intermediate-risk treatments, and for all femoral head values (excepting left head values for high-risk treatments). M3D results were better than both CR and M3D results for bladder values in low- and intermediate-risk treatments and for

TABLE 8. Calculations and comparisons of gEUD values for 100 analyzed prostate treatments, 25 from each group (high-, intermediate-, low-risk, and prostate bed patients).

		<i>Prostate</i>	<i>Rectum</i>	<i>Bladder</i>	<i>R Fem Head</i>	<i>L Fem Head</i>	
HR	gEUD mean value (Gy)	83.1±0.5	57.2±2.9	52.1±3.0	27.5±2.3	26.7±2.4	
	Mean	M3D	-1.0±0.7	2.8±2.0	-1.3±0.8	1.9±1.6	2.0±1.7
	difference (%)	CC	-0.9±0.4	0.1±0.7	0.4±0.6	3.2±0.7	3.6±0.7
		CR	-1.3±0.7	3.2±1.1	-1.3±1.0	3.6±1.0	3.1±1.3
	p-values	M3D vs. CC	0.4	< 0.05	< 0.05	< 0.05	< 0.05
		M3D vs. CR	0.12	0.23	0.97	< 0.05	< 0.05
		CC vs. CR	< 0.05	< 0.05	< 0.05	< 0.05	0.07
IR	gEUD mean value (Gy)	83.3±0.6	57.3±2.8	37.2±6.8	21.8±3.2	21.2±3.0	
	Mean	M3D	-0.1±0.6	3.1±1.6	-0.5±1.4	1.6±1.7	1.3±1.1
	difference (%)	CC	-0.1±0.3	0.3±0.7	1.3±0.7	2.4±0.7	2.3±0.8
		CR	0.1±0.3	4.7±1.3	-2.5±1.3	4.2±1.4	3.6±1.1
	p-values	M3D vs. CC	0.53	< 0.05	< 0.05	< 0.05	< 0.05
		M3D vs. CR	0.12	< 0.05	< 0.05	< 0.05	< 0.05
		CC vs. CR	0.13	< 0.05	< 0.05	< 0.05	< 0.05
LR	gEUD mean value (Gy)	83.2±0.7	58.6±2.4	37.0±8.0	20.8±2.8	20.7±2.9	
	Mean	M3D	-0.3±0.7	3.4±2.0	-0.3±1.8	0.7±1.2	0.7±1.3
	difference (%)	CC	-0.1±0.3	0.5±0.7	1.4±0.8	2.0±0.6	2.0±0.5
		CR	0.2±0.5	4.7±1.3	-1.9±1.2	3.7±0.9	2.9±1.3
	p-values	M3D vs. CC	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
		M3D vs. CR	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
		CC vs. CR	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05
Bed	gEUD mean value (Gy)	75.7±0.5	58.8±2.1	44.1±5.9	23.2±2.6	23.5±3.1	
	Mean	M3D	-0.2±0.9	3.1±1.7	-1.3±1.0	1.2±1.2	1.2±1.2
	difference (%)	CC	-0.3±0.5	0.3±0.9	0.4±0.9	2.3±0.6	2.2±0.7
		CR	0.0±0.5	3.7±1.2	-1.2±1.2	3.9±1.3	3.7±1.1
	p-values	M3D vs. CC	0.71	< 0.05	< 0.05	< 0.05	< 0.05
		M3D vs. CR	0.08	< 0.05	0.75	< 0.05	< 0.05
		CC vs. CR	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05

M3D = Mobius3D; CC = COMPASS dose calculation; CR = COMPASS dose reconstruction; HR = high-risk; IR = intermediate-risk; LR = low-risk.

all femoral head values. In addition, M3D results were better than CR ones for rectum values in intermediate-risk treatments. Finally, CR results were better than M3D results for prostate gland values in low-risk treatments and also better than CC results for prostate gland values in prostate bed treatments. The remaining differences were not statistically significant.

TCP comparisons between the TPS, M3D, and COMPASS for all the previously described models are shown in Table 9. The maximum mean difference was found for high-risk COMPASS reconstructed values using the Poisson model and taking the values for radiobiological parameters from the study by Cheung et al.<sup>(54)</sup> ( $-2.4\% \pm 1.1\%$ ). CC results were better than CR results for all high-risk treatments (except those obtained using the sigmoidal model and taking the values from the study by Okunieff et al.<sup>(45)</sup>) and prostate bed values. CC and CR results were better than M3D ones for all low-risk values. In addition, CC results were better than M3D results using the Niemierko model with any set of parameters and using the sigmoidal model with the parameters by Levegrun et al.<sup>(57)</sup> The remaining differences for TCP comparisons were not statistically significant.

NTCP comparisons between the TPS and the two systems are shown in Table 10. The worst mean difference was found for COMPASS reconstructed rectum values in low-risk treatments ( $1.5\% \pm 0.5\%$ ). Discrepancies for bladder and femoral heads were better than 0.1% and 0.01%, respectively, in all cases. CC results were better than both CR and M3D results for rectum values in all cases, for bladder values in high-risk and prostate bed treatments and also for intermediate-risk treatments using the Niemierko model. In addition, CC results were better



TABLE 9. Calculations and comparisons of TCP values for 100 analyzed prostate treatments, 25 from each group (high-, intermediate-, low-risk, and prostate bed patients), using the three models and set of parameters described in the text.

Parameters	Model	Mean Differences (%)			p-values		
		M3D	CC	CR	M3D vs. CC	M3D vs. CR	CC vs. CR
<i>High-risk Prostate</i>							
Cheung	Poisson	-2.0±1.2	-1.6±0.8	-2.4±1.1	0.13	0.19	< 0.05
	Sigmoidal	-1.5±1.2	-1.5±0.8	-1.9±1.2	0.90	0.12	< 0.05
	Niemierko	-1.5±1.0	-1.4±0.6	-1.9±1.0	0.41	0.12	< 0.05
Okunieff	Poisson	$(-2.4±1.5) \cdot 10^{-3}$	$(-1.9±1.0) \cdot 10^{-3}$	$(-2.8±1.4) \cdot 10^{-3}$	0.15	0.19	< 0.05
	Sigmoidal	-0.4±0.3	-0.3±0.2	-0.5±0.3	0.99	0.12	0.06
	Niemierko	-0.3±0.2	-0.3±0.1	-0.4±0.2	0.42	0.12	< 0.05
Levegrun	Poisson	-0.7±0.4	-0.5±0.3	-0.8±0.4	0.14	0.19	< 0.05
	Sigmoidal	-1.3±0.9	-1.2±0.5	-1.6±0.8	0.49	0.12	< 0.05
	Niemierko	-1.3±0.9	-1.1±0.5	-1.6±0.8	0.47	0.12	< 0.05
<i>Intermediate-risk Prostate</i>							
Cheung	Poisson	$(-0.4±2.1) \cdot 10^{-1}$	$(0.6±13) \cdot 10^{-2}$	$(0.5±1.4) \cdot 10^{-1}$	0.21	0.06	0.14
	Sigmoidal	$(-1.3±7.4) \cdot 10^{-1}$	$(-0.7±3.8) \cdot 10^{-1}$	$(0.7±3.8) \cdot 10^{-1}$	0.63	0.19	0.17
	Niemierko	$(-1.3±6.7) \cdot 10^{-1}$	$(-0.6±3.6) \cdot 10^{-1}$	$(0.8±3.6) \cdot 10^{-1}$	0.54	0.13	0.12
Okunieff	Poisson	$(-0.3±1.5) \cdot 10^{-3}$	$(0.5±9.4) \cdot 10^{-4}$	$(0.4±1.0) \cdot 10^{-3}$	0.21	0.06	0.14
	Sigmoidal	$(-0.4±3.0) \cdot 10^{-1}$	$(-0.3±1.8) \cdot 10^{-1}$	$(0.1±1.6) \cdot 10^{-1}$	0.77	0.38	0.38
	Niemierko	$(-0.4±2.2) \cdot 10^{-1}$	$(-0.2±1.1) \cdot 10^{-1}$	$(0.3±1.2) \cdot 10^{-1}$	0.53	0.13	0.12
Levegrun	Poisson	$(-0.8±4.1) \cdot 10^{-1}$	$(0.1±2.6) \cdot 10^{-1}$	$(1.0±2.8) \cdot 10^{-1}$	0.21	0.05	0.14
	Sigmoidal	$(-1.7±8.4) \cdot 10^{-1}$	$(-0.8±4.2) \cdot 10^{-1}$	$(1.0±4.5) \cdot 10^{-1}$	0.56	0.14	0.12
	Niemierko	$(-1.6±8.4) \cdot 10^{-1}$	$(-0.7±4.1) \cdot 10^{-1}$	$(1.0±4.5) \cdot 10^{-1}$	0.55	0.14	0.12
<i>Low-risk Prostate</i>							
Cheung	Poisson	$(-0.9±2.0) \cdot 10^{-2}$	$(-4.0±8.1) \cdot 10^{-3}$	$(0.4±1.7) \cdot 10^{-2}$	0.13	< 0.05	< 0.05
	Sigmoidal	$(-1.7±4.8) \cdot 10^{-1}$	$(-0.6±2.9) \cdot 10^{-1}$	$(1.5±3.8) \cdot 10^{-1}$	0.09	< 0.05	< 0.05
	Niemierko	$(-1.7±3.6) \cdot 10^{-1}$	$(-0.7±1.9) \cdot 10^{-1}$	$(1.0±3.0) \cdot 10^{-1}$	< 0.05	< 0.05	< 0.05
Okunieff	Poisson	$(-0.7±1.6) \cdot 10^{-3}$	$(-3.1±6.2) \cdot 10^{-4}$	$(0.3±1.3) \cdot 10^{-3}$	0.13	< 0.05	< 0.05
	Sigmoidal	$(-1.0±3.2) \cdot 10^{-1}$	$(-0.3±2.1) \cdot 10^{-1}$	$(1.0±2.5) \cdot 10^{-1}$	0.12	< 0.05	< 0.05
	Niemierko	$(-1.1±2.2) \cdot 10^{-1}$	$(-0.5±1.1) \cdot 10^{-1}$	$(0.6±1.8) \cdot 10^{-1}$	< 0.05	< 0.05	< 0.05
Levegrun	Poisson	$(-2.0±4.3) \cdot 10^{-1}$	$(-0.9±1.7) \cdot 10^{-1}$	$(0.9±3.7) \cdot 10^{-1}$	0.12	< 0.05	< 0.05
	Sigmoidal	$(-4.0±8.6) \cdot 10^{-1}$	$(-1.7±4.4) \cdot 10^{-1}$	$(2.4±7.0) \cdot 10^{-1}$	< 0.05	< 0.05	< 0.05
	Niemierko	$(-4.0±8.5) \cdot 10^{-1}$	$(-1.7±4.4) \cdot 10^{-1}$	$(2.3±7.0) \cdot 10^{-1}$	< 0.05	< 0.05	< 0.05
<i>Prostate Bed</i>							
King	Poisson	-0.3±1.6	$(-3.5±7.2) \cdot 10^{-1}$	$(0.3±8.0) \cdot 10^{-1}$	0.73	0.23	< 0.05
	Sigmoidal	-0.3±1.5	$(-4.3±8.0) \cdot 10^{-1}$	$(0.5±8.2) \cdot 10^{-1}$	0.65	0.09	< 0.05
	Niemierko	-0.3±1.5	$(-4.4±7.7) \cdot 10^{-1}$	$(0.4±8.1) \cdot 10^{-1}$	0.68	0.10	< 0.05
Okunieff	Poisson	$(-0.2±1.5) \cdot 10^{-2}$	$(-3.1±6.5) \cdot 10^{-3}$	$(0.3±7.3) \cdot 10^{-3}$	0.71	0.24	< 0.05
	Sigmoidal	$(-0.9±5.2) \cdot 10^{-1}$	$(-1.3±3.5) \cdot 10^{-1}$	$(0.4±3.3) \cdot 10^{-1}$	0.62	0.08	< 0.05
	Niemierko	$(-0.9±3.8) \cdot 10^{-1}$	$(-1.1±2.0) \cdot 10^{-1}$	$(0.1±2.0) \cdot 10^{-1}$	0.70	0.09	< 0.05
Levegrun	Poisson	-0.6±3.4	-0.8±1.5	0.1±1.7	0.75	0.21	< 0.05
	Sigmoidal	-0.5±2.0	-0.6±1.1	0.1±1.1	0.66	0.09	< 0.05
	Niemierko	-0.5±2.1	-0.6±1.1	0.1±1.1	0.69	0.09	< 0.05

M3D = Mobius3D; CC = COMPASS dose calculation; CR = COMPASS dose reconstruction.

than CR ones for left femoral head values and better than M3D results for bladder values in intermediate-risk treatments using the Niemierko model. M3D results were better than both CC and CR results in the same previous case using both LKB and Niemierko models. M3D results were better than CC ones for left femoral head values in high-risk (Niemierko model) and prostate bed treatments (LKB and Niemierko models). Finally, M3D results were better than CR results for rectum values in intermediate- and low-risk, as well as prostate bed treatments, for right femoral head values using both models and for left femoral head values using the Niemierko model in prostate bed treatments. There was no statistically significant difference for other NTCP comparisons.

TABLE 10. Calculations and comparisons of NTCP values for 100 analyzed prostate treatments, 25 from each group (high-, intermediate-, low-risk, and prostate bed patients), using the two models described in the text.

Volume	Model	Mean Differences (%)			p-values		
		M3D	CC	CR	M3D vs. CC	M3D vs. CR	CC vs. CR
<i>High-risk Prostate</i>							
Rectum	LKB	$(8.7 \pm 8.7) \cdot 10^{-1}$	$(0.3 \pm 2.6) \cdot 10^{-1}$	$(9.1 \pm 5.0) \cdot 10^{-1}$	< 0.05	0.68	< 0.05
	Niemierko	$(2.4 \pm 3.5) \cdot 10^{-1}$	$(0.5 \pm 8.9) \cdot 10^{-2}$	$(2.4 \pm 2.2) \cdot 10^{-1}$	< 0.05	0.97	< 0.05
Bladder	LKB	$(-3.7 \pm 5.2) \cdot 10^{-2}$	$(0.9 \pm 1.4) \cdot 10^{-2}$	$(-2.8 \pm 3.2) \cdot 10^{-2}$	< 0.05	0.49	< 0.05
	Niemierko	$(-3.4 \pm 4.2) \cdot 10^{-2}$	$(0.9 \pm 1.3) \cdot 10^{-2}$	$(-2.7 \pm 3.0) \cdot 10^{-2}$	< 0.05	0.55	< 0.05
LF head	LKB	$(0.6 \pm 2.0) \cdot 10^{-4}$	$(0.9 \pm 2.5) \cdot 10^{-4}$	$(3.7 \pm 8.2) \cdot 10^{-5}$	0.06	0.48	0.23
	Niemierko	$(0.6 \pm 1.7) \cdot 10^{-4}$	$(0.9 \pm 2.2) \cdot 10^{-4}$	$(4.1 \pm 7.6) \cdot 10^{-5}$	< 0.05	0.54	0.22
RF head	LKB	$(1.5 \pm 6.1) \cdot 10^{-4}$	$(1.3 \pm 4.0) \cdot 10^{-4}$	$(1.0 \pm 2.4) \cdot 10^{-4}$	0.65	0.55	0.45
	Niemierko	$(1.3 \pm 4.8) \cdot 10^{-4}$	$(1.2 \pm 3.0) \cdot 10^{-4}$	$(1.0 \pm 1.9) \cdot 10^{-4}$	0.76	0.64	0.50
<i>Intermediate-risk Prostate</i>							
Rectum	LKB	$(8.6 \pm 5.1) \cdot 10^{-1}$	$(0.9 \pm 2.0) \cdot 10^{-1}$	1.2±0.6	< 0.05	< 0.05	< 0.05
	Niemierko	$(2.2 \pm 1.9) \cdot 10^{-1}$	$(2.4 \pm 5.8) \cdot 10^{-2}$	$(3.1 \pm 2.3) \cdot 10^{-1}$	< 0.05	< 0.05	< 0.05
Bladder	LKB	$(-0.1 \pm 1.5) \cdot 10^{-3}$	$(0.4 \pm 1.5) \cdot 10^{-3}$	$(-1.2 \pm 3.2) \cdot 10^{-3}$	< 0.05	0.14	0.08
	Niemierko	$(-0.2 \pm 2.1) \cdot 10^{-3}$	$(0.7 \pm 1.9) \cdot 10^{-3}$	$(-1.9 \pm 4.1) \cdot 10^{-3}$	< 0.05	0.08	< 0.05
LF head	LKB	$(2.0 \pm 6.2) \cdot 10^{-6}$	$(2.6 \pm 7.2) \cdot 10^{-6}$	$(3.4 \pm 8.6) \cdot 10^{-6}$	< 0.05	< 0.05	0.05
	Niemierko	$(2.7 \pm 8.1) \cdot 10^{-6}$	$(3.6 \pm 9.4) \cdot 10^{-6}$	$(0.5 \pm 1.1) \cdot 10^{-5}$	< 0.05	< 0.05	< 0.05
RF head	LKB	$(2.9 \pm 6.3) \cdot 10^{-6}$	$(0.5 \pm 1.2) \cdot 10^{-5}$	$(0.8 \pm 2.3) \cdot 10^{-5}$	0.15	0.14	0.14
	Niemierko	$(3.7 \pm 7.8) \cdot 10^{-6}$	$(0.6 \pm 1.4) \cdot 10^{-5}$	$(1.0 \pm 2.7) \cdot 10^{-5}$	0.12	0.12	0.12
<i>Low-risk Prostate</i>							
Rectum	LKB	1.1±0.7	$(1.5 \pm 2.9) \cdot 10^{-1}$	1.5±0.5	< 0.05	< 0.05	< 0.05
	Niemierko	$(3.1 \pm 3.6) \cdot 10^{-1}$	$(0.5 \pm 1.2) \cdot 10^{-1}$	$(4.0 \pm 3.1) \cdot 10^{-1}$	< 0.05	< 0.05	< 0.05
Bladder	LKB	$(1.5 \pm 7.1) \cdot 10^{-3}$	$(1.6 \pm 5.7) \cdot 10^{-3}$	$(-2.0 \pm 5.5) \cdot 10^{-3}$	0.87	0.15	0.11
	Niemierko	$(1.7 \pm 8.2) \cdot 10^{-3}$	$(2.1 \pm 6.4) \cdot 10^{-3}$	$(-2.7 \pm 6.3) \cdot 10^{-3}$	0.41	0.12	0.06
LF head	LKB	$(0.3 \pm 1.1) \cdot 10^{-5}$	$(2.5 \pm 9.1) \cdot 10^{-6}$	$(0.3 \pm 1.0) \cdot 10^{-5}$	0.71	0.49	0.12
	Niemierko	$(0.3 \pm 1.3) \cdot 10^{-5}$	$(0.3 \pm 1.1) \cdot 10^{-5}$	$(0.4 \pm 1.2) \cdot 10^{-5}$	0.85	0.35	0.12
RF head	LKB	$(1.9 \pm 8.5) \cdot 10^{-6}$	$(2.0 \pm 6.9) \cdot 10^{-6}$	$(0.3 \pm 1.3) \cdot 10^{-5}$	0.88	0.09	0.23
	Niemierko	$(0.2 \pm 1.0) \cdot 10^{-5}$	$(2.6 \pm 8.5) \cdot 10^{-6}$	$(0.4 \pm 1.6) \cdot 10^{-5}$	0.70	0.07	0.21
<i>Prostate Bed</i>							
Rectum	LKB	1.1±0.6	$(1.1 \pm 3.8) \cdot 10^{-1}$	1.3±0.6	< 0.05	< 0.05	< 0.05
	Niemierko	$(2.9 \pm 1.8) \cdot 10^{-1}$	$(0.3 \pm 1.2) \cdot 10^{-1}$	$(3.6 \pm 2.1) \cdot 10^{-1}$	< 0.05	< 0.05	< 0.05
Bladder	LKB	$(-3.3 \pm 7.5) \cdot 10^{-3}$	$(1.8 \pm 5.1) \cdot 10^{-3}$	$(-0.4 \pm 1.1) \cdot 10^{-2}$	< 0.05	0.76	< 0.05
	Niemierko	$(-4.2 \pm 8.3) \cdot 10^{-3}$	$(2.2 \pm 5.5) \cdot 10^{-3}$	$(-0.5 \pm 1.2) \cdot 10^{-2}$	< 0.05	0.71	< 0.05
LF head	LKB	$(0.6 \pm 2.3) \cdot 10^{-5}$	$(1.0 \pm 2.7) \cdot 10^{-5}$	$(1.5 \pm 4.7) \cdot 10^{-5}$	< 0.05	0.08	0.22
	Niemierko	$(0.7 \pm 2.3) \cdot 10^{-5}$	$(1.1 \pm 2.7) \cdot 10^{-5}$	$(1.7 \pm 4.8) \cdot 10^{-5}$	< 0.05	< 0.05	0.19
RF head	LKB	$(1.6 \pm 3.1) \cdot 10^{-6}$	$(0.6 \pm 1.8) \cdot 10^{-5}$	$(5.0 \pm 7.2) \cdot 10^{-6}$	0.20	< 0.05	0.65
	Niemierko	$(2.3 \pm 4.3) \cdot 10^{-6}$	$(0.8 \pm 1.9) \cdot 10^{-5}$	$(7.0 \pm 9.2) \cdot 10^{-6}$	0.17	< 0.05	0.80

M3D = Mobius3D; CC = COMPASS dose calculation; CR = COMPASS dose reconstruction.

The correlation between the changes in DVH parameters and the corresponding changes in radiobiological outcomes is summarized in Tables 11 and 12 for targets and normal tissues, respectively. Selected DVH parameters for prostate volumes were  $D_{98}$ ,  $D_{2}$ , and  $D_{50}$ . Correlation ranged from moderate ( $r > 0.500$ ) to strong ( $r > 0.700$ ) for gEUD comparisons, although CC results seemed to have a weaker correlation ( $r < 0.500$ ). For OARs, selected DVH parameters were  $V_{70}$  and mean dose for rectum and bladder and maximum dose for femoral heads. For gEUD, moderate to strong correlation was observed, although some results exhibited a weaker correlation (CC and CR results for rectum and bladder). NTCP results seemed to have stronger correlation for the LKB model than the Niemierko model for rectum. Correlation was weaker for mean dose compared to  $V_{70}$  values for bladder. Low-risk cases exhibited a stronger correlation for bladder NTCP results compared with the other cases. Femoral heads showed no correlation for either LKB or Niemierko calculations.

TABLE 11. Pearson correlation coefficient comparing gEUD and TCP for prostate target volumes. The selected DVH metrics were  $D_{98}$ ,  $D_2$ , and  $D_{50}$ .

		<i>Cheung/King Parameters</i>			<i>Okunieff Parameters</i>			<i>Levegrun Parameters</i>			
		<i>TCP</i>	<i>TCP</i>	<i>TCP</i>	<i>TCP</i>	<i>TCP</i>	<i>TCP</i>	<i>TCP</i>	<i>TCP</i>	<i>TCP</i>	
		<i>Poiss</i>	<i>Sigm</i>	<i>Niem</i>	<i>Poiss</i>	<i>Sigm</i>	<i>Niem</i>	<i>Poiss</i>	<i>Sigm</i>	<i>Niem</i>	
		<i>gEUD</i>									
<i>High-risk Prostate</i>											
$D_{98}$	M3D	0.582	0.599	0.527	0.581	0.593	0.516	0.581	0.595	0.579	0.579
	CC	0.335	0.396	0.254	0.336	0.393	0.242	0.337	0.394	0.338	0.340
	CR	0.339	0.469	0.242	0.344	0.471	0.227	0.346	0.470	0.353	0.356
$D_2$	M3D	0.847	0.721	0.799	0.841	0.704	0.786	0.838	0.709	0.823	0.818
	CC	0.368	0.063	0.348	0.358	0.046	0.337	0.354	0.051	0.340	0.327
	CR	0.752	0.630	0.650	0.747	0.614	0.627	0.745	0.619	0.732	0.727
$D_{50}$	M3D	0.917	0.889	0.863	0.916	0.879	0.852	0.916	0.882	0.913	0.912
	CC	0.783	0.886	0.645	0.786	0.880	0.624	0.787	0.882	0.788	0.793
	CR	0.883	0.835	0.736	0.880	0.817	0.708	0.880	0.823	0.869	0.868
<i>Intermediate-risk Prostate</i>											
$D_{98}$	M3D	0.593	0.596	0.574	0.590	0.596	0.529	0.591	0.597	0.589	0.589
	CC	0.476	0.664	0.389	0.486	0.662	0.211	0.481	0.667	0.480	0.489
	CR	0.246	0.463	0.220	0.276	0.464	0.074	0.260	0.461	0.280	0.287
$D_2$	M3D	0.749	0.752	0.694	0.734	0.750	0.593	0.741	0.753	0.731	0.729
	CC	0.409	0.370	0.352	0.406	0.367	0.230	0.408	0.373	0.405	0.404
	CR	0.662	0.651	0.582	0.655	0.648	0.364	0.659	0.655	0.653	0.652
$D_{50}$	M3D	0.846	0.877	0.811	0.843	0.877	0.724	0.845	0.878	0.842	0.842
	CC	0.641	0.863	0.536	0.656	0.860	0.308	0.648	0.866	0.652	0.661
	CR	0.764	0.934	0.671	0.776	0.932	0.382	0.770	0.935	0.775	0.780
<i>Low-risk Prostate</i>											
$D_{98}$	M3D	0.559	0.585	0.466	0.555	0.585	0.435	0.556	0.590	0.547	0.549
	CC	0.415	0.541	0.337	0.415	0.541	0.315	0.416	0.541	0.409	0.414
	CR	0.549	0.646	0.464	0.556	0.646	0.425	0.554	0.646	0.558	0.562
$D_2$	M3D	0.701	0.691	0.626	0.700	0.691	0.596	0.700	0.694	0.698	0.697
	CC	0.402	0.539	0.281	0.397	0.539	0.249	0.398	0.539	0.389	0.392
	CR	0.700	0.758	0.590	0.703	0.758	0.540	0.702	0.758	0.703	0.705
$D_{50}$	M3D	0.853	0.858	0.756	0.851	0.857	0.719	0.852	0.862	0.846	0.846
	CC	0.530	0.882	0.320	0.531	0.882	0.262	0.531	0.882	0.525	0.532
	CR	0.856	0.956	0.711	0.860	0.955	0.648	0.859	0.957	0.861	0.864
<i>Prostate Bed</i>											
$D_{98}$	M3D	0.848	0.885	0.833	0.850	0.883	0.740	0.849	0.888	0.840	0.850
	CC	0.484	0.563	0.425	0.472	0.556	0.260	0.479	0.573	0.447	0.475
	CR	0.535	0.614	0.525	0.543	0.616	0.441	0.539	0.610	0.531	0.541
$D_2$	M3D	0.660	0.655	0.634	0.652	0.650	0.533	0.657	0.663	0.644	0.654
	CC	0.382	0.465	0.340	0.380	0.460	0.188	0.382	0.472	0.358	0.381
	CR	0.591	0.706	0.544	0.595	0.701	0.316	0.593	0.712	0.567	0.594
$D_{50}$	M3D	0.926	0.944	0.907	0.925	0.941	0.806	0.926	0.949	0.915	0.926
	CC	0.715	0.876	0.654	0.718	0.870	0.424	0.717	0.882	0.682	0.718
	CR	0.796	0.900	0.751	0.800	0.895	0.526	0.798	0.904	0.773	0.799

M3D = Mobius3D; CC = COMPASS dose calculation; CR = COMPASS dose reconstruction; Poiss = Poisson; Sigm = sigmoidal; Niem = Niemierko.

TABLE 12. Pearson correlation coefficient comparing gEUD and NTCP for OARs. The selected DVH metrics were  $V_{70}$  and mean dose ( $D_M$ ) for rectum and bladder, and maximum dose ( $D_{max}$ ) for femoral heads.

	High-risk Prostate						Intermediate-risk Prostate						Low-risk Prostate						Prostate Bed					
	gEUD		NTCP		Niemi		gEUD		NTCP		Niemi		gEUD		NTCP		Niemi		gEUD		NTCP		Niemi	
	LKB	LKB	NTCP	NTCP	Niemi	Niemi	LKB	LKB	NTCP	NTCP	Niemi	Niemi	LKB	LKB	NTCP	NTCP	Niemi	Niemi	LKB	LKB	NTCP	NTCP	Niemi	Niemi
$V_{70}$	M3D	0.820	0.944	0.847	0.847	0.847	0.603	0.893	0.795	0.795	0.795	0.502	0.889	0.889	0.889	0.766	0.766	0.766	0.839	0.938	0.938	0.804	0.804	0.804
	CC	0.769	0.788	0.741	0.741	0.741	0.655	0.733	0.720	0.720	0.720	0.655	0.489	0.489	0.489	0.470	0.470	0.470	0.826	0.820	0.820	0.781	0.781	0.781
	CR	0.397	0.915	0.796	0.796	0.796	0.165	0.893	0.866	0.866	0.866	0.017	0.881	0.881	0.881	0.784	0.784	0.784	0.680	0.906	0.906	0.823	0.823	0.823
	M3D	0.881	0.859	0.729	0.729	0.729	0.777	0.564	0.351	0.351	0.351	0.506	0.775	0.775	0.775	0.608	0.608	0.608	0.743	0.736	0.736	0.600	0.600	0.600
	CC	0.729	0.603	0.514	0.514	0.514	0.765	0.653	0.504	0.504	0.504	0.778	0.575	0.575	0.575	0.344	0.344	0.344	0.578	0.493	0.493	0.438	0.438	0.438
	CR	0.706	0.666	0.470	0.470	0.470	0.672	0.492	0.335	0.335	0.335	0.707	0.269	0.269	0.269	0.132	0.132	0.132	0.620	0.189	0.189	0.002	0.002	0.002
$V_{70}$	M3D	0.777	0.639	0.662	0.662	0.662	0.643	0.206	0.266	0.266	0.266	0.748	0.815	0.815	0.815	0.816	0.816	0.816	0.594	0.616	0.616	0.594	0.594	0.594
	CC	0.163	0.142	0.162	0.162	0.162	0.114	0.160	0.189	0.189	0.189	0.196	0.578	0.578	0.578	0.607	0.607	0.607	0.445	0.097	0.097	0.445	0.445	0.445
	CR	0.735	0.641	0.668	0.668	0.668	0.505	0.379	0.48	0.48	0.48	0.587	0.518	0.518	0.518	0.581	0.581	0.581	0.719	0.542	0.542	0.719	0.719	0.719
	M3D	0.853	0.462	0.497	0.497	0.497	0.522	0.130	0.143	0.143	0.143	0.820	0.452	0.452	0.455	0.455	0.455	0.455	0.664	-0.16	-0.16	-0.153	-0.153	-0.153
	CC	0.449	0.158	0.210	0.210	0.210	0.424	-0.306	-0.335	-0.335	-0.335	0.416	-0.342	-0.342	-0.374	-0.374	-0.374	-0.374	0.397	-0.098	-0.098	-0.072	-0.072	-0.072
	CR	0.825	0.557	0.594	0.594	0.594	0.748	-0.027	0.056	0.056	0.056	0.751	0.298	0.298	0.298	0.336	0.336	0.336	0.723	0.389	0.389	0.418	0.418	0.418
$D_{max}$	M3D	0.847	0.104	0.115	0.115	0.115	0.418	0.351	0.350	0.350	0.693	0.190	0.190	0.190	0.190	0.190	0.190	0.689	-0.171	-0.171	-0.155	-0.155	-0.155	
	CC	0.675	0.022	0.030	0.030	0.030	0.421	0.483	0.478	0.478	0.432	-0.085	-0.085	-0.085	-0.084	-0.084	-0.084	0.704	-0.184	-0.184	-0.167	-0.167	-0.167	
	CR	0.851	-0.338	-0.288	-0.288	-0.288	0.319	0.285	0.280	0.280	0.844	-0.122	-0.122	-0.122	-0.120	-0.120	-0.120	0.637	-0.309	-0.309	-0.311	-0.311	-0.311	
	M3D	0.901	0.067	0.091	0.091	0.091	0.552	0.139	0.190	0.190	0.496	0.077	0.077	0.077	0.078	0.078	0.078	0.745	0.183	0.183	0.203	0.203	0.203	
	CC	0.705	0.066	0.072	0.072	0.072	0.589	-0.048	-0.017	-0.017	0.376	-0.198	-0.198	-0.198	-0.181	-0.181	-0.181	0.524	0.057	0.057	0.067	0.067	0.067	
	CR	0.815	0.049	0.070	0.070	0.070	0.531	-0.160	-0.143	-0.143	0.507	-0.174	-0.174	-0.174	-0.176	-0.176	-0.176	0.723	-0.313	-0.313	-0.287	-0.287	-0.287	

M3D = Mobius3D; CC = COMPASS dose calculation; CR = COMPASS dose reconstruction; Niemi = Niemierko.

## IV. DISCUSSION

### A. Remarks and limitations of the present study

#### **A.1 Variability between radiobiological response parameters and models**

Radiobiological models are powerful evaluation tools because values such as TCP and NTCP are related to the clinical outcome of a treatment. Therefore, biological-based evaluation becomes an interesting metric in order to evaluate a treatment plan. However, all biological models have uncertainties in the values of the parameters chosen. These metrics should be used with caution, due to these uncertainties. The use of biological models in plan evaluation requires accurate TCP/NTCP models and parameter estimation. The users of biological metrics could derive model parameters based on their own experience by calibrating selected models against observed clinical outcomes. Another option is to cautiously use published parameter values, as these data are available for many tumor and normal tissue sites. These known ideas have been expressed in reference reports, such as AAPM TG-166 or ICRU 83 reports.<sup>(20,21)</sup> The first option could not be feasible for our institution, as it requires expertise in outcome modeling and sufficient patient throughput. In order to overcome these possible uncertainties, published parameters taken from reference reports and studies were used, as in the TG-166 report,<sup>(21)</sup> the study by Okunieff et al.,<sup>(45)</sup> or the study by Burman et al.<sup>(58)</sup> The possible variation due to the model selection was overcome by using three models for TCP calculations (Poisson, sigmoidal, and Niemierko) and two models for NTCP calculations (LKB and Niemierko). In both cases, results between the models were comparable.

A limitation of the present study was related with the TCP calculation using the Eq. (3), where the dose protraction factor has not been considered in order to model the TCP. Dose protraction factor ( $G$ ) modifies the quadratic term of the linear-quadratic expression in order to take into account the sublethal damage repair of protracting the dose delivery. If the delivery takes a short time (instantaneous),  $G = 1$ . For any other dose delivery pattern,  $G < 1$ . This study was performed considering  $G \sim 1$  since treatment times are about few minutes, as in the prostate treatments using VMAT techniques.

#### **A.2 Use of AAPM TG-166 test cases**

The aim of the present study was to evaluate the implementation of radiobiological metrics in the patient-specific QA workflow. TG-166 test cases have been designed to assess the biological modeling implemented in several commercial TPSs. In particular, benchmark phantom structures have been specifically developed to address this problem. In addition, the dose calculation and reconstruction capabilities of the systems should be tested with clinical cases. The other TG-166 cases (H&N, prostate, and brain) have been intentionally used for this purpose.

#### **A.3 Relevance of radiobiological metrics**

The main aim of patient-specific QA for modulated treatments is to ensure the quality of each individual patient treatment. The pretreatment QA measurement-based process must be considered to ensure the correct information flow from TPS plan calculation to treatment delivery in the linac by means of the record and verify (R&V) system. Such patient-specific QA is conventionally performed by delivering the patient plans to a phantom with detectors, and comparing the calculated and measured dose in the phantom. Recent studies have investigated estimating the delivered patient dose from QA measurement, resulting in several new QA tools. These new approaches to patient-specific QA open the possibility of adopting patient dose-based metrics that are more relevant to the expected treatment outcome. Therefore, new QA metrics should be introduced in order to effectively take into account the clinical impact of possible calculation/delivery mistakes on the treatment outcome for modulated treatments. Considering radiobiological parameters as potential indicators of this clinical outcome for radiation therapy

treatments, such parameters could be included in these metrics. The previous statement goes beyond the usual patient-specific QA flow, which is based in dose distribution (2D or 3D) comparisons. In this way, patient-specific QA metrics could directly evaluate the variation of expected clinical outcome between the planned dose and delivered dose.

In addition to the previous discussion, radiobiological metrics could be used to evaluate the robustness of a radiotherapy treatment from the point of view of its sensitivity to possible perturbations. From the sigmoid shape of TCP and NTCP curves, the high region for TCP and the low region for NTCP are the ideal regions for these indices, because results are less sensitive to possible changes. Determining the robustness of a plan could be used to prevent the influence of different possible sources of error and could lead to improvements in the quality of the plan before treatment.

#### **A.4 Limitation in the definition of action levels**

Any QA metric should have sensitivity in order to show the impact of possible errors on the patient treatment. Action levels for classical metrics, as gamma analysis, have been widely studied. However, gamma passing rates could not catch clinically relevant patient dose errors.<sup>(16,17)</sup> TCP and NTCP models have been introduced in order to take into account the treatment outcome. In this way, these metrics permit to concentrate on the errors that are of clinical importance. However, the impact of an error source on the clinical outcome depends on the location/characteristics of the PTV and the possible OARs. As an example, the traditional action levels for gamma analysis (3% of dose difference) could be applied to evaluate the observed difference in a serial normal tissue, as the spinal cord. If the maximum dose to spinal cord is, for example, 5 Gy, an error of 3%, 5% or even 10% probably will have no impact on the NTCP for this tissue. When the NTCP slope starts to be steep, the possible impact on NTCP becomes increasingly important. In order to implement new patient-specific QA protocols, action levels for new metrics should be defined. Nevertheless, the definition of action levels might be beyond the scope of the present study. Future application of these metrics to a large amount of modulated treatments/disease sites/PTVs/OARs could give us enough information and statistics in order to define a correct level for the acceptability/rejection of gEUD/TCP/NTCP variations, depending on the case.

#### **B. Verifications of biological metrics for both systems with TG-166**

Discrepancies between TPS and the two systems were in good agreement for gEUD calculations. Maximum discrepancies were found for small volume structures, like inner ear, where small differences could lead to poor results. For TCP and NTCP comparisons, discrepancies were better than the previous results. The worst results were found in the benchmark phantom test case.

### C. Evaluation of prostate treatments with classical (gamma) and DVH-based dose-volume metrics.

COMPASS-calculated and reconstructed passing rates were slightly better than those from Mobius3D. Results were above the TG-119 action level of 88% for composite dose gamma analysis,<sup>(62)</sup> excluding some results with the 2%/2 mm gamma criterion: COMPASS calculated mean values for high-risk pelvic lymph node PTV, COMPASS reconstructed mean values for high-risk prostate PTV, and some Mobius3D results (high-risk prostate and pelvic lymph node PTVs, prostate bed PTV, rectum for the bed case, and bladder for the high- and low-risk and bed cases).

DVH comparisons were comparable to results from other studies.<sup>(63-65)</sup> Both systems led to similar results for dose-volume parameters. CC calculations improved the M3D values, and Mobius3D results were better than CR values.

### D. Biological metrics applied prostate treatments

Radiobiological calculations led to comparable discrepancies for the three systems. The statistically significant differences were shown in the previous section. For gEUD values, differences were larger for OARs than for target volumes. Considering the TCP calculations, absolute TCP values calculated for the TPS with the Poisson model were, in general, slightly larger than those from the sigmoidal and Niemierko models. These absolute values were comparable between the models, with the exception of the results for the high-risk prostate case taking the values from Cheung et al.<sup>(54)</sup> (sigmoidal and Niemierko values were very close, but they were about 20% lower than Poisson results), and the prostate bed case taking the values from King et al.<sup>(56)</sup> (sigmoidal and Niemierko values were also very close but about 15% lower than Poisson results) and Levegrun et al.<sup>(57)</sup> (Poisson results were about 10%–15% higher than those from sigmoidal and Niemierko values.) The observed discrepancies in absolute TCP values come from the model/parameter selection and they were also derived for M3D, CC, and CR results, preserving the low differences obtained in the comparisons. For NTCP calculations, absolute values were comparable between both the LKB and Niemierko models and also preserved the low discrepancies in the comparisons.

Correlation results between DVH and gEUD/TCP/NTCP differences were analyzed in the previous section. Correlation between  $D_{50}$  and TCP was stronger than results obtained comparing  $D_{98}$  and  $D_2$ . TCP increase/decrease was highly correlated with  $D_{50}$ , which is related to the mean dose. Correlation between NTCP and dose-volume parameters for OARs was larger for rectum than bladder or femoral head values. These results could be explained using ideas from the study by Zhen et al.,<sup>(66)</sup> based on the discussion expressed in Discussion section A.3 above. If TCP/NTCP values are located in low-gradient regions of the sigmoid curve (high TCP or low NTCP), the clinical outcome is less sensitive to changes in the plan, like dose-volume discrepancies. NTCP values for bladder and femoral heads are low and extremely low, respectively (Fig. 4). This could explain the lack of correlation between changes in DVH parameters and NTCP results in these cases.

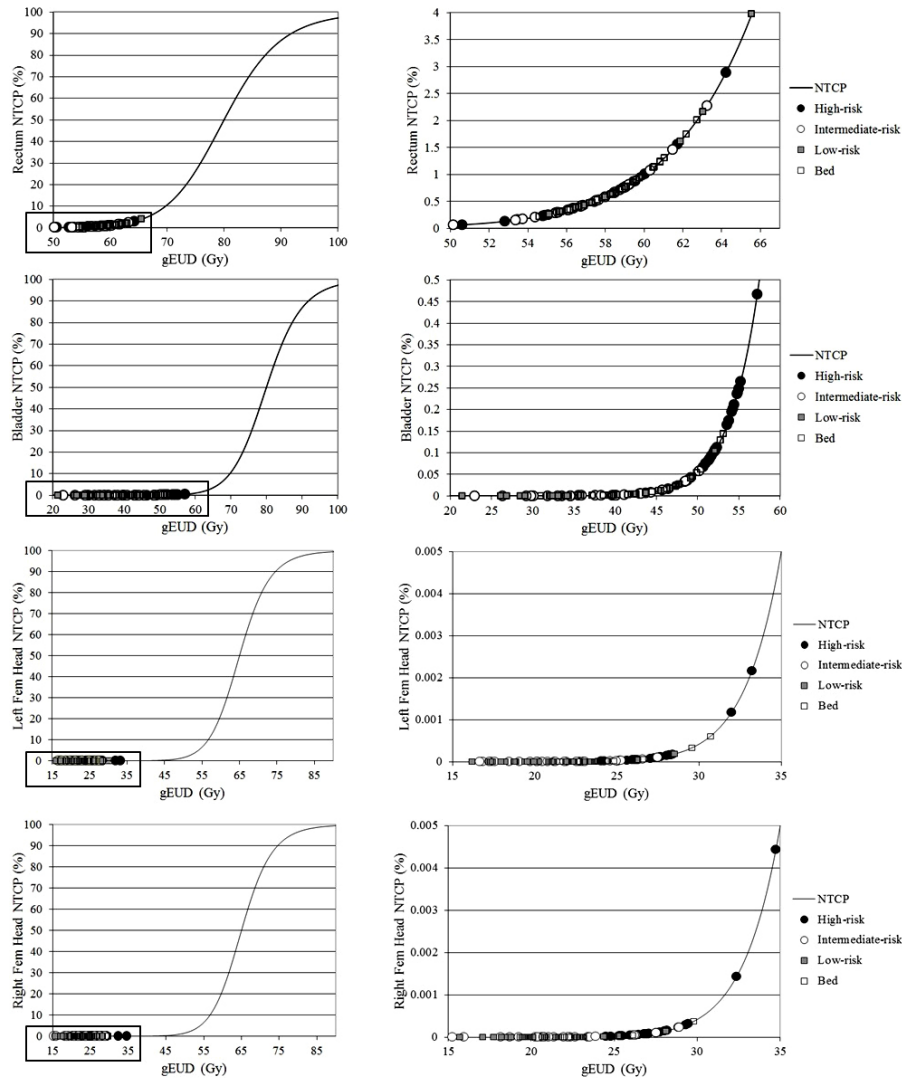


FIG. 4. NTCP values for rectum, bladder, and femoral heads calculated by the TPS for the high-, intermediate-, and low-risk and prostate bed cases. NTCP values are located in low-gradient regions of the sigmoid curve; therefore, the clinical outcome is less sensitive to changes in DVH parameters.

## V. CONCLUSIONS

This study has explored the potential usefulness of biological metrics in patient-specific QA process. Initial evaluation of radiobiological data extracted from dose calculation and reconstruction performed by Mobius3D and COMPASS systems were carried out by means of TG-166 test cases. The capabilities of the systems for 3D dose calculations and reconstructions were assessed with classical metrics, obtaining comparable results between both systems. Radiobiological metrics expressed in terms of comparisons between different indices (gEUD, TCP, and NTCP) were applied to a paradigmatic case of VMAT delivery (prostate treatment). The possibility of using radiobiological calculations as alternative metrics was introduced in order to evaluate the expected clinical outcome of radiotherapy treatments in the scope of pretreatment patient-specific QA.



## COPYRIGHT

This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/).

## REFERENCES

1. Peeters ST, Heemsbergen WD, Koper PC, et al. Dose-response in radiotherapy for localized prostate cancer: results of the Dutch multicenter randomized phase III trial comparing 68 Gy of radiotherapy with 78 Gy. *J Clin Oncol*. 2006;24(13):1990–96.
2. Zelefsky MJ, Fuks Z, Happersett L, et al. Clinical experience with intensity modulated radiation therapy (IMRT) in prostate cancer. *Radiother Oncol*. 2000;55(3):241–49.
3. Brenner DJ and Hall EJ. Fractionation and protraction for radiotherapy of prostate carcinoma. *Int J Radiat Oncol Biol Phys*. 1999;43(5):1095–101.
4. Otto K. Volumetric modulated arc therapy: IMRT in a single gantry arc. *Med Phys*. 2008;35(1):310–17.
5. Bzdusek K, Friberger H, Eriksson K, Hårdemark B, Robinson D, Kaus M. Development and evaluation of an efficient approach to volumetric arc therapy planning. *Med Phys*. 2009;36(6):2328–39.
6. Teoh M, Cark CH, Wood K, Whitaker S, Nisbet A. Volumetric modulated arc therapy: a review of current literature and clinical use in practice. *Br J Radiol*. 2011;84(1007):967–96.
7. Davidson MT, Blake SJ, Batchelar DL, Cheung P, Mah K. Assessing the role of volumetric modulated arc therapy (VMAT) relative to IMRT and helical tomotherapy in the management of localized, locally advanced, and post-operative prostate cancer. *Int J Radiat Oncol Biol Phys*. 2011;80(5):1550–58.
8. Fogarty GB, Ng D, Liu G, Haydu LE, Bhandari N. Volumetric modulated arc therapy is superior to conventional intensity modulated radiotherapy – a comparison among prostate cancer patients treated in an Australian centre. *Radiat Oncol*. 2011;6:108.
9. Clemente F, Perez-Vara C, Couselo M, et al. Hypofractionated prostate treatment in a single VMAT arc [abstract]. *Med Phys*. 2013;40(6):351.
10. Alber M, Broggi S, De Wagter C, et al. Guidelines for the verification of IMRT. Mijnheer G, editor. *ESTRO Physics Booklet 9*. Brussels, Belgium: ESTRO; 2008.
11. Ezzell G, Galvin J, Low D, et al. Guidance document on delivery, treatment planning, and clinical implementation of IMRT: report of the IMRT Subcommittee of the AAPM Radiation Therapy Committee. *Med Phys*. 2003;30(8):2089–115.
12. Low D, Harms W, Mutic S, Purdy J. A technique for the quantitative evaluation of dose distributions. *Med Phys*. 1998;25(5):656–61.
13. Nelms B and Simon J. A survey on planar IMRT QA analysis. *J Appl Clin Med Phys*. 2007;8(3):76–90.
14. Bedford J, Lee Y, Wai P, South C, Warrington A. Evaluation of the Delta4 phantom for IMRT and VMAT verification. *Phys Med Biol*. 2009;54(9):N167–76.
15. Létourneau D, Publicover J, Kozelka J, Moseley D, Jaffray D. Novel dosimetric phantom for quality assurance of volumetric modulated arc therapy. *Med Phys*. 2009;36(5):1813–21.
16. Nelms B, Zhen H, Tomé W. Per-beam, planar IMRT QA passing rates do not predict clinically relevant patient dose errors. *Med Phys*. 2011;38(2):1037–44.
17. Zhen H, Nelms B, Tomé W. Moving from gamma passing rates to patient DVH-based QA metrics in pre-treatment dose QA. *Med Phys*. 2011;38(10):5477–89.
18. Boggula R, Lorenz F, Mueller L, et al. Experimental validation of a commercial 3D dose verification system for intensity-modulated arc therapies. *Phys Med Biol*. 2010;55(19):5619–33.
19. Ling CC and Li XA. Over the next decade the success of radiation treatment planning will be judged by the immediate biological response of tumor cells rather than by surrogate measures such as dose maximization and uniformity. *Med Phys*. 2005;32(7):2189–92.
20. ICRU. Prescribing, recording, and reporting intensity-modulated photon-beam therapy (IMRT). ICRU Report 83. Bethesda, MD: ICRU; 2010.
21. Allen Li X, Alber M, Deasy JO, et al. The use and QA of biologically related models for treatment planning: short report of the TG-166 of the therapy physics committee of the AAPM. *Med Phys*. 2012;39(3):1386–409.
22. Njeh CF, Parker BC, Orton CG. Point/Counterpoint: Evaluation of treatment plans using target and normal tissue DVHs is no longer appropriate. *Med Phys*. 2015;42(5):2099–102.
23. Deasy JO, Mayo CS, Orton CG. Treatment planning evaluation and optimization should be biologically and not dose/volume based. *Med Phys*. 2015;42(6):2753–56.
24. Niemierko A. Reporting and analyzing dose distributions: a concept of equivalent uniform dose. *Med Phys*. 1997;24(1):103–10.
25. Niemierko A. A generalized concept of equivalent uniform dose (EUD). *Med Phys*. 1999;26(6):1100.
26. Munro TR and Gilbert CW. The relation between tumour lethal doses and the radiosensitivity of tumour cells. *Br J Radiol*. 1961;34:246–51.
27. Dale RG. The application of the linear-quadratic dose-effect equation to fractionated and protracted radiotherapy. *Br J Radiol*. 1985;58(690):518–28.
28. Niemierko A and Goitein M. Implementation of a model for estimating tumor control probability for an inhomogeneously irradiated tumor. *Radiother Oncol*. 1993;29(2):140–47.

29. Lyman JT. Complication probability as assessed from dose-volume histograms. *Radiat Res Suppl.* 1985;8:S13–S19.
30. Lyman JT and Wolbarst AB. Optimization of radiation therapy, IV: a dose–volume histogram reduction algorithm. *Int J Radiat Oncol Biol Phys.* 1989;17(2):433–36.
31. Kutcher GJ and Burman C. Calculation of complication probability factors for non-uniform normal tissue irradiation: the effective volume method. *Int J Radiat Oncol Biol Phys.* 1989;16(6):1623–30.
32. Kutcher GJ, Burman C, Brewster L, Goitein M, Mohan R. Histogram reduction method for calculating complication probabilities for three-dimensional treatment planning evaluations. *Int J Radiat Oncol Biol Phys.* 1991;21(1):137–46.
33. Niemierko A and Goitein M. Calculation of normal tissue complication probability and dose-volume histogram reduction schemes for tissues with a critical element architecture. *Radiother Oncol.* 1991;20(3):166–76.
34. Niemierko A and Goitein M. Modeling of normal tissue response to radiation: the critical volume model. *Int J Radiat Oncol Biol Phys.* 1993;25(1):135–45.
35. Moiseenko V, Battista J, Van Dyk J. Normal tissue complication probabilities: dependence on choice of biological model and dose–volume histogram reduction scheme. *Int J Radiat Oncol Biol Phys.* 2000;46(4):983–95.
36. Liu H, Mackie T, McCullough E. Correction kernel tilting and hardening in convolution/superposition dose calculation for clinical divergent and polychromatic photon beams. *Med Phys.* 1997;24(11):1729–41.
37. Lu W, Olivera G, Chen M, Reckwerdt P, Mackie T. Accurate convolution/superposition for multi-resolution dose calculation using cumulative tabulated kernels. *Phys Med Biol.* 2005;50(4):655–80.
38. Jacques R, Wong J, Taylor R, McNutt T. Real-time dose computation: GPU-accelerated source modeling and superposition/convolution. *Med Phys.* 2011;38(1):294–305.
39. Chen Q, Chen M, Lu W. Ultrafast convolution/superposition using tabulated and exponential kernels on GPU. *Med Phys.* 2011;38(3):1150–61.
40. Huang J, Childress N, Kry S. Calculation of high resolution and material-specific photon energy deposition kernels [abstract]. *Med Phys.* 2012;39(6):3822.
41. Huang J, Eklund D, Childress N, et al. Investigation of various energy deposition kernel refinements for the convolution/superposition method. *Med Phys.* 2013;40(12):121721.
42. Boggula R, Lorenz F, Mueller L, et al. Experimental validation of a commercial 3D dose verification system for intensity-modulated arc therapies. *Phys Med Biol.* 2010;55(19):5619–33.
43. Stavrev P, Stavreva N, Niemierko A, Goitein M. Generalization of a model of tissue response to radiation based on the idea of functional subunits and binomial statistics. *Phys Med Biol.* 2001;46(5):1501–18.
44. Stavreva N, Stavrev P, Warkentin B, Fallone BG. Derivation of the expression for gamma50 and D50 for different individual TCP and NTCP models. *Phys Med Biol.* 2002;47(20):3591–604.
45. Okunieff P, Morgan D, Niemierko A, Suit HD. Radiation dose-response of human tumors. *Int J Radiat Oncol Biol Phys.* 1995;32(4):1227–37.
46. Warkentin B, Stavrev P, Stavreva N, Field C, Fallone BJ. A TCP-NTCP estimation module using DVHs and known radiobiological models and parameter sets. *J Appl Clin Med Phys.* 2004;5(1):50–63.
47. Emami B, Lyman J, Brown A, et al. Tolerance of normal tissue to therapeutic irradiation. *Int J Radiat Oncol Biol Phys.* 1991;21(1):109–22.
48. Marks L, Yorke E, Jackson A, et al. Use of normal tissue complication probability models in the clinic. *Int J Radiat Oncol Biol Phys.* 2010;76(3 Suppl):S10–S19.
49. Deacon J, Peckham MJ, Steel GG. The radioresponsiveness of human tumors and the initial slope of the cell survival curve. *Radiother Oncol.* 1984;2(4):317–23.
50. Brenner DJ and Hall EJ. Fractionation and protraction for radiotherapy of prostate carcinoma. *Int J Radiat Oncol Biol Phys.* 1999;43(5):1095–101.
51. Fowler J, Chappell R, Ritter M. Is alpha/beta for prostate tumors really low? *Int J Radiat Oncol Biol Phys.* 2001;50(4):1021–31.
52. King CR and Fowler JF. A simple analytic derivation suggest that prostate cancer alpha/beta ratio is low. *Int J Radiat Oncol Biol Phys.* 2001;51(1):213–14.
53. Brenner DJ, Martinez AA, Edmunson GK, Mitchell C, Thames HD, Armour EP. Direct evidence that prostate tumors show high sensitivity to fractionation (low alpha/beta ratio), similar to late-responding normal tissue. *Int J Radiat Oncol Biol Phys.* 2002;52(1):6–13.
54. Cheung R, Tucker SL, Dong L, Kuban D. Dose-response for biochemical control among high-risk prostate cancer patients after external beam radiotherapy. *Int J Radiat Oncol Biol Phys.* 2003;56(5):1234–40.
55. Cheung R, Tucker SL, Lee AK, et al. Dose-response characteristics of low- and intermediate-risk prostate cancer treated with external beam radiotherapy. *Int J Radiat Oncol Biol Phys.* 2005;61(4):993–1002.
56. King CR and Kapp DS. Radiotherapy after prostatectomy: is the evidence for dose escalation out there? *Int J Radiat Oncol Biol Phys.* 2008;71(2):346–50.
57. Levegrün S, Jackson A, Zelefsky MJ, et al. Risk group dependence of dose-response for biopsy outcome after three-dimensional conformal radiation therapy of prostate cancer. *Radiother Oncol.* 2002;63(1):11–26.
58. Burman C, Kutcher GJ, Emami B, Goitein M. Fitting of normal tissue tolerance data to an analytic function. *Int J Radiat Oncol Biol Phys.* 1991;21(1):123–35.
59. Stavrev P, Niemierko A, Stavreva N, Goitein M. The application of biological models to clinical data. *Phys Med.* 2001;17(2):1–12.
60. Huang E, Teh BS, Strother DR, et al. Intensity-modulated radiation therapy for pediatric medulloblastoma: early report on the reduction of ototoxicity. *Int J Radiat Oncol Biol Phys.* 2002;52(3):599–605.
61. Lee TF, Chao PJ, Ting HM, et al. Comparative analysis of SmartArc-based dual arc volumetric-modulated arc radiotherapy (VMAT) versus intensity-modulated radiotherapy (IMRT) for nasopharyngeal carcinoma. *J Appl Clin Med Phys.* 2011;12(4):3587.

62. Ezzell G, Burmeister J, Dogan N, et al. IMRT commissioning: multiple institution planning and dosimetry comparisons, a report from AAPM Task Group 119. *Med Phys.* 2009;36(11):5359–73.
63. Nelms B, Chan M, Jarry G, et al. Evaluating IMRT and VMAT dose accuracy: practical examples of failure to detect systematic errors when applying a commonly used metric and action levels. *Med Phys.* 2013;40(11):111722.
64. Boggula R, Jahnke L, Wertz H, Lohr F, Wenz F. Patient-specific 3D pretreatment and potential 3D online dose verification of Monte Carlo-calculated IMRT prostate treatment plans. *Int J Radiat Oncol Biol Phys.* 2011;81(4):1168–75.
65. Clemente-Gutiérrez F and Pérez-Vara C. Dosimetric validation and clinical implementation of two 3D dose verification systems for quality assurance in volumetric-modulated arc therapy techniques. *J Appl Clin Med Phys.* 2015;16(2):5190.
66. Zhen H, Nelms BE, Tomé WA. On the use of biomathematical models in patient-specific IMRT dose QA. *Med Phys.* 2013;40(7):071702.