Physica Medica 67 (2019) 132–140

Contents lists available at ScienceDirect

Physica Medica

journal homepage: www.elsevier.com/locate/ejmp

Original paper

Inter-planner variation in treatment-plan quality of plans created with a knowledge-based treatment planning system

Kazuki Kubo[a], Hajime Monzen[b,*], Kentaro Ishii[a], Mikoto Tamura[b], Yuta Nakasaka[a], Masayuki Kusawake[a], Shun Kishimoto[a], Ryuta Nakahara[a], Shogo Matsuda[a], Toshifumi Nakajima[a], Ryu Kawamorita[a]

[a] Department of Radiation Oncology, Tane General Hospital, 1-12-21 Kujo-minami, Nishi, Osaka 550-0025, Japan
[b] Department of Medical Physics, Graduate School of Medical Sciences, Kindai University, 377-2 Ohno-higashi, Osaka-sayama, Osaka 589-8511, Japan

ARTICLE INFO

Keywords:
Knowledge-based treatment planning system
Inter-planner variation
VMAT
Prostate

ABSTRACT

Purpose: This study aimed to clarify the inter-planner variation of plan quality in knowledge-based plans created by nine planners.
Methods: Five hypofractionated prostate-only (HPO) volumetric modulated arc therapy (VMAT) plans and five whole-pelvis (WP) VMAT plans were created by each planner using a knowledge-based planning (KBP) system. Nine planners were divided into three groups of three planners each: Senior, Junior, and Beginner. Single optimization with only priority modification for all objectives was performed to stay within the dose constraints. The coefficients of variation (CVs) for dosimetric parameters were evaluated, and a plan quality metric (PQM) was used to evaluate comprehensive plan quality.
Results: Lower CVs (< 0.05) were observed at dosimetric parameters in the planning target volume for both HPO and WP plans, while the CVs in the rectum and bladder for WP plans (< 0.91) were greater than those for HPO plans (< 0.17). The PQM values of HPO plans for Cases1–5 (average ± standard deviation) were 41.2 ± 7.1, 40.9 ± 5.6, and 39.9 ± 4.6 in the Senior, Junior, and Beginner groups, respectively. For the WP plans, the PQM values were 51.9 ± 6.3, 47.5 ± 4.3, and 40.0 ± 6.6, respectively. The number of clinically acceptable HPO and WP plans were 13/15 and 11/15 in the Senior group, 13/15 and 10/15 plans in the Junior group, and 8/15 and 2/15 plans in the Beginner group, respectively.
Conclusion: Inter-planner variation in the plan quality with RapidPlan remains, especially for the complicated VMAT plans, due to planners' heuristics.

1. Introduction

Dependence on the skill and experience of a planner or institution is one of the most serious issues in intensity-modulated radiation therapy (IMRT) and volumetric-modulated arc therapy (VMAT) planning [1–4]. Knowledge-based treatment planning (KBP) has been introduced to many radiotherapy centers to overcome this problem. The RapidPlan (Varian Medical Systems, Palo Alto, CA, USA) can predict an achievable dose–volume histogram (DVH) (i.e., prediction range) for organs at risk (OARs) in a patient using a model that is trained with data from patients previously treated for the same disease [5]. Some studies have investigated whether the knowledge-based plan had plan quality comparable to that of the clinical plan [6,7] and concluded that the RapidPlan was a useful tool for consistent plan quality [8–12]. In addition, mechanical performance and dosimetric accuracy for the RapidPlan-based plan were also comparable to those of the clinical plan, which indicated the RapidPlan could be safely administered to a patient [13]. Moreover, trained models from an experienced center can be shared with other centers and can be used to identify systematic variations between multiple campuses of a single institution [14,15]. Hence, the RapidPlan may provide uniform plan quality across many centers. However, the RapidPlan has a weak point. Chang et al. showed that the performance of a RapidPlan to achieve dose constraints is still behind that of an experienced planner, and manual touch-up is necessary, although RapidPlan-based plans with a single optimization without any modifications could produce clinically acceptable plans [16]. Fogliata et al. also demonstrated that auto-generated priorities of an upper constraint for the planning target volume (PTV) and clinical target volume (CTV) require manual modification to reach institutionally acceptable dose homogeneity during the interactive phase of the

* Corresponding author.
E-mail address: hmon@med.kindai.ac.jp (H. Monzen).

https://doi.org/10.1016/j.ejmp.2019.10.032
Received 16 June 2019; Received in revised form 15 October 2019; Accepted 17 October 2019
Available online 06 November 2019

Table 1
Data for the training and trial sets.

	HPO VMAT		WP VMAT	
Volume (cc)	Training set (n = 44)	Trial set (n = 5)	Training set (n = 77)	Trial set (n = 5)
PTV				
Total	91.3 ± 31.8	89.9 ± 24.1	116.6 ± 36.5	113.7 ± 17.6
Nodal PTV				
Total			789.6 ± 159.9	689.1 ± 120.2
Rectum				
Total	62.8 ± 16.3	70.5 ± 16.1	70.0 ± 22.4	93.1 ± 26.5
Overlap with PTVs	2.9 ± 0.9	1.5 ± 0.5	5.5 ± 2.1	5.1 ± 2.5
Bladder				
Total	232.6 ± 68.0	232.0 ± 57.1	244.1 ± 65.2	190.9 ± 46.6
Overlap with PTVs	6.4 ± 2.5	7.3 ± 3.2	33.3 ± 9.5	17.7 ± 4.4
Bowel				
Total			981.6 ± 387.5	730.1 ± 435.9
Overlap with PTVs			73.5 ± 32.0	38.2 ± 35.4
Rt. femoral head				
Total	155.0 ± 18.4	168.7 ± 19.2	153.4 ± 18.2	164.4 ± 14.9
Lt. femoral head				
Total	153.2 ± 18.4	167.8 ± 22.1	153.2 ± 17.5	169.8 ± 15.4

PTV: planning target volume, HPO: hypofractionated prostate only, WP: whole pelvis

optimization for lung cancer patients, while fixed priorities of upper and lower objectives could generate acceptable plans for prostate cancer patients [17]. They showed the inaccuracy of priority given by the RapidPlan. Thus, manual modification of priorities and adding objectives should still be performed to achieve the best dose constraints in clinical practice, even though the RapidPlan was introduced. However, manual modification of the priority and objectives by the planner may nullify RapidPlan's main advantage—i.e., overcoming the differences in planning skills for planners or institutions. Scaggion et al. investigated inter-planner variability in prostate VMAT plans generated with RapidPlan assistance and demonstrated that the RapidPlan significantly reduced inter-planner variability [18]. However, their study investigated only simple VMAT plans and was performed under optimization methods limited to use and modification to know the capability of the RapidPlan.

The purpose of this study was to clarify the variations in plan quality among planners using RapidPlan-based simple and complex VMAT plans when planners with different experience optimized with the RapidPlan with modification. Forty-five hypofractionated prostate-only (HPO) VMAT plans for low- and intermediate-risk prostate cancer and 45 whole-pelvis (WP) VMAT plans for high-risk prostate cancer were created with RapidPlan by nine planners with different levels of planning experience for five patients per disease. Their plans were selected to confirm whether the plans' complexity affected the inter-planner variation. The planners were allowed only "modification of priorities" of all objectives during the optimization [17]. The plan quality was scored with a plan quality metric (PQM) [1] to evaluate the inter-planner variation.

2. Materials and methods

2.1. Patient characteristics and contouring

Data from 131 patients with prostate cancer who were consecutively treated with HPO or WP VMAT between January 2016 and March 2019 were used in this study. Of these patients, 49 classified as being at low- or intermediate-risk underwent HPO VMAT, and 82 classified as being at high-risk underwent WP VMAT. The Eclipse treatment planning system version 15.6 (Varian Medical Systems, Palo Alto, CA, USA) was used for this study. The CTV was defined as the entire prostate with a 1.5 cm proximal area of seminal vesicles for intermediate- and high-risk patients, while CTV was defined as the entire prostate for a low-risk patient. For HPO VMAT, the PTV was generated by adding a 0.6 cm margin to the CTV in all directions, except posteriorly, where a 0.4 cm margin was added. In the case of a high-risk patient, the margins were 0.8 cm and 0.5 cm, respectively. For WP VMAT, a pelvic lymph node was also defined as another target in a high-risk patient. The obturator vessels, including the common, external, and internal iliac vessels, were delineated and then expanded by adding an isotropic 0.7 cm margin to generate the nodal CTV, excluding the adjacent bone, muscle, bowel, and bladder. The nodal PTV was generated adding a 0.5 cm margin to the nodal CTV. The rectum, bladder, femoral heads, and bowels were contoured as the OARs. Details of the contouring were reported elsewhere [19]. Written informed consent was obtained from all patients, and our institution's ethics committee approved the study (IRB number 201901-48).

2.2. Model configuration

Two models for HPO and WP plans were trained based on 44 and 77 plans, respectively. These clinically delivered plans were created by senior planners between April 2016 and December 2018 using Eclipse version 10.0, and the dose distributions were not recalculated with Eclipse version 15.6. They have been accepted by both physicians and medical physicists. The structures used for training were the PTV, rectum, bladder, and femoral heads for the HPO model and the PTV, nodal PTV, rectum, bladder, bowels, and femoral heads for the WP model. The volumes of all structures for the training set and trial set are shown in Table 1. More detail is shown in Supplementary file. The structures indicated as outliers were determined based on analysis using RapidPlan version 15.6. A few outlier structures may not affect the quality of the plan made by the model [20], but all outliers were excluded in the model to eliminate their effect. Four cases of the bowels and two cases of the left femoral head were excluded as outliers from the WP model. Two cases of the bladder were excluded from the HPO model. The objective setting for each model is shown in Table 2. The models were validated through an open loop process before starting this study. Fixed priorities were used for upper and lower objectives in the targeted structures, and the line objective and priority given by the RapidPlan were used for the OARs. The line objective was placed just below the inferior boundary of the DVH estimated range. The values of fixed priorities were determined from the optimization results of the validation set, which were tuned to achieve our institution's acceptance criteria. An upper objective was placed at desirable maximum dose for the OARs.

Table 2
Objectives settings and priorities for HPO and WP models.

	HPO model				WP model			
Structure	Objective	Volume (%)	Dose (Gy)	Priority	Objective	Volume (%)	Dose (Gy)	Priority
PTV	Upper	0.0	53.5	105	Upper	0.0	54.4	105
	Lower	100.0	52.8	80	Lower	100.0	52.0	95
Nodal PTV	N/A				Upper	1.0	51.7	120
	N/A				Lower	100.0	46.8	105
Bladder	Upper	0.0	51.6	Generated	Upper	0.0	52.0	Generated
	Line			Generated	Line			Generated
Bowels	N/A				Upper	0.0	50.0	Generated
	N/A				Line			Generated
Rectum	Upper	0.0	51.6	Generated	Upper	0.0	52.0	Generated
	Line			Generated	Line			Generated
Femoral head (Each)	Upper	0.0	41.8	Generated	Upper	0.0	37.5	Generated
NTO	Automatic			100	Automatic			100

PTV: planning target volume, NTO: normal tissue objective, HPO: hypofractionated prostate only, WP: whole pelvis.

2.3. Treatment planning with RapidPlan

Of 131 patients, 10 who were consecutively treated between January 2019 and March 2019 were studied to compare the variations in plan quality among planners. Five patients were at low or intermediate risk, and five were at high risk. Two full arcs (181°–179°, clockwise and counterclockwise) and 10-MV photons were used for both the HPO and WP VMAT plans. The same collimator angles were used by the planners for each patient. Planning techniques such as the number of arcs, energy, and collimator angle were the same as those in the clinically accepted human-generated plan. The prescribed doses were 51.6 Gy (12 fractions) in the PTV for the HPO VMAT plan and 52.0 Gy in the PTV and 46.8 Gy (26 fractions) to the nodal PTV for WP VMAT using a simultaneous integrated boost technique. A boost VMAT plan delivering 26.0 Gy in 13 fractions to the PTV is concomitant with the initial WP VMAT in clinical practice, but only the WP VMAT plan was employed in this study. Dose calculations were performed using an anisotropic analytical algorithm with a grid size of 0.25 cm. Dose normalization was performed at 95% of the PTV volume and the mean dose of the PTV for the HPO and WP VMAT plans, respectively.

Nine planners with different levels of planning experience participated in this study and were divided into three groups of three planners per group (Senior, Junior, and Beginner) according to the number of VMAT planning cases they performed per week. Five HPO and five WP VMAT plans were generated by each planner using the HPO and WP models, respectively. Single optimization was performed with only modification from preset priorities (Table 2) during optimization to generate the best plan that would meet the dose constraints in our center. Any other modification—such as modification of the dose or volume of the objectives or adding an objective—was not allowed during the optimization. The dose constraints are shown in Table 3. In addition, HPO and WP VMAT plans were generated using the RapidPlan without modification during optimization to compare with the plans generated by the planners. We counted the number of clinically acceptable HPO and WP VMAT plans, which were judged according to the dose constraints in Table 3.

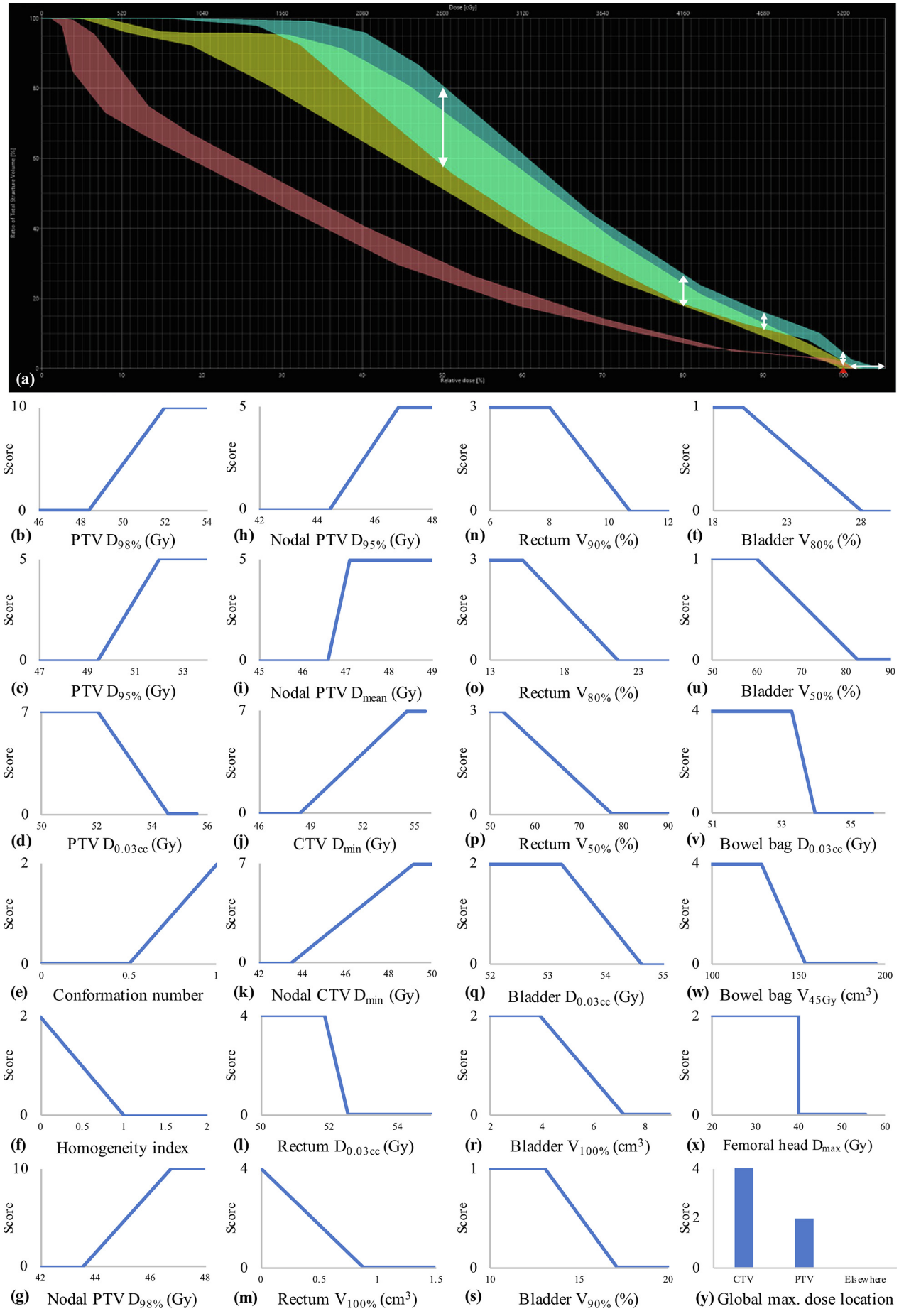
2.4. Evaluation

The coefficients of variation (CVs) for dosimetric parameters were evaluated for the HPO and WP plans. In the HPO VMAT plan, the parameters were $D_{99\%}$ (dose received at least a volume of 99%) and $D_{0.03cc}$ for the PTV; minimum dose for the CTV; $D_{0.03cc}$ and $V_{95\%}$ (volume receiving at least 95% of the prescribed dose), $V_{90\%}$, $V_{80\%}$, and $V_{50\%}$ for the rectum and bladder; and $V_{27.86Gy}$ and maximum dose for each femoral head. For the WP VMAT plan, the parameters were $D_{98\%}$, $D_{95\%}$, and $D_{0.03cc}$ for the PTV; $D_{98\%}$, $D_{95\%}$, and mean dose for the nodal PTV; minimum doses for the CTV and nodal CTV; $D_{0.03cc}$, $V_{100\%}$, $V_{90\%}$, $V_{80\%}$, and $V_{50\%}$ for the rectum and bladder; $D_{0.03cc}$ and V_{45Gy} for the bowels; and maximum dose for each femoral head. The homogeneity index (HI: defined as $HI = [D_{2\%} - D_{98\%}]/D_{50\%}$, where $D_{2\%}$, $D_{98\%}$, and $D_{50\%}$ are minimum doses to 2%, 98%, and 50% of the PTV volume, respectively) [21] and the conformation number (CN: defined as $CN = V_{PTV,\ pre}{}^2/[V_{PTV} \times V_{pre}]$, where $V_{PTV,\ pre}$ is the volume covered by the prescribed dose in the PTV, V_{PTV} is the volume of the PTV, and V_{pre} is the volume covered by the prescribed dose) [22] were also evaluated.

Table 3
Clinical goals of the HPO and WP VMAT plans in our institution.

	HPO VMAT plan		WP VMAT plan	
Structure	Dosimetric parameter	Clinical goal	Dosimetric parameter	Clinical goal
PTV	$D_{99\%}$	≥ 49.05 Gy	$D_{98\%}$	≥ 48.36 Gy
	N/A		$D_{95\%}$	≥ 49.40 Gy
	$D_{0.03cc}$	< 55.21 Gy	$D_{2\%}$	< 53.04 Gy
Nodal PTV	N/A		$D_{98\%}$	≥ 43.50 Gy
	N/A		$D_{95\%}$	≥ 44.40 Gy
	N/A		D_{mean}	46.60 – 47.10 Gy
CTV	D_{min}	≥ 50.57 Gy	D_{min}	≥ 48.36 Gy
Nodal CTV	N/A		D_{min}	≥ 43.50 Gy
Bladder	$V_{50\%}$	< 50%	$V_{60\%}$	< 55%
	$V_{80\%}$	< 20%	N/A	
	$V_{90\%}$	< 10%	$V_{90\%}$	< 30%
	$V_{95\%}$	< 3.0 cc	N/A	
	$D_{0.03cc}$	< 54.18 Gy	$D_{0.03cc}$	< 53.04 Gy
Bowels	N/A		$D_{2\%}$	< 50.0 Gy
	N/A		V_{45Gy}	< 195.0 cc
Rectum	$V_{50\%}$	< 50%	$V_{60\%}$	< 50%
	$V_{80\%}$	< 20%	N/A	
	$V_{90\%}$	< 10%	$V_{90\%}$	< 25%
	$V_{95\%}$	< 3.0 cc	N/A	
	$D_{0.03cc}$	< 54.18 Gy	$D_{0.03cc}$	< 53.04 Gy
Femoral head (Each)	D_{max}	< 41.80 Gy	D_{max}	< 40.0 Gy
	$V_{27.86Gy}$	< 10 cc	N/A	

$V_{n\%}$: volume received (n%) of the prescribed dose, $D_{n\%}$: minimum dose receiving at least the n% volume; HP VMAT: hypofractionated prostate-only VMAT, WP VMAT: whole-pelvis VMAT.



(caption on next page)

Fig. 1. Example of score assignments for the plan quality metric (PQM) for whole-pelvis volumetric modulated arc therapy (WP VMAT) plans. To assign numerical scores to patients with clearly different achievable sparing of organs at risk (OARs), prediction ranges generated by the model were used in each patient, and our clinical dose constraints were used for the targets in all patients. (a) Prediction ranges for the rectum (yellow), bladder (cyan), and bowels (pink). White arrows from right to left show prediction ranges to assign the scores at dosimetric parameters ($D_{0.03cc}$, $V_{100\%}$, $V_{90\%}$, $V_{80\%}$, and $V_{50\%}$) for the bladder. Details of score assignments are shown for (b) $D_{98\%}$ of the PTV; (c) $D_{95\%}$ of the PTV; (d) $D_{0.03cc}$ of the PTV; (e) conformation number; (f) homogeneity index; (g) $D_{98\%}$ of the nodal PTV; (h) $D_{95\%}$ of the nodal PTV; (i) D_{mean} of the nodal PTV; (j) D_{min} of the CTV; (k) D_{min} of the nodal CTV; (l) $D_{0.03cc}$ of the rectum; (m) $V_{100\%}$ of the rectum; (n) $V_{90\%}$ of the rectum; (o) $V_{80\%}$ of the rectum; (p) $V_{50\%}$ of the rectum; (q) $D_{0.03cc}$ of the bladder; (r) $V_{100\%}$ of the bladder; (s) $V_{90\%}$ of the bladder; (t) $V_{80\%}$ of the bladder; (u) $V_{50\%}$ of the bladder; (v) $D_{0.03cc}$ of the bowels; (w) V_{45Gy} of the bowels; (x) D_{max} of each femoral head; (y) global maximum dose location. The final PQM score is the sum of all components (maximum possible PQM = 100.0). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 4
Coefficient of variations to dosimetric parameters from nine planners of five HPO VMAT and five WP VMAT cases.

Dosimetric parameter	HPO VMAT plan						Dosimetric parameter	WP VMAT plan					
	Case 1	Case 2	Case 3	Case 4	Case 5	Average		Case 1	Case 2	Case 3	Case 4	Case 5	Average
PTV							PTV						
$D_{99\%}$	0.00	0.00	0.01	0.01	0.01	0.01	$D_{98\%}$	0.01	0.04	0.01	0.01	0.02	0.02
							$D_{95\%}$	0.01	0.02	0.01	0.01	0.01	0.01
$D_{0.03cc}$	0.01	0.01	0.01	0.01	0.01	0.01	$D_{0.03cc}$	0.00	0.01	0.00	0.01	0.01	0.01
CN	0.02	0.03	0.02	0.02	0.02	0.02	CN	0.05	0.06	0.03	0.05	0.02	0.04
HI	0.16	0.13	0.17	0.20	0.13	0.16	HI	0.24	0.46	0.18	0.23	0.30	0.28
							Nodal PTV						
							$D_{98\%}$	0.07	0.01	0.02	0.02	0.01	0.03
							$D_{95\%}$	0.05	0.01	0.01	0.01	0.01	0.02
							D_{mean}	0.02	0.00	0.01	0.01	0.00	0.01
CTV							CTV						
D_{min}	0.01	0.01	0.01	0.01	0.01	0.01	D_{min}	0.00	0.01	0.01	0.02	0.01	0.01
							Nodal CTV						
							D_{min}	0.13	0.02	0.04	0.06	0.03	0.05
Rectum							Rectum						
$D_{0.03cc}$	0.01	0.01	0.01	0.01	0.01	0.01	$D_{0.03cc}$	0.00	0.00	0.00	0.01	0.02	0.01
$V_{95\%}$	0.05	0.05	0.13	0.17	0.06	0.09	$V_{100\%}$	0.53	0.53	0.72	0.37	0.91	0.61
$V_{90\%}$	0.03	0.04	0.09	0.16	0.05	0.07	$V_{90\%}$	0.11	0.23	0.11	0.19	0.25	0.18
$V_{80\%}$	0.02	0.03	0.05	0.12	0.04	0.05	$V_{80\%}$	0.06	0.18	0.07	0.13	0.21	0.13
$V_{50\%}$	0.02	0.03	0.03	0.10	0.03	0.04	$V_{50\%}$	0.01	0.18	0.07	0.18	0.10	0.11
Bladder							Bladder						
$D_{0.03cc}$	0.01	0.01	0.01	0.01	0.01	0.01	$D_{0.03cc}$	0.01	0.00	0.01	0.01	0.00	0.01
$V_{95\%}$	0.02	0.03	0.03	0.05	0.04	0.04	$V_{100\%}$	0.48	0.26	0.35	0.35	0.26	0.34
$V_{90\%}$	0.02	0.03	0.03	0.06	0.03	0.04	$V_{90\%}$	0.05	0.09	0.10	0.10	0.33	0.13
$V_{80\%}$	0.02	0.04	0.03	0.05	0.02	0.03	$V_{80\%}$	0.05	0.11	0.12	0.09	0.33	0.14
$V_{50\%}$	0.02	0.02	0.03	0.05	0.02	0.03	$V_{50\%}$	0.01	0.06	0.10	0.05	0.34	0.11
							Bowels						
							$D_{0.03cc}$	0.02	0.01	0.01	0.03	0.00	0.01
							V_{45Gy}	0.36	0.10	0.14	0.14	0.09	0.17
Rt. femoral head							Rt. femoral head						
D_{max}	0.03	0.11	0.07	0.10	0.11	0.08	D_{max}	0.03	0.05	0.04	0.04	0.05	0.04
$V_{27.86Gy}$	2.02	3.00	1.60	1.19	2.49	2.06							
Lt. femoral head							Lt. femoral head						
D_{max}	0.10	0.14	0.10	0.04	0.07	0.09	D_{max}	0.04	0.06	0.07	0.04	0.03	0.05
$V_{27.86Gy}$	0.00	3.00	2.63	1.09	3.00	1.94							

PTV: planning target volume, CTV: clinical target volume, CN: conformation number, HI: homogeneity index, HPO VMAT: hypofractionated prostate only VMAT, WP VMAT: whole-pelvis VMAT.

The PQM was used to evaluate the comprehensive plan quality. The PQM is a user-defined metric intended to quantify and compare the plan quality by mimicking the judgment of a physician, which consisted of a set of clear and specific plan objectives. Each plan objective is scored according to a mathematically described physician judgment criterion. Scoring of the above dosimetric parameters was performed as follows: The scores for the PTV, nodal PTV, CTV, nodal CTV, HI, and CN were assigned according to the dose constraints used in our center. The scores for the OARs were assigned based on prediction ranges that were estimated as being achievable by the RapidPlan for each patient. The score for the global maximum dose location was assigned according to Nelms et al. [1]. The final PQM score is the sum of all components (maximum possible PQM = 100.0). Fig. 1 shows an example of the scoring details for the WP VMAT plan. (An example of the HPO VMAT plan is shown in a Supplementary File.) PQM scores were converted to percentage PQM (PQM%), representing the ratio of the achieved score to the maximum achievable score [18,23]. The Steel–Dwass test was used to compare the PQM scores of the three groups. Differences were considered significant at a value of $p < 0.05$. Clinically accepted manual plans created without the RapidPlan (i.e., human-generated plan devised via a trial-and-error process) were also evaluated using the PQM score for 10 patients.

3. Results

Five HPO VMAT plans and five WP VMAT plans were created by each planner, and the total 45 HPO VMAT plans and 45 WP VMAT plans were then analyzed. Table 4 shows the CVs of dosimetric parameters from nine planners for each case included in the HPO and WP VMAT plans. Except for the HI in the PTV, the CVs at dosimetric parameters were < 0.05 in both HPO and WP VMAT plans. The average CVs at $D_{0.03cc}$, $V_{100\%}$, $V_{90\%}$, $V_{80\%}$, and $V_{50\%}$ for the rectal doses were 0.01, 0.09, 0.07, 0.05, and 0.04, respectively, in the HPO VMAT plans vs. 0.01, 0.61, 0.18, 0.13, and 0.11, respectively, in the WP VMAT plans. For the bladder doses, the average CVs at $D_{0.03cc}$, $V_{100\%}$, $V_{90\%}$, $V_{80\%}$, and $V_{50\%}$ were 0.01, 0.04, 0.04, 0.03, and 0.03, respectively, in

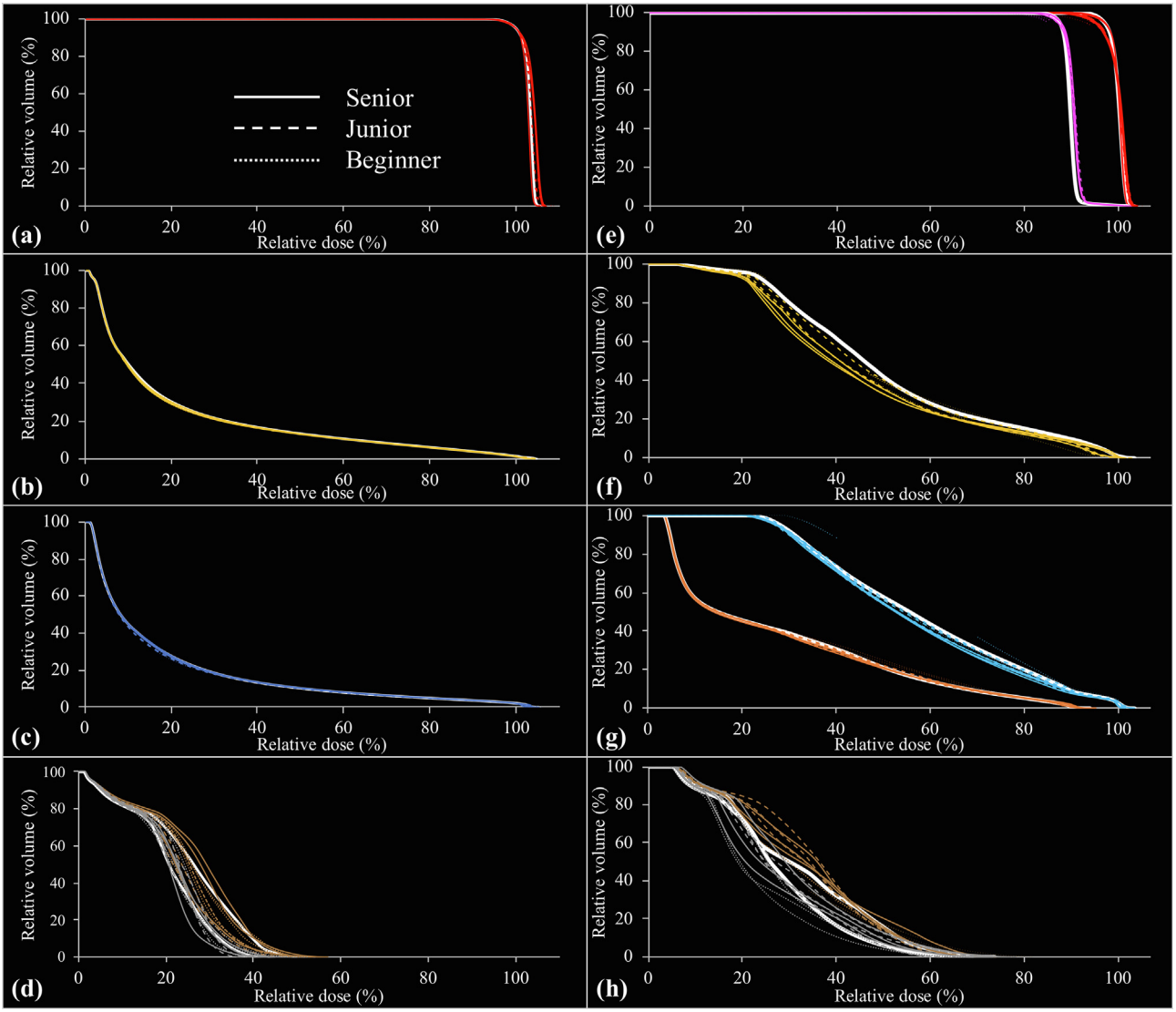


Fig. 2. Example comparison of dose–volume histograms (DVH) among planners. (a–d) Case 1 in the HPO VMAT plan. (e–h) Case 3 in the WP VMAT plan. The red, yellow, cyan, brown, gray, magenta, and orange lines indicate the PTV, rectum, bladder, right femoral head, left femoral head, nodal PTV, and bowels, respectively. The white lines are those of the RapidPlan-based plan without modifications. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the HPO VMAT plans vs. 0.01, 0.34, 0.13, 0.14, and 0.11, respectively, in the WP VMAT plans. The CVs of the OARs for the WP VMAT plans were greater than those for the HPO VMAT plans. The DVHs of the HPO VMAT (Case 1) and WP VMAT (Case 3) plans are shown in Fig. 2.

The comparison of PQM scores in each case are shown in Fig. 3. Average scores for Cases 1–5 in the HPO and WP VMAT plans are summarized in Table 5. Widely ranging scores were observed in the WP VMAT plans relative to the HPO VMAT plans. The Senior group's average PQM scores were higher than those in the Beginner group in all WP VMAT cases ($p < 0.05$), although similar scores were observed in all groups in the HPO VMAT plans ($p > 0.05$). With RapidPlan-based plans without modification, PQM scores were 46, 47, 43, 34, and 46, respectively, in Cases 1–5 in the HPO VMAT plans and 53, 42, 50, 46, and 47 in Cases 1–5 in the WP VMAT plans. The PQM scores of manual plans without the RapidPlan were 41, 50, 38, 31, and 41 for the HPO patients, and 46, 46, 55, 50, and 49 for the WP patients. The PQM scores are also summarized in Table 5.

There were 34 and 23 clinically acceptable HPO VMAT and WP VMAT plans (pass rates of 75.6% and 51.1%) in the RapidPlan-based plans with modification, respectively. The clinically acceptable plans were distributed among Cases 1–5, respectively, as follows: HPO VMAT: 13 (3, 3, 3, 1, 3) and WP VMAT: 11 (2, 1, 2, 3, 3), with pass rates of 86.7% and 73.3%, respectively, in the Senior planner group; 13 (2, 3, 3, 2, 3) and 10 (0, 2, 2, 3, 3), with pass rates of 86.7% and 66.7%, respectively, in the Junior planner group; and 8 (1, 1, 2, 1, 3) and 2 (0, 1, 0, 0, 1), with pass rates of 53.3% and 13.3% in the Beginner planner group. These results were achieved with the dose constraints for all structures in the HPO and WP VMAT plans. In the RapidPlan-based plans without modifications, the number of clinically acceptable plans were 4 and 2 (pass rates 80.0% and 40.0%), respectively. Most of the failed plans did not meet the dose constraints at $D_{0.03cc}$ in the rectal or bladder doses in the HPO VMAT plans. The failed plans did not meet the dose constraints at D_{mean} to the nodal PTV or the D_{max} to the femoral heads in the WP VMAT plans.

4. Discussion

To evaluate the variations in plan quality of the HPO and WP VMAT plans generated using the RapidPlan by nine planners with different levels of planning experience, the plan quality of 45 HPO and 45 WP

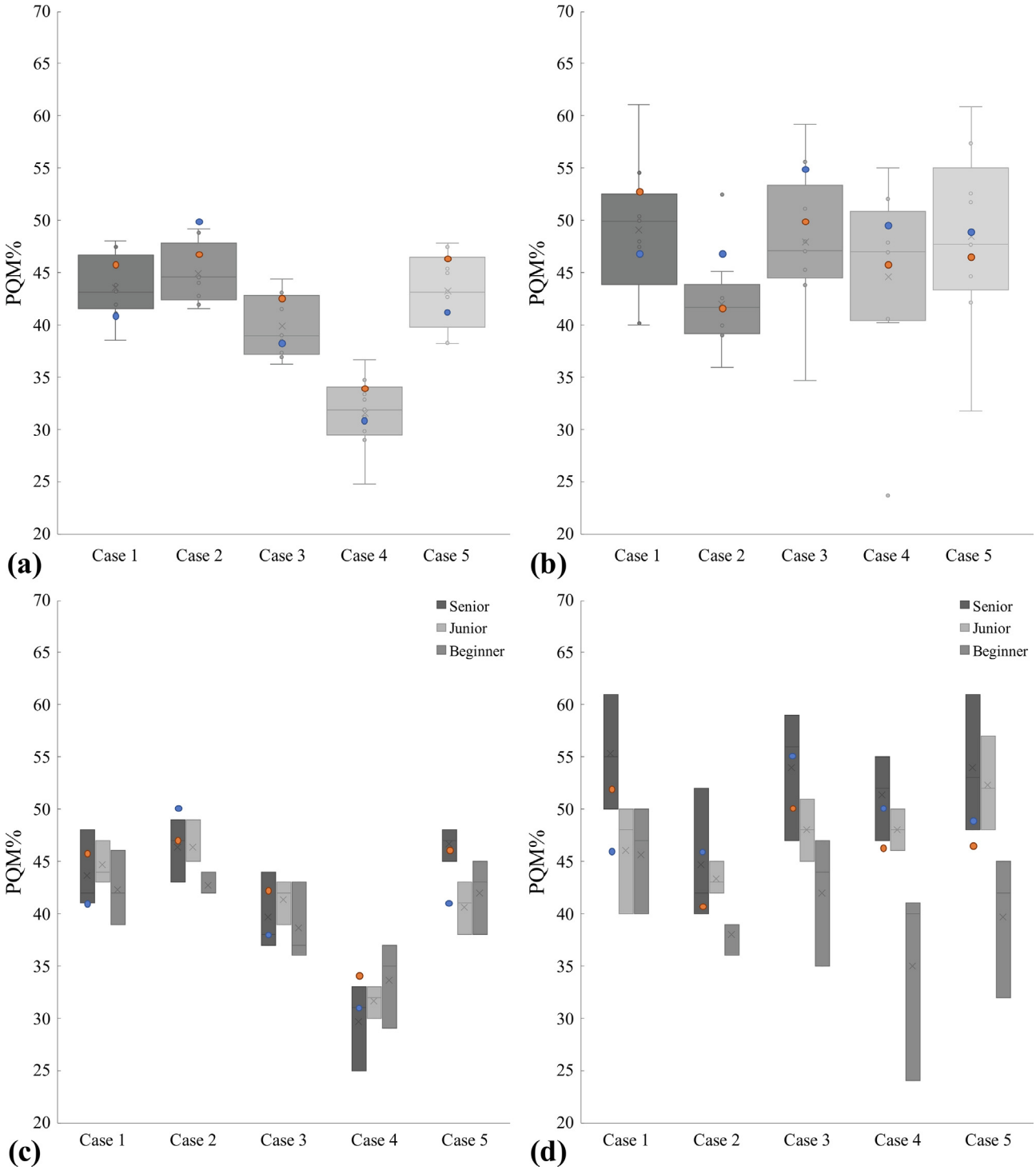


Fig. 3. Comparison of score variations in the plan quality metric (PQM) for (a) HPO VMAT plan cases and (b) WP VMAT plan cases. Comparison of score variations among Senior, Junior, and Beginner planner groups in (c) the HPO VMAT plan and (d) WP VMAT plan cases. The central line in the box plot indicates the median value, and lower and upper lines are the first and third quartiles, respectively. The data are plotted as individual filled circles, and the filled circles outside whisker show outliers. Whisker values are not considered outliers. The cross shows the mean value. The blue and orange circles show PQM scores for the RapidPlan based plan without modifications and the human-generated plan, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

VMAT plans were analyzed using the PQM score and CVs. Small variations in PQM scores were observed in the HPO VMAT plans, whereas PQM scores were wide-ranging in the WP VMAT plans, even when RapidPlan was used. Additionally, the PQM score was significantly higher in the Senior group than in the Beginner group with the WP VMAT plan, whereas there were no significant differences when using the HPO VMAT plan (Table 5). Thus, the inter-planner variation in plan quality due to the planner's skill and experience was apparent with the manual modification of priorities being performed to produce the best plan to meet the dose constraints in the complicated VMAT plan.

Most HPO VMAT plans have been more clinically acceptable than the WP VMAT plans (pass rates 75.6% vs. 51.1%) in RapidPlan-based

Table 5
Summary of mean PQM scores among three groups.

	HPO VMAT					WP VMAT				
	Senior	Junior	Beginner	Auto[a]	Manual[a]	Senior	Junior	Beginner	Auto[a]	Manual[a]
Case 1	44	45	42	46	41	55	46	46	53	46
Case 2	46	46	43	47	50	45	43	38	42	46
Case 3	40	41	39	43	38	54	48	42	50	55
Case 4	30	32	34	34	31	51	48	35	46	50
Case 5	47	41	42	46	41	54	52	40	47	49
p-value	N/A	> 0.05	> 0.05	> 0.05	> 0.05	N/A	> 0.05	< 0.05	> 0.05	> 0.05

PQM: plan quality metric.
[a] Auto: RapidPlan without modification; Manual: human-generated plan.

plans with modification. On the other hand, the number of clinically acceptable plans were 4 and 2 (pass rates 80.0% and 40.0%), for the RapidPlan-based HPO and WP VMAT plans without modification, respectively, which was comparable to that for human-generated plans based on PQM scores. Although manual modification of priorities was nearly unnecessary in the HPO VMAT plan, the Beginner planners modified the priorities actively during optimization. Consequently, there were fewer clinically acceptable plans generated by the Beginner planners than by those using the RapidPlan-based plans without modification (53.3% vs. 80.0%), although the same model was used. In contrast, Senior and Junior planners carried out priority modification successfully in HPO VMAT plans (86.7% vs. 80.0%). Alternatively, WP VMAT plans required modified priorities to meet dose constraints during optimization. More correct modifications were performed by Senior planners, and the clinically acceptable plans increased relative to cases without modification (73.3% vs. 40.0%). Castriconi et al. also showed that the plan quality of the RapidPlan generated plans with planner interaction improved compared with those with the RapidPlan without planner interaction [24]. In our study, the plan quality depended on planner experience when manual modification of priority was required for RapidPlan-based optimization of a complicated plan. Scaggion et al. studied inter-planner variability using simple prostate VMAT plans and concluded that the RapidPlan could reduce inter-planner variability [18]. This result is consistent with the results using the HPO VMAT plan. In addition, we found that more complex VMAT plans (i.e., the WP VMAT plan) depended on the planner's skill and experience.

The PQM used in this study removes any ambiguity of the plan objectives and provides a fair platform on which to compare plan results [1]. To assign scores to patients with clearly different achievable OAR sparing, we used ranges estimated by the model to be applied to the OARs in each patient, and our clinical dose constraints were used for the targeted tissues in all patients. In this study, only modification of priority was allowed during optimization because it is required to create a plan achieving the dose constraints in RapidPlan-based optimization [16,17,25,26].

Different variations of the CVs were observed between the HPO and WP VMAT plans. Larger variations were seen at the rectal and bladder doses than with the PTV and CTV doses in the WP VMAT plans, although the line objectives were placed to the rectal and bladder structures. This result may be due to the differences in the number of OARs and targets because more structures increase the complexity of the dose constraints for all structures. The largest variations were observed at $V_{27.86Gy}$ of the femoral head in the HPO VMAT plans, but they were caused by placed the upper objective at the maximum dose point only. Other larger variations were observed at $V_{95\%}$ to the rectum and bladder. As the line objective is unable to function in the overlap region with the target volume, the difference between priorities for the upper objective at maximum dose affected the $V_{95\%}$.

The following methods may be useful for reducing inter-planner variation for consistent plan quality. 1) An experienced planner defines preset objectives at the model configuration to eliminate priority modification [27]. 2) Beginner planners are required to train for optimization without the RapidPlan. 3) Plans optimized with RapidPlan should be checked by experienced planners. The RapidPlan has a heuristic aspect, which leads to inter-planner variation during complicated VMAT planning. Planners should unify on how to use the RapidPlan, such as defining a priority modification procedure and adding an objective rule. The development of a fully automated treatment planning system with a deep learning and neural network may improve the inter-planner variation.

There are some limitations to this study. Only priority modification was allowed, and objectives placed at fixed points were used. In clinical practice, inter-planner variation may be increased by changing the point of the objective. HPO and WP VMAT plans were created for five patients each. More samples are required to perform statistical analyses. Beginners should not generate plans to be delivered clinically. The KBP should be useful only for training and for didactic reasons. All PQM scores were low compared to the maximum in this study, although the quality of plans was not absolutely low. It was caused by assigned scores to the targets. Maximum scores for $D_{98\%}$ and $D_{95\%}$ of the PTV were assigned to 52 Gy and maximum scores for $D_{98\%}$ and $D_{95\%}$ of the nodal PTV were assigned to 46.8 Gy in WP VMAT plans. These doses are the same to the prescription doses. The results may suffer of the limited range of the PQM scores.

5. Conclusion

Inter-planner variation in the plan quality with RapidPlan remains, especially for the complicated VMAT plans, due to planners' heuristics.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

We thank Ashleigh Cooper, PhD, from Edanz Group (www.edanzediting.com/ac) for editing a draft of this manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ejmp.2019.10.032.

References

[1] Nelms BE, Robinson G, Markham J, Velasco K, Boyd S, Narayan S, et al. Variation in external beam treatment plan quality: an inter-institutional study of planners and planning systems. Pract Radiat Oncol 2012;2(4):296–305. https://doi.org/10.1016/j.prro.2011.11.012.

[2] Chung HT, Lee B, Park E, Lu JJ, Xia P. Can all centers plan intensity-modulated radiotherapy (IMRT) effectively? An external audit of dosimetric comparisons between three-dimensional conformal radiotherapy and IMRT for adjuvant chemoradiation for gastric cancer. Int J Radiat Oncol Biol Phys 2008;71(4):1167–74. https://doi.org/10.1016/j.ijrobp.2007.11.040.
[3] Batumalai V, Jameson MG, Forstner DF, Vial P, Holloway LC. How important is dosimetrist experience for intensity modulated radiation therapy? A comparative analysis of a head and neck case. Pract Radiat Oncol 2013;3(3):e99–106. https://doi.org/10.1016/j.prro.2012.06.009.
[4] Villaggi E, Hernandez V, Fusella M, Moretti E, Russo S, Vaccara EML, et al. Plan quality improvement by DVH sharing and planner's experience: results of a SBRT multicentric planning study on prostate. Phys Med 2019;62:73–82. https://doi.org/10.1016/j.ejmp.2019.05.003.
[5] Fogliata A, Wang PM, Francesca B, Clivio A, Nicolini G, Vanetti E, et al. Assessment of a model based optimization engine for volumetric modulated arc therapy for patients with advanced hepatocellular cancer. Radiat Oncol 2014;9:236. https://doi.org/10.1186/s13014-014-0236-0.
[6] Good D, Lo J, Lee WR, Wu J, Yin FF, Das SK. A knowledge-based approach to improving and homogenizing intensity modulated radiation therapy planning quality among treatment centers: an example application to prostate cancer planning. Int J Radiat Oncol Biol Phys 2013;87(1):176–81. https://doi.org/10.1016/j.ijrobp.2013.03.015.
[7] Nwankwo O, Mekdash H, Sihono DSK, Wenz F, Glatting G. Knowledge-based radiation therapy (KBRT) treatment planning versus planning by experts: validation of a KBRT algorithm for prostate cancer treatment planning. Radiat Oncol 2015;10:111. https://doi.org/10.1186/s13014-015-0416-6.
[8] Tol JP, Delaney AR, Dahele M, Slotman BJ, Verbakel WF. Evaluation of a knowledge-based planning solution for head and neck cancer. Int J Radiat Oncol Biol Phys 2015;91(3):612–20. https://doi.org/10.1016/j.ijrobp.2014.11.014.
[9] Fogliata A, Nicolini G, Clivio A, Vanetti E, Laksar S, Tozzi A, et al. A broad scope knowledge based model for optimization of VMAT in esophageal cancer: validation and assessment of plan quality among different treatment centers. Radiat Oncol 2015;10:220. https://doi.org/10.1186/s13014-015-0530-5.
[10] Kubo K, Monzen H, Ishii K, Tamura M, Kawamorita R, Sumida I, et al. Dosimetric comparison of RapidPlan and manually optimized plans in volumetric modulated arc therapy for prostate cancer. Phys Med 2017;44:199–204. https://doi.org/10.1016/j.ejmp.2017.06.026.
[11] Fogliata A, Reggiori G, Stravato A, Lobefalo F, Franzese C, Franceschini D, et al. RapidPlan head and neck model: the objectives and possible clinical benefit. Radiat Oncol 2017;12(1):73. https://doi.org/10.1186/s13014-017-0808-x.
[12] Wang J, Hu W, Yang Z, Chen X, Wu Z, Yu X, et al. Is it possible for knowledge-based planning to improve intensity modulated radiation therapy plan quality for planners with different planning experiences in left-sided breast cancer patients? Radiat Oncol 2017;12(1):85. https://doi.org/10.1186/s13014-017-0822-z.
[13] Tamura M, Monzen H, Matsumoto K, Kubo K, Otsuka M, Inada M, et al. Mechanical performance of a commercial knowledge-based VMAT planning for prostate cancer. Radiat Oncol 2018;13(1):163. https://doi.org/10.1186/s13014-018-1114-y.
[14] Ueda Y, Fukunaga J, Kamima T, Adachi Y, Nakamatsu K, Monzen H. Evaluation of multiple institutions' models for knowledge-based planning of volumetric modulated arc therapy (VMAT) for prostate cancer. Radiat Oncol 2018;13(1):46. https://doi.org/10.1186/s13014-018-0994-1.
[15] Berry S, Ma R, Boczkowski A, Jackson A, Zhang P, Hunt M. Evaluating inter-campus plan consistency using a knowledge based planning model. Radiother Oncol 2016;120(2):349–55. https://doi.org/10.1016/j.radonc.2016.06.010.
[16] Chang ATY, Hung AWM, Cheung FWK, Lee MCH, Chan OSH, Philips H, et al. Comparison of planning quality and efficiency between conventional and knowledge-based algorithms in nasopharyngeal cancer patients using intensity modulated radiation therapy. Int J Radiat Oncol Biol Phys 2016;95(3):981–90. https://doi.org/10.1016/j.ijrobp.2016.02.017.
[17] Fogliata A, Belosi F, Clivio A, Navarria P, Nicolini G, Scorsetti M, et al. On the pre-clinical validation of a commercial model-based optimization engine: application to volumetric modulated arc therapy for patients with lung or prostate cancer. Radiother Oncol 2014;113(3):385–91. https://doi.org/10.1016/j.radonc.2014.11.009.
[18] Scaggion A, Fusella M, Roggio A, Basso S, Pivato N, Rossato MA, et al. Reducing inter- and intra-planner variability in radiotherapy plan output with a commercial knowledge-based planning solution. Phys Med 2018;53:86–93. https://doi.org/10.1016/j.ejmp.2018.08.016.
[19] Ishii K, Ogino R, Hosokawa Y, Fujioka C, Okada W, Nakahara R, et al. Comparison of dosimetric parameters and acute toxicity after whole-pelvic vs prostate-only volumetric-modulated arc therapy with daily image guidance for prostate cancer. Br J Radiol 2016;89(1062):20150930. https://doi.org/10.1259/bjr.20150930.
[20] Delaney AR, Tol JP, Dahele M, Cuijpers J, Slotman BJ, Verbakel WF. Effect of dosimetric outliers on the performance of a commercial knowledge-based planning solution. Int J Radiat Oncol Biol Phys 2016;94(3):469–77. https://doi.org/10.1016/j.ijrobp.2015.11.011.
[21] Kataria T, Sharma K, Subramani V, Karrthick KP, Bisht SS. Homogeneity index: an objective tool for assessment of conformal radiation treatments. J Med Phys 2012;37(4):207–13. https://doi.org/10.4103/0971-6203.103606.
[22] van't Riet A, Mak AC, Moerland MA, Elders LH, van der Zee W. A conformation number to quantify the degree of conformality in brachytherapy and external beam irradiation: application to the prostate. Int J Radiat Oncol Biol Phys 1997;37(3):731–6. https://doi.org/10.1016/S0360-3016(96)00601-3.
[23] Fusella M, Scaggion A, Pivato N, Rossato MA, Zorz A, Paiusco M. Efficiently train and validate a RapidPlan model through APQM scoring. Med Phys 2018;45(6):2611–9. https://doi.org/10.1002/mp.12896.
[24] Castriconi R, Fiorino C, Broggi S, Cozzarini C, Muzio ND, Calandrino R, et al. Comprehensive intra-institution stepping validation of knowledge-based models for automatic plan optimization. Phys Med 2019;57:231–7. https://doi.org/10.1016/j.ejmp.2018.12.002.
[25] Snyder KC, Kim J, Reding A, Fraser C, Gordon J, Ajiouni M, et al. Development and evaluation of a clinical model for lung cancer patients using stereotactic body radiotherapy (SBRT) within a knowledge-based algorithm for treatment planning. J Appl Clin Med Phys 2016;17(6):263–75. https://doi.org/10.1120/jacmp.v17i6.6429.
[26] Yu G, Li Y, Feng Z, Tao C, Yu Z, Li B, et al. Knowledge-based IMRT planning for individual liver cancer patients using a novel specific model. Radiat Oncol 2018;13(1):52. https://doi.org/10.1186/s13014-018-0996-z.
[27] Kamima T, Ueda Y, Fukunaga J, Shimizu Y, Tamura M, Ishikawa K, et al. Multi-institutional evaluation of knowledge-based planning performance of volumetric modulated arc therapy (VMAT) for head and neck cancer. Phys Med 2019;64:174–81. https://doi.org/10.1016/j.ejmp.2019.07.004.