

Article

Automated Knowledge-Based Intensity-Modulated Proton Planning: An International Multicenter Benchmarking Study

Alexander R. Delaney ^{1,*}, Lei Dong ² , Anthony Mascia ³, Wei Zou ² , Yongbin Zhang ³, Lingshu Yin ², Sara Rosas ⁴, Jan Hrbacek ⁴, Antony J. Lomax ⁴, Ben J. Slotman ¹ , Max Dahele ¹ and Wilko F. A. R. Verbakel ¹ 

¹ Cancer Center Amsterdam, Department of Radiation Oncology, VU University Medical Center, De Boelelaan 1117, 1081 HV Amsterdam, The Netherlands; bj.slotman@vumc.nl (B.J.S.); m.dahele@vumc.nl (M.D.); w.verbakel@vumc.nl (W.F.A.R.V.)

² Department of Radiation Oncology, University of Pennsylvania, Philadelphia, PA 19104, USA; Lei.Dong@uphs.upenn.edu (L.D.); Wei.Zou@uphs.upenn.edu (W.Z.); Lingshu.Yin@uphs.upenn.edu (L.Y.)

³ Department of Radiation Oncology, University of Cincinnati Medical Center, 234 Goodman Street, Cincinnati, OH 45219, USA; Anthony.Mascia@cchmc.org (A.M.); Yongbin.Zhang@UCHealth.com (Y.Z.)

⁴ Paul Scherrer Institute, Center for Proton Radiotherapy, 5232 Villigen, Switzerland; Sara.Rosas@psi.ch (S.R.); jan.hrbacek@psi.ch (J.H.); tony.lomax@psi.ch (A.J.L.)

* Correspondence: alexrdelaney@gmail.com; Tel.: +31-2-044-41571

Received: 28 August 2018; Accepted: 29 October 2018; Published: 2 November 2018



Abstract: Background: Radiotherapy treatment planning is increasingly automated and knowledge-based planning has been shown to match and sometimes improve upon manual clinical plans, with increased consistency and efficiency. In this study, we benchmarked a novel prototype knowledge-based intensity-modulated proton therapy (IMPT) planning solution, against three international proton centers. **Methods:** A model library was constructed, comprising 50 head and neck cancer (HNC) manual IMPT plans from a single center. Three external-centers each provided seven manual benchmark IMPT plans. A knowledge-based plan (KBP) using a standard beam arrangement for each patient was compared with the benchmark plan on the basis of planning target volume (PTV) coverage and homogeneity and mean organ-at-risk (OAR) dose. **Results:** PTV coverage and homogeneity of KBPs and benchmark plans were comparable. KBP mean OAR dose was lower in 32/54, 45/48 and 38/53 OARs from center-A, -B and -C, with 23/32, 38/45 and 23/38 being >2 Gy improvements, respectively. In isolated cases the standard beam arrangement or an OAR not being included in the model or being contoured differently, led to higher individual KBP OAR doses. Generating a KBP typically required <10 min. **Conclusions:** A knowledge-based IMPT planning solution using a single-center model could efficiently generate plans of comparable quality to manual HNC IMPT plans from centers with differing planning aims. Occasional higher KBP OAR doses highlight the need for beam angle optimization and manual review of KBPs. The solution furthermore demonstrated the potential for robust optimization.

Keywords: proton therapy; IMPT; head and neck cancer; knowledge-based planning; model-based planning; automated

1. Introduction

The number of proton centers is increasing [1]. However, treatment planning studies for established photon modalities such as intensity-modulated radiation therapy (IMRT) and volumetric-modulated arc therapy (VMAT) indicate substantial intra- and inter-institutional

variation [2]. This will presumably also be the case for intensity-modulated proton therapy (IMPT) since IMPT treatment planning is complex and involves many steps, including: determining beam arrangements, using appropriate optimization objectives and deciding whether to use of a range shifter or bolus. While certain centers are using robust optimization, which adds further complexity, the majority of proton therapy treatments, to-date, have been delivered without this [3–6]; and although online adaptive proton therapy has been proposed to deal with daily anatomical changes, this requires faster planning solutions to be practical.

A desire to reduce variation, and produce consistently “good” plans, has largely spurred the development of automated treatment planning for photon therapy [7–10]. One such approach is knowledge-based planning, an example of which is RapidPlan™ (Varian Medical Systems, Palo Alto, CA). RapidPlan utilizes the geometries and associated dosimetry of previously created treatment plans to construct a model which can be then used to predict a range of achievable organ at risk (OAR) dose-volume histograms (DVHs) for future patients. Provided the library of plans used to derive the model is of a “high enough” quality, improvements in photon treatment plan quality and reductions in planning time and variability are attainable. Knowledge-based planning has also demonstrated a role in patient-specific plan quality assurance (QA) [11–14].

We previously illustrated the principle of using the photon-RapidPlan solution to automatically create IMPT treatment plans [15]. Since then, we have collaborated on the development of a prototype knowledge-based IMPT (proton-specific) planning solution, RapidPlanPT, in which the physical characteristics of proton beams have been appropriately modeled. In this novel study, we benchmark automated knowledge-based plans (KBPs) from this prototype solution against manual head and neck cancer (HNC) IMPT plans from 3 international protons centers. By using a model based on plans from a single center we also evaluate the versatility of such an approach.

2. Results

Summarized results for the 21 patients from all external centers are presented in Figure 1. Generally, target coverage and homogeneity were similar between KBPs and benchmark plans. Aside from the oral cavity and ipsilateral submandibular gland, KBP OAR mean dose was >2 Gy lower, on average, than in benchmark plans, with statistically significant differences in a number of OARs including the contralateral submandibular gland (8.6 Gy) and larynx (8 Gy). Figures 2 and 3 show instances of similarities and differences in dose distributions and DVHs between benchmark and KBPs. Generation of DVH-predictions, subsequent optimization and dose calculation of KBPs required 8.3 min ($n = 6$ plans). Table 1 shows the dosimetry for the seven benchmark plans and KBPs from each external center.

For center-A, all but one KBP met the aims for PTV_B Dmax of $\leq 110\%$ (patient 1 had a Dmax of 111% delivered to a volume of 0.1 cm³, however center-A accepts >110% in 0.03 cm³). While benchmark plans had, on average, lower parotid gland and oral cavity mean dose, KBPs had lower submandibular gland, esophagus and larynx mean dose. A “continue optimization” was performed to improve PTV homogeneity in four cases while in two cases it was performed for non-modeled OARs not meeting the criteria in Table 2 (assigned the maximum accepted dose as an objective). Figure 4 illustrates the similarity between the mean dose values of OARs (included in Table 2) in benchmark plans and KBPs, with the R² value, 0.91, and slope, being close to 1. One OAR, for which the KBP resulted in considerably higher dose than the respective benchmark plan, is circled in Figure 4. This was a left ear with a KBP mean dose of 31 Gy (aim <30 Gy). The largest increase in mean dose to a salivary gland, when using KBPs over benchmark plans for center-A, concerned an ipsilateral parotid gland (6.6 Gy). This increase was due to the standard, Y-shaped, beam arrangement used in the KBP compared with a five-field technique, including two oblique dorsal fields, in the benchmark plan.

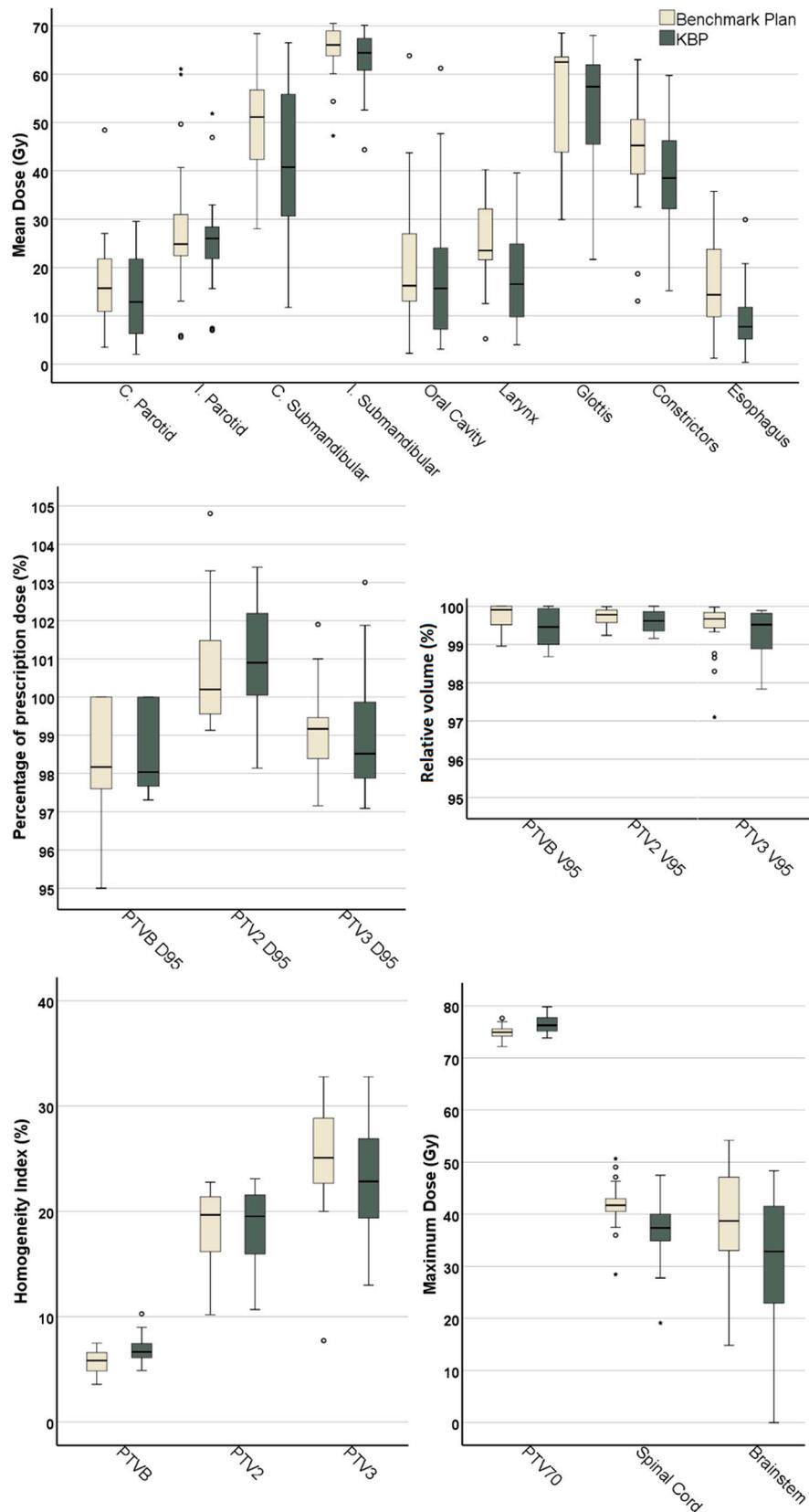


Figure 1. Box-whisker plots for all 21 patients from external centers. Dark-line in the middle of box denotes the median while the top/bottom of the box indicates 75th/25th percentile. A structure denoted with "*" indicates a statistically significant difference between benchmark plans and knowledge-based plans (KBPs) for said structure.

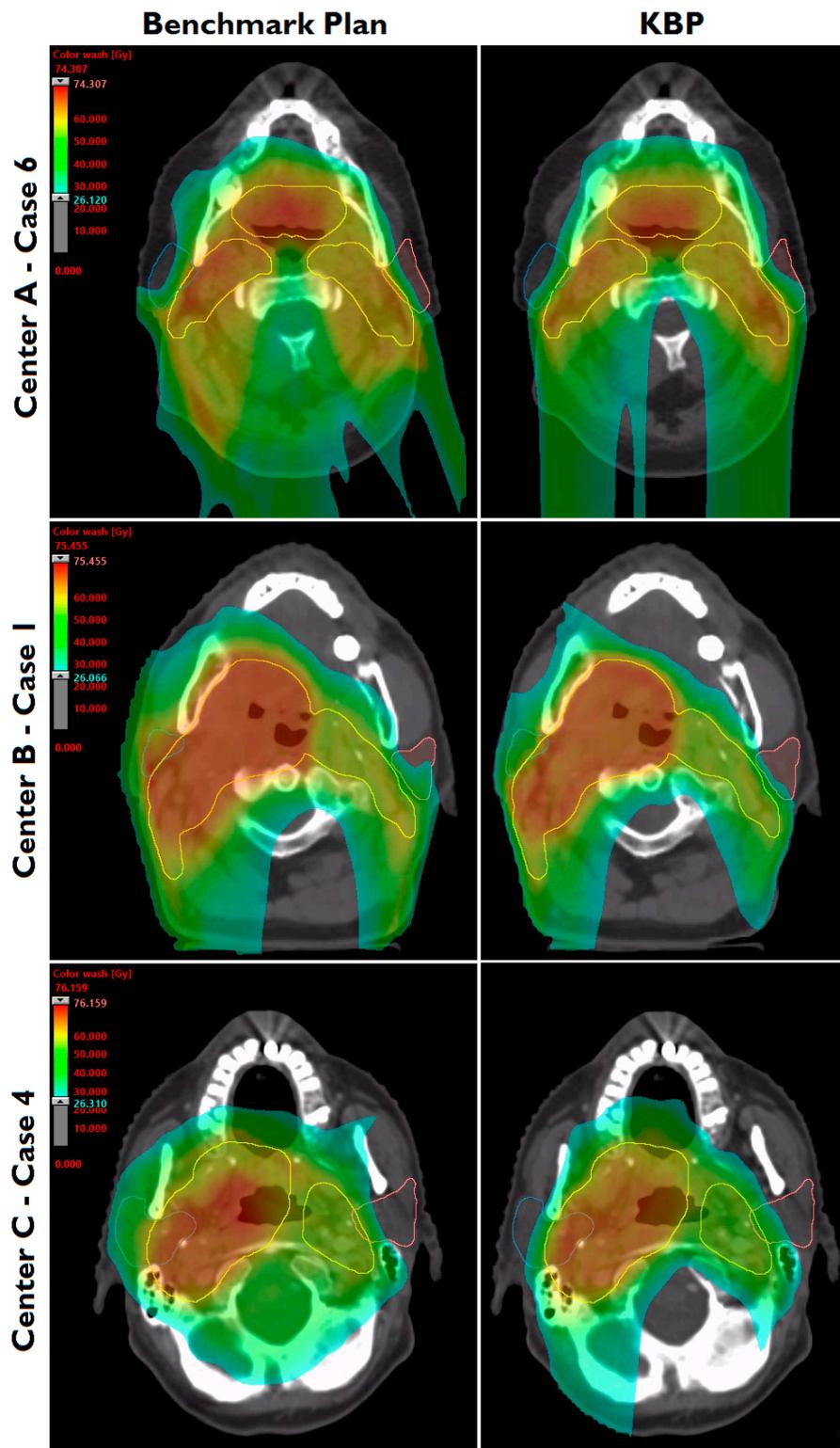


Figure 2. Dose distributions of benchmark plans and knowledge-based plans (KBPs) for one evaluation patient from each external center. Parotid glands (blue, pink) and the composite planning target volume (yellow) are shown.

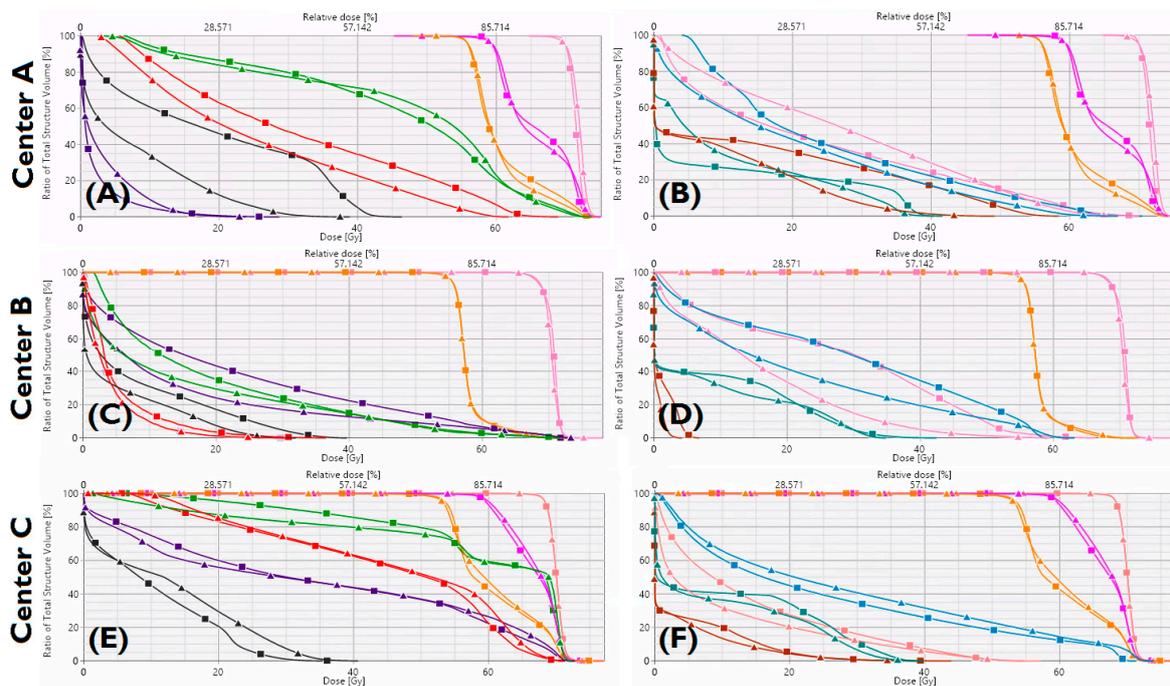


Figure 3. Benchmark plan (■) and KBP (▲) dose-volume histograms for one patient from Center-A (A,B), Center-B (C,D), Center-C (E,F). Targets: PTV_B(Pink), PTV_{E1}(magenta), PTV_{E2}(orange). Organs-at-risk: brainstem (dark-grey), oral cavity (purple), constrictors (green), larynx (red), esophagus (brown), left parotid (pink), right parotid (blue), spinal cord (petroleum-blue).

Table 1. Dosimetry averaged for the 7 benchmark plans and KBPs from each of the three external centers. Wilcoxon signed rank tests were used to identify significant ($p < 0.05$) differences between benchmark plans and KBPs, with * indicative of such differences.

PTVs	Center A		Center B		Center C	
	Benchmark	KBP	Benchmark	KBP	Benchmark	KBP
PTVB D95 (%)	100 ± 0 (100–100)	99.8 ± 0.6 (98.36–100)	97.5 ± 1.2 (95–98.4)	97.6 ± 0.4 (97.3–98.3)	97.8 ± 0.5 (97.2–98.6)	98 ± 0.3 (97.6–98.6)
PTVB V95 (%)	100 ± 0 (100–100)	99.8 ± 0.5 (98.7–100)	99.5 ± 0.4 (99–100)	99.1 ± 0.2 * (99–99.5)	99.7 ± 0.3 (99.2–100)	99.6 ± 0.4 (99–100)
PTVB MaxDose (Gy)	75.4 ± 0.9 (74.2–77)	75.9 ± 0.9 (75.2–77.5)	74.7 ± 1.7 (72.2–77.6)	77.7 ± 1.4 * (75.1–79.8)	74.5 ± 1.2 (72.7–76.2)	75.8 ± 1.8 (73.9–78.9)
PTVB MinDose (Gy)	66 ± 2.3 (62–68.7)	64.9 ± 2.8 * (60–67.7)	62.2 ± 2 (59.6–65.2)	56.5 ± 4 * (52–62.9)	62.1 ± 2.3 (58.9–64.7)	60.5 ± 4.8 (53–65.8)
HIB (%)	6.2 ± 1 (4.9–7.5)	7.6 ± 1.5 (6.2–10.3)	5.3 ± 1.2 (3.9–7.2)	7 ± 0.7 * (6–8.1)	5.7 ± 1.3 (3.6–7.2)	5.8 ± 1.3 (4.9–8.4)
PTVE1 D95 (%)	100.6 ± 0.9 (99.4–101.8)	101.3 ± 1 (100–102.5)	99.5 ± 0 (99.5–99.5)	98.1 ± 0 (98.1–98.1)	101.2 ± 2.2 (99.1–104.8)	101.1 ± 1.8 (98.7–103.4)
PTVE1 V95 (%)	99.8 ± 0.2 (99.5–99.9)	99.6 ± 0.3 (99.2–99.9)	100 ± 0 (100–100)	99.3 ± 0 (99.3–99.3)	99.6 ± 0.3 (99.2–100)	99.6 ± 0.3 (99.2–100)
HIE1 (%)	20.7 ± 2.6 (15.4–22.8)	21.3 ± 2.4 (16.1–23.1)	10.2 ± 0 (10.2–10.2)	10.7 ± 0 (10.7–10.7)	18 ± 2.7 (13.7–20.9)	17.4 ± 2 (15.4–20)
PTVE2 D95 (%)	99.9 ± 1.5 (97.2–101.9)	100.7 ± 1.7 (98.3–103)	98.8 ± 0.4 (98.1–99.3)	98 ± 0.7 * (97.1–98.9)	98.7 ± 1.1 (97.4–100.4)	98.5 ± 0.8 (97.6–99.9)
PTVE2 V95 (%)	99.4 ± 1 (97.1–100)	99.8 ± 0.1 (99.7–99.9)	99.7 ± 0.2 (99.3–99.9)	98.9 ± 0.6 * (97.8–99.5)	99.2 ± 0.7 (98.3–99.9)	99.2 ± 0.6 (98.4–99.9)

Table 1. Cont.

PTVs	Center A		Center B		Center C	
	Benchmark	KBP	Benchmark	KBP	Benchmark	KBP
HIE2 (%)	26.4 ± 3.6 (20.3–29.3)	24.3 ± 5.5 (15.8–31.2)	22.7 ± 2 (20–25.1)	22.5 ± 1.7 (19.4–24.7)	25.6 ± 8.8 (7.7–32.8)	24.1 ± 8.4 (13–32.8)
OAR Mean Doses (Gy)						
Contra. Parotid	16.5 ± 7.3 (6.4–25.7)	17.7 ± 8.7 (6.3–29.5)	20.5 ± 14.4 (3.5–48.4)	12.3 ± 8.7 * (2–26.8)	14.07 ± 7.55 (4–23.6)	13.2 ± 8.7 (4.3–27.8)
Ipsi. Parotid	25.1 ± 3.2 (21.7–31.5)	28 ± 3.3 * (24–33)	31.2 ± 11.6 (13.1–49.7)	20.8 ± 7.6 * (7.4–30.8)	28.9 ± 23 (5.6–61.1)	27.3 ± 17.5 (7–51.8)
Contra. Sub	45.1 ± 14.9 (28–64.9)	42.5 ± 15.6 (24–63.3)	51.5 ± 11.8 (39–68.4)	35.2 ± 24.8 * (11.7–66.5)	53.6 ± 7.3 (43.1–64.7)	45.5 ± 12.1 * (32.3–63)
Ipsi. Sub	63.1 ± 8.1 (47.3–68.9)	60 ± 7.81 (44.4–65)	67.9 ± 2.7 (65.5–70.5)	66.9 ± 4.2 (61.1–70.1)	64.2 ± 5.9 (54.4–69.9)	63.5 ± 6.4 (52.6–68.8)
Oral Cavity	10.5 ± 5.5 (2.2–16.2)	13.5 ± 7.7 * (3.2–22.4)	23.2 ± 18.4 (10.3–63.8)	16.5 ± 20.3 * (3.1–61.2)	28.4 ± 12.2 (8.5–43.7)	27.9 ± 15.5 (3.7–47.7)
Constrictors	40.6 ± 5.4 (32.5–48.5)	38.5 ± 6.7 (31–47.3)	42.2 ± 15.7 (18.7–51.4)	30.4 ± 11 (15.2–41.4)	50.1 ± 8.5 (40.8–63)	39.5 ± 16.7 * (18–59.8)
Esophagus	17 ± 9.7 (8.3–35.8)	12.8 ± 9.1 * (5.9–30)	16.7 ± 8.3 (1.2–25.4)	7.1 ± 4.8 * (0.4–15.8)	12.4 ± 13.4 (1.5–31)	7.2 ± 6.3 (0.4–13.2)
Larynx	28.5 ± 6.9 (21.6–40.3)	25.2 ± 7 * (18.7–39.6)	22.2 ± 10.8 (5.2–36)	9.6 ± 3.4 * (4–14.3)		
Glottis					53.7 ± 16.3 (29.9–68.5)	50.9 ± 18.3 (21.7–68.1)
OAR Max Doses (Gy)						
Spinal Cord	39 ± 5.2 (28.5–42)	38.8 ± 3.4 (34.7–43.1)	45 ± 3.2 (41.9–50.7)	34.4 ± 4.5 * (27.8–38.7)	40.9 ± 3.8 (37.5–49)	36.7 ± 8.7 (19.1–47.5)
Brainstem	43.4 ± 7.3 (34.4–53)	42.5 ± 5.1 (33.7–48.3)	34.2 ± 12.6 (15.8–54.2)	27.3 ± 6.7 (18.3–36.7)	34.5 ± 15.3 (14.8–52.4)	22 ± 18.8 * (0–41.5)

Table 2. General planning-aims for the three external centers. *: Center A constraints for organs-at-risk (OARs) are on OARminusPTV delineations. We report dosimetric results based on entire OAR structures.

PTVs	Center A *	Center B	Center C
D95%	≥100%		≥95%
V95%		≥99%	
Dmax	≤110%	≤115%	
D99%	≥93%		≥90%
D2%			≤107%
D5%			≤105%
Max Dose (Gy)			
Spinal Cord	≤45	≤45	≤41
Spinal Cord PRV	≤50		
Brainstem	≤54	≤54	≤62
Brain		ALARA	≤65
Optic Nerve	≤54	≤54	≤60
Chiasm	≤54	≤54	≤60
Esophagus			≤71.4

Table 2. Cont.

PTVs	Center A *	Center B	Center C
Larynx			≤71.4
TMJoint			≤70
Trachea			≤71.4
Eye			≤50
Mandible-PTV	≤70		
Mean Dose (Gy)			
Parotid	<26	<26	<24
Submandibular	<30	<26	
Oral Cavity	<20	<30	<45 (<35 if distal to target)
Pharynx	<50		<45
Larynx	<20	<30	
Esophagus	<20	<30	<40 if not in PTV
Middle Ears	<30	<30	
Eye	<35	ALARA	<35
Glottis			<45 (<20 if distal to target)
Trachea			<40

All center-B KBPs met PTV_B planning-aims despite a small, but statistically significant, decrease in PTV_B V95% (Table 1). On average, KBP PTV_{E1}/PTV_{E2} V95% was 0.7/0.8% lower than in the benchmark plan. KBPs improved all OAR metrics in Table 1, with a ≥8.2 Gy statistically significant reduction in mean dose to the parotid glands. The largest, statistically significant, improvements in sparing, on average, were in the contralateral submandibular gland (16.3 Gy) and larynx (12.6 Gy). A “continue optimization” was performed to improve PTV homogeneity in all cases however it was not required in any case to improve dosimetry for OARs excluded from the model. For instance, in three patients PTV_{E2} did not meet the planning-aims of V95% ≥99% after one round of optimization. This was resolved by, after optimization, converting any cold spots in PTV_{E2} into contours and running the “continue optimization”, including additional objectives for the cold spots, ensuring planning-aims were met.

All center-C KBPs met PTV coverage-aims (D95% ≥95%), with minimal differences in PTV coverage and homogeneity relative to benchmark plans. Averaged over seven patients, KBPs improved all OAR sparing (Table 1); noticeable improvements occurred in the contralateral submandibular gland (8.1 Gy) and constrictor muscles (10.6 Gy). There was good correspondence of mean dose to the oral cavity between KBPs and benchmark plans (Figure 4). A “continue optimization” was performed to improve PTV homogeneity in all cases, two of which it was also used to improve dosimetry for OARs excluded from the model. However, for these two instances, in contrast to the benchmark plans, the KBP did not meet planning-aims and resulted in inferior sparing: (1) the V_{54 Gy} for the left brachial plexus of patient six was 1.1 cm³ (aim ≤0.05 cm³) and (2) the V_{52 Gy} for the right brachial plexus of patient seven was 0.6 cm³ (aim ≤0.05 cm³) when using the KBP.

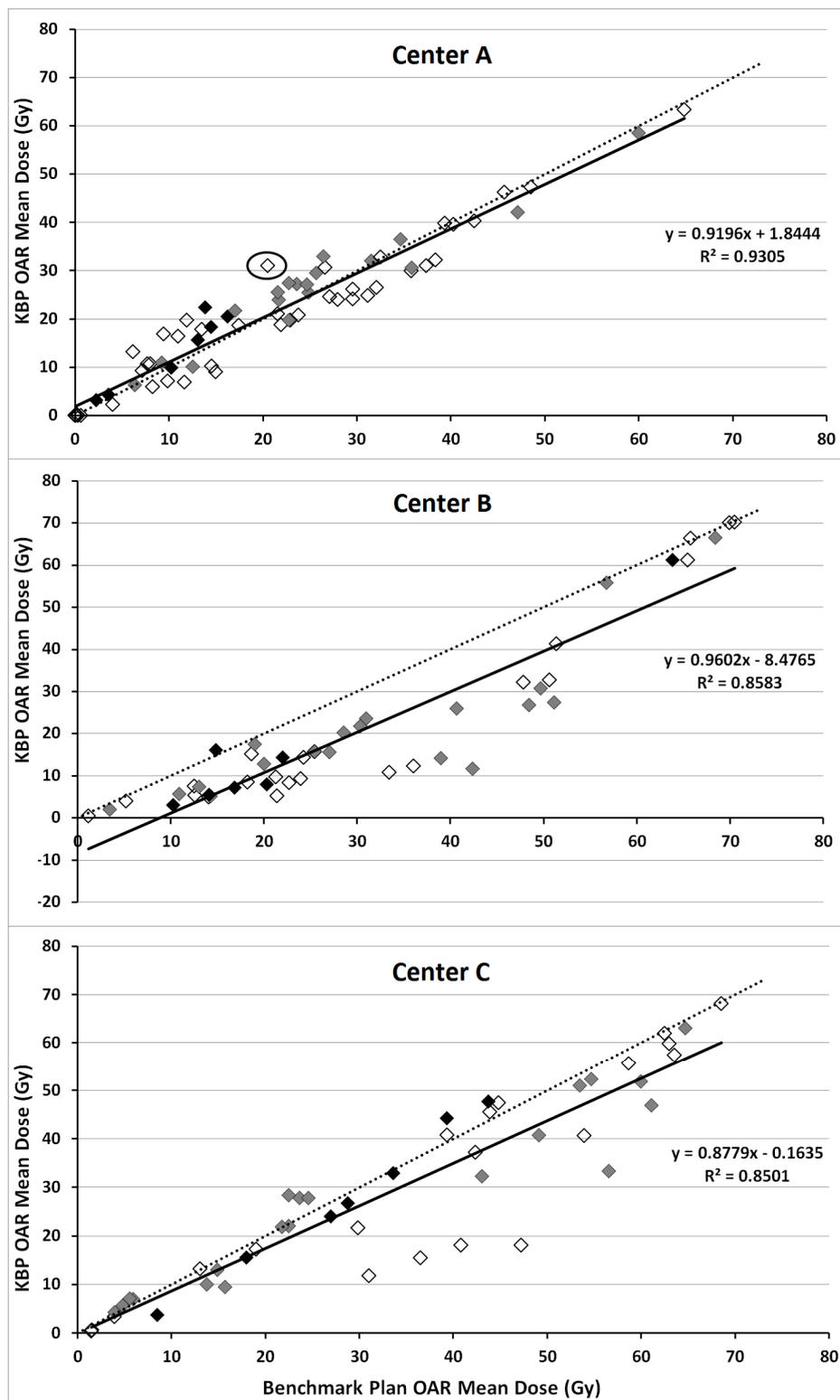


Figure 4. Correlation between mean dose of the salivary glands (grey), oral cavity (black) and all other OARs listed under “mean dose” from Table 2 (white) in knowledge-based plans (KBPs) (vertical-axis) and respective benchmark plans (horizontal-axis). Dotted line indicates line through the origin with slope of 1 while solid line represents linear fit through data points.

3. Discussion

This study, aimed at providing proof-of-principle, demonstrated that a knowledge-based planning solution, using a model library comprising plans from a single center with a standard planning technique, could create good quality HNC IMPT plans when benchmarked against plans from three experienced proton centers. The slightly different planning-aims of each center demonstrated the versatility of this automated approach. Creation of KBPs was efficient, requiring <10 min for generating predictions, subsequent optimization and dose calculation. Occasional higher OAR doses in KBPs highlights the need for manual review of KBPs and an understanding of the limitations of such a knowledge-based approach using fixed beam orientations and clinic-specific planning-aims. Knowing how “good” the plans are in a model library remains a limitation with approaches like that used in RapidPlan/RapidPlanPT. This is one reason why plans from multiple experienced proton centers have been used to benchmark the KBPs in this study. The fact that KBPs were at least comparable to or better than the benchmark plans, indicates that model library plans were of a reasonable quality.

For OARs, common to both the model and benchmark plans, KBPs provided comparable or improved sparing over respective benchmark plans. However, certain OARs/PTVs had an inferior dose distribution in KBPs. In some cases, this concerned OARs not included in the model, such as the ear of center-A and brachial plexus of center-C. For such OARs the maximum accepted dose in the planning aim was used as an objective in the KBP. However, because it was not attempted to reduce dose below this, and because in some cases the optimization weighting may have been too low, dose may have been higher than necessary in the KBP. This reflects a limitation when using a clinic-specific model for other centers: not all OARs may be the same. Furthermore, center-A oral cavity delineation was inconsistent, due to inter-physician differences, and the higher oral cavity mean dose (13.5 Gy versus 10.5 Gy) for KBPs of center-A may be attributable to the delineation of the oral cavity in patients 1–4. These oral cavity structures were >200% larger than cases 5–7, resulting in on average 4.4 Gy higher doses in the KBPs compared with benchmark plans. Although, visually, these contours differed from those in the model, the software did not flag them as volumetric outliers. Patients 5–7 had, on average, only 1.1 Gy higher oral cavity dose in KBPs. While these increases in mean dose may not be substantial, users should nonetheless check that OAR delineation/geometry of a prospective patient is similar to that in the model.

Since modeling partly depends on the beam-arrangement of plans in the model library, it was decided that in order to obtain the most accurate predictions for prospective patients, KBPs should have the same beam-arrangement as those in the model-library. Therefore, this study shows the result of straight-forward, standard three-field automated IMPT optimization without beam angle optimization. This could be a limitation of the study and KBP plan quality may be further improved with patient-specific beam-arrangements (as in the benchmark plans). Future investigations should test the applicability of such a model for differing beam arrangements [16] and include differing beam-arrangements in the model library. Optimized beam angle selection in IMPT is under investigation [17,18] and could potentially be integrated into future IMPT planning solutions. Since all external centers used the same treatment planning system, the versatility of our approach, with respect to plans from multiple vendors, was not tested. Furthermore, external centers used a bolus/range shifter, and differing beam data [19,20] and target margins. While examining the effect of such parameters on the results was beyond the scope of this study, previous work has alluded to these issues. Langner et al. compared commissioning beam data for two separate proton treatment centers which obtained their beam data using independent dosimeters and found that spot profiles were very similar both intra-institutionally (between treatment rooms) and inter-institutionally [21]. Minimal outlier removal was performed in the current analysis and while previous work has indicated that outlier removal may not always be necessary for RapidPlan photon KBPs [22,23], this has not yet been tested for RapidPlanPT. Finally, we note that the software used in this study is still under refinement and improvements are possible.

The KBPs in this study were non-robustly optimized, consistent with general clinical practice at the outset of this study [3–6]. The results are therefore relevant to contemporary practice. Whether RapidPlanPT can also be used for robust optimization requires further investigation. Since center-A began using robust optimization during the course of this study, and interest in robust optimization is increasing, we performed a preliminary test of the potential of using RapidPlanPT with a model based on non-robustly optimized plans to be used for robust optimization. Center-A provided clinical, robustly optimized plans for three new cases. Using the same model (comprising non-robustly optimized treatment plans), we created robustly optimized KBPs for these patients. All three KBPs met center-A's robustness criteria and plan quality was largely comparable to that of the clinical plans (Supplementary Materials: Figures S1 and S2, Tables S1 and S2). These initial results are encouraging and merit investigation with a larger sample size.

We previously showed the feasibility of using a photon-specific knowledge-based solution for IMPT [15]. The present study incorporates proton-specific software which appropriately characterizes proton behavior. Literature on automated proton planning solutions is relatively sparse and we are currently unaware of similar multi-institutional comparisons. However, Hall et al. used OAR-PTV geometry to predict achievable doses for OARs relevant to the treatment of skull-base tumors in order to facilitate comparison of proton and photon plans [24]. Bijman et al. looked at uncertainties in model-based patient selection for IMRT or IMPT, using automatically planned IMPT plans: the approach to automation differed from this study, in that it used a pre-defined wish-list of hard constraints and hierarchical OAR objectives (tackled in order of priority) [25]. Lomax and colleagues have developed a tool to automatically pre-calculate feasible planning solutions specifically for uveal melanomas [26].

4. Conclusions

This proof-of-principle study has shown that a knowledge-based approach to IMPT planning is feasible. Such a solution could aid both experienced and more recently established proton centers, with a number of potential applications including implementation of consistent, efficient treatment planning; fast (re)planning for adaptive IMPT; and QA of IMPT plans in the clinic and trial setting [27].

5. Materials and Methods

5.1. Treatment Planning for Populating the Knowledge-Based Planning Model

A constant radiobiological effectiveness (RBE) of 1.1 was assumed for IMPT plans, where all doses reported in "Gy" are understood to represent Gy_{RBE} [28]. The 50 non-nasopharynx, locally advanced HNC IMPT plans in the RapidPlanPT model were coplanar and used a simultaneous integrated boost (SIB) technique to deliver 70/54.25 Gy to the boost/elective planning target volume (PTV_B/PTV_E) in 35 fractions. PTV margins were 4–5 mm [29]. A 5 mm transition-region facilitated dose fall-off between PTVs. Plans typically aimed to spare multiple salivary glands, swallowing muscles and the oral cavity [30], although certain OARs could be excluded due to the extent of overlap with PTVs. Maximum point dose objectives were used for spinal cord, brainstem and planning-at-risk volumes (isotropic 3 mm OAR expansion). The aim was to deliver 95% of prescribed dose ($V_{95\%}$) to $\geq 99\%/98\%$ of PTV_B/PTV_E while limiting PTV volume receiving $>107\%$ of prescription dose. IMPT plans were created using the Varian Eclipse non-linear universal proton optimizer (NUPO) and proton convolution superposition algorithm (PCS) v13.7.14 (Varian Medical Systems, Palo Alto, CA, USA) with a 2.5 mm calculation grid. Plans were made with a standard three-field, multi-field optimization (MFO) technique, with gantry angles at 35° – 55° , 180° and 305° – 330° . Gantry variations in these three directions were determined by PTV geometry. A range shifter of 5.7 cm water equivalent thickness was used for proximal PTV irradiation [31]. Each field included proximal, distal and lateral target margins of 0.2 cm, 0.3 cm and 0.5 cm, respectively. A non-robust optimization was performed interactively by manually adjusting optimization objectives to maintain an approximately fixed diagonal distance

to DVH-lines in the optimization-window. In certain cases, a subsequent optimization or “continue optimization” was performed to improve PTV dose homogeneity [32]. A basic refinement of the model was carried out in which regression, residual and DVH-plots, as well as statistical metrics provided by the planning software were used to remove obvious outliers with DVHs a considerable distance above the predicted curve [14,22]. The minimal refinement is reflected in the number of structures matched to each OAR in the model (average: 47, minimum: 37).

5.2. External Center Treatment Planning

Typical planning-aims for PTVs and OARs are shown in Table 2. Center-A defined dose-volume constraints on “OAR minus PTV” structures whilst center-B and -C used the entire OAR. For analysis, all dosimetric data is reported on the entire OAR structure. Up to two elective PTVs, which could differ in prescription dose per patient, were used. Therefore, we refer to the boost PTV as PTV_B, the mid-dose elective PTV as PTV_{E1}, and the low-dose elective PTV as PTV_{E2}. This study intends to demonstrate the feasibility of knowledge-based planning for IMPT, regardless of the beam data used. All reporting is done on the PTVs provided by each center. Although differences in beam data and the choice of proximal/distal spot placement margins may have some effect on the dose distribution and amount of OAR sparing, such effects are expected to be small for state of the art clinical IMPT facilities, and the choice of optimization objectives is likely to have more impact on the OAR sparing. For all these reasons the KBP is considered relatively insensitive to the vendor-specific beamline or beam model.

5.2.1. Center-A

IMPT plans utilized a SIB, delivering 70 Gy, 63/59.5 Gy and 56/54.3 Gy to PTV_B, PTV_{E1} and PTV_{E2} in 35 fractions, respectively. PTV aims were such that dose to 95% of the volume (D95%) was $\geq 100\%$ of the prescribed dose. IMPT plans were created using Varian Eclipse NUPO and PCS v13.7.15 with a 2.5 mm dose calculation grid. Plans utilized a three or five-field MFO technique, including a beam from the anterior direction to irradiate the neck and shoulder nodal regions and two oblique posterior-anterior beams utilized for the head, however, beam angles varied according to tumor location. A range-shifting bolus of 7.5 cm water equivalent thickness was used in all plans. Each field included proximal, distal and lateral target margins of typically 0.5 cm, 0.5 cm and 0.3–0.5 cm, respectively.

5.2.2. Center-B

IMPT plans utilized a SIB, delivering 70 Gy, 63 Gy and 56 Gy to PTV_B, PTV_{E1} and PTV_{E2} in 35 fractions, respectively. The aim was that all PTVs had a V95% $\geq 99\%$. IMPT plans were created using Varian Eclipse NUPO and PCS v13.7.15 with a 2.5 mm dose calculation grid. Plans utilized a three-field MFO with typical gantry angles of 35°, 180° and 325°. A range shifter of 5.7 cm water-equivalent thickness was used for anterior-oblique fields. Typical proximal, distal and lateral target margins were 0.5 cm, 0.5 cm and 1 cm, respectively.

5.2.3. Center-C

IMPT plans utilized a SIB, delivering 70 Gy, 60/58 Gy and 54 Gy to PTV_B, PTV_{E1} and PTV_{E2} in 35 fractions, respectively. PTV D95% aim was $\geq 95\%$ of the prescribed dose. IMPT plans were created using Varian Eclipse NUPO and PCS v13.7.15 with a 2.5 mm dose calculation grid. Plans typically utilized a four-field MFO with varying gantry angles to cover both cranial and caudal portions of a patient. Beam-arrangements were cross shaped in cranial portions to avoid the shoulders. A range shifter of 5 cm water equivalent thickness was used for all fields. Each field included proximal, distal and lateral target margins of 0.8 cm, 0.8 cm and 1.7 cm, respectively.

5.3. Evaluation Patients

Each external center provided 7 HNC, non-robustly optimized, coplanar IMPT benchmark plans. Center-A provided clinical plans while center-B and -C created plans solely for the purpose of this study. Benchmark plans were deemed clinically acceptable by each of the respective external centers. All centers used their own clinical protocol and beam arrangement. For each patient, and without reference to the benchmark plans, RapidPlanPT was used to create non-robust KBPs using NUPO v15.5.99 (facilitates use of line objectives) and PCS v15.0.17 using a standard Y-shaped beam arrangement and the same target margins as used in the model-plans. Dose values of PTV_{E1}/PTV_{E2} optimization objectives were assigned manually, prior to optimization, as the prescription dose varied per center. Optimization objective priorities mirrored those of model library plans. A subsequent optimization or “continue optimization” was performed if (i) PTV homogeneity was not acceptable (priorities increased/objectives added for PTVs) or (ii) KBPs did not meet criteria in Table 2 for OARs not included in the model (objectives added for any non-compliant OAR outside of the modeled structures). External centers could utilize a number of volume constraints for OARs not included in Table 2, such as the brachial plexus. Where necessary, additional manual constraints were used in the KBP planning process to address this.

5.4. Study Endpoints

KBPs were normalized such that $PTV_B D95\% \geq 100\%$, $V95\% \geq 99\%$ and $D95\% \geq 95\%$ for center-A, -B and -C, respectively. Since external centers had differing PTV and OAR planning aims (Table 2), comparisons between benchmark plans and KBPs were performed altogether and per-center, on the basis of (1) target coverage and homogeneity (PTV D95%, PTV V95% and PTV homogeneity index (HI), where $HI = 100\% \times (D2\% - D98\%)/D50\%$) (2) mean dose to OARs. For combined data and data per center, paired 2-sided Student t-tests and Wilcoxon signed rank tests were used to determine whether differences between benchmark plans and KBPs were significant ($p < 0.05$), respectively. The time required to create KBPs was also measured.

5.5. Ethical Statement

All patients signed a consent form, where they agreed to the use of their data for scientific purposes at PSI. The institutional review board (IRB) of the University of Cincinnati Medical Center determined the research proposal does not meet the regulatory criteria for research involving human subjects and thus ongoing IRB oversight was not required. The retrospective planning study was approved at the University of Pennsylvania under IRB # 830036 (1 June 2018).

Supplementary Materials: The following are available online at <http://www.mdpi.com/2072-6694/10/11/420/s1>, Figure S1: Dose-volume histograms (DVHs, solid lines) and respective DVH-predictions (shaded regions) of patient 1 for (LEFT) the non-robust IMPT KBP when using the CTV as a target volume and (RIGHT) the robustly-optimized (4 mm, 3%) IMPT KBP when using the CTV as a target volume, Figure S2: DVH-predictions (shaded regions) and respective DVHs (solid lines) for the robustly optimized (4 mm, 3%) IMPT KBP of patient 1 when using the CTV (▲) and CTV + 2 mm (■) as model targets. Both plans were optimized using the original CTV (Same patient as in Figure S1), Table S1: Robustness evaluation (isocenter shift 4 mm, calibration curve error 3%) and nominal dosimetry for robust IMPT KBPs created using the CTV, CTV + 1 mm (CTV + 1), CTV + 2 mm (CTV + 2), CTV + 3mm (CTV + 3) and CTV + 4 mm (CTV + 4) to generate OAR DVH-predictions, averaged over three validation patients, Table S2: Robustness evaluation (isocenter shift 3 mm, calibration curve error 3.5%) and nominal dosimetry for robust clinical plans and KBPs for three patients from center-A.

Author Contributions: Conceptualization, M.D. and W.F.A.R.V.; Data curation, A.R.D., W.Z., L.Y. and S.R.; Formal analysis, A.R.D., L.D., A.M., W.Z., Y.Z., L.Y., S.R., J.H., A.J.L., M.D. and W.V.; Funding acquisition, M.D. and W.V.; Investigation, A.D., L.D., A.M., W.Z., Y.Z., L.Y., S.R., J.H., A.L., M.D. and W.F.A.R.V.; Methodology, A.R.D., M.D. and W.F.A.R.V.; Project administration, A.R.D., M.D. and W.F.A.R.V.; Resources, B.J.S., M.D. and W.F.A.R.V.; Supervision, M.D. and W.F.A.R.V.; Writing—original draft, A.R.D.; Writing—review & editing, A.R.D., L.D., A.M., W.Z., Y.Z., L.Y., S.R., J.H., A.J.L., B.J.S., M.D. and W.F.A.R.V.

Funding: This research was supported by a grant from Varian Medical Systems. SNF R’Equip grant number 326030_157749 partially supported the contribution of PSI.

Acknowledgments: This research was supported by a grant from Varian Medical Systems.

Conflicts of Interest: The Department of Radiation Oncology, VU University Medical Center, has a research collaboration with Varian Medical Systems and Dahele, Slotman and Verbakel have received honoraria/travel support from Varian medical systems. SNF R'Equip grant number 326030_157749 partially supported the contribution of PSI.

References

1. Particle Therapy Co-Operative Group Particle therapy Facilities under Construction. 2017. Available online: <https://www.ptcog.ch/index.php/facilities-under-construction> (accessed on 17 November 2017).
2. Nelms, B.E.; Robinson, G.; Markham, J.; Velasco, K.; Boyd, S.; Narayan, S.; Wheeler, J.; Sobczak, M.L. Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems. *Pract. Radiat. Oncol.* **2012**, *2*, 296–305. [[CrossRef](#)] [[PubMed](#)]
3. McDonald, M.W.; Liu, Y.; Moore, M.G.; Johnstone, P.A.S. Acute toxicity in comprehensive head and neck radiation for nasopharynx and paranasal sinus cancers: Cohort comparison of 3D conformal proton therapy and intensity modulated radiation therapy. *Radiat. Oncol.* **2016**, *11*, 1–10. [[CrossRef](#)] [[PubMed](#)]
4. Romesser, P.B.; Cahlon, O.; Scher, E.; Zhou, Y.; Berry, S.L.; Rybkin, A.; Sine, K.M.; Tang, S.; Sherman, E.J.; Wong, R.; et al. Proton beam radiation therapy results in significantly reduced toxicity compared with intensity-modulated radiation therapy for head and neck tumors that require ipsilateral radiation. *Radiother. Oncol.* **2016**, *118*, 286–92. [[CrossRef](#)] [[PubMed](#)]
5. Holliday, E.B.; Garden, A.S.; Rosenthal, D.I.; Fuller, C.D.; Morrison, W.H.; Gunn, G.B.; Phan, J.; Beadle, B.M.; Zhu, X.R.; Zhang, X.; et al. Proton therapy reduces treatment-related toxicities for patients with nasopharyngeal cancer: A case-match control study of intensity-modulated proton therapy and intensity-modulated photon therapy. *Int. J. Part. Ther.* **2015**, *2*, 19–28. [[CrossRef](#)]
6. Blanchard, P.; Garden, A.S.; Gunn, G.B.; Rosenthal, D.I.; Morrison, W.H.; Hernandez, M.; Crutison, J.; Lee, J.J.; Ye, R.; Fuller, C.D.; et al. Head and neck radiotherapy Intensity-modulated proton beam therapy (IMPT) versus intensity-modulated photon therapy (IMRT) for patients with oropharynx cancer—A case matched analysis. *Radiother. Oncol.* **2016**, *120*, 48–55. [[CrossRef](#)] [[PubMed](#)]
7. Quan, E.M.; Chang, J.Y.; Liao, Z.; Xia, T.; Yuan, Z.; Liu, H.; Li, X.; Wages, C.A.; Mohan, R.; Zhang, X. Automated volumetric modulated arc therapy treatment planning for stage III lung cancer: How does it compare with intensity-modulated radio therapy? *Int. J. Radiat. Oncol.* **2012**, *84*, e69–e76. [[CrossRef](#)] [[PubMed](#)]
8. Wu, B.; Ricchetti, F.; Sanguineti, G.; Kazhdan, M.; Simari, P.; Jacques, R.; Taylor, R.; McNutt, T. Data-driven approach to generating achievable dose-volume histogram objectives in intensity-modulated radiotherapy planning. *Int. J. Radiat. Oncol. Biol. Phys.* **2011**, *79*, 1241–1247. [[CrossRef](#)] [[PubMed](#)]
9. Wu, B.; Pang, D.; Simari, P.; Taylor, R.; Sanguineti, G.; McNutt, T. Using overlap volume histogram and IMRT plan data to guide and automate VMAT planning: A head-and-neck case study. *Med. Phys.* **2013**, *40*, 021714. [[CrossRef](#)] [[PubMed](#)]
10. Breedveld, S.; Storchi, P.R.M.; Voet, P.W.J.; Heijmen, B.J.M. ICycle: Integrated, multicriterial beam angle, and profile optimization for generation of coplanar and noncoplanar IMRT plans. *Med. Phys.* **2012**, *39*, 951–963. [[CrossRef](#)] [[PubMed](#)]
11. Tol, J.P.; Delaney, A.R.; Dahele, M.; Slotman, B.J.; Verbakel, W.F.A.R. Evaluation of a knowledge-based planning solution for head and neck cancer. *Int. J. Radiat. Oncol. Biol. Phys.* **2015**, *91*, 612–620. [[CrossRef](#)] [[PubMed](#)]
12. Fogliata, A.; Belosi, F.; Clivio, A.; Navarria, P.; Nicolini, G.; Scorsetti, M.; Vanetti, E.; Cozzi, L. On the pre-clinical validation of a commercial model-based optimisation engine: Application to volumetric modulated arc therapy for patients with lung or prostate cancer. *Radiother. Oncol.* **2014**, *113*, 385–391. [[CrossRef](#)] [[PubMed](#)]
13. Tol, J.P.; Dahele, M.; Delaney, A.R.; Slotman, B.J.; Verbakel, W.F.A.R. Can knowledge-based DVH predictions be used for automated, individualized quality assurance of radiotherapy treatment plans? *Radiat. Oncol.* **2015**, *10*. [[CrossRef](#)] [[PubMed](#)]

14. Li, N.; Carmona, R.; Sirak, I.; Kasaova, L.; Followill, D.; Michalski, J.; Bosch, W.; Straube, W.; Mell, L.K.; Moore, K.L. Highly Efficient Training, Refinement, and Validation of a Knowledge-based planning quality-control system for radiation therapy clinical trials. *Int. J. Radiat. Oncol. Biol. Phys.* **2017**, *97*, 164–172. [[CrossRef](#)] [[PubMed](#)]
15. Delaney, A.R.; Dahele, M.; Tol, J.P.; Kuijper, I.T.; Slotman, B.J.; Verbakel, W.F.A.R. Using a knowledge-based planning solution to select patients for proton therapy. *Radiother. Oncol.* **2017**, *124*, 263–270. [[CrossRef](#)] [[PubMed](#)]
16. Fogliata, A.; Wang, P.-M.; Belosi, F.; Clivio, A.; Nicolini, G.; Vanetti, E.; Cozzi, L. Assessment of a model based optimization engine for volumetric modulated arc therapy for patients with advanced hepatocellular cancer. *Radiat. Oncol.* **2014**, *9*, 236. [[CrossRef](#)] [[PubMed](#)]
17. Gu, W.; O'Connor, D.; Nguyen, D.; Yu, V.Y.; Ruan, D.; Sheng, K. Integrated beam angle and scanning spot optimization for intensity modulated proton therapy. *Int. J. Radiat. Oncol.* **2017**, *99*, S107. [[CrossRef](#)]
18. Cao, W.; Lim, G.J.; Li, Y.; Zhu, X.R.; Zhang, X. Improved beam angle arrangement in intensity modulated proton therapy treatment planning for localized prostate cancer. *Cancer* **2015**, *7*, 574–584. [[CrossRef](#)] [[PubMed](#)]
19. Michiels, S.; Barragán, A.M.; Souris, K.; Poels, K.; Crijs, W.; Lee, J.A.; Sterpin, E.; Nuyts, S.; Haustermans, K.; Depuydt, T. Patient-specific bolus for range shifter air gap reduction in intensity-modulated proton therapy of head-and-neck cancer studied with Monte Carlo based plan optimization. *Radiother. Oncol.* **2017**, *128*, 161–166. [[CrossRef](#)] [[PubMed](#)]
20. Widesott, L.; Lomax, A.J.; Schwarz, M. Is there a single spot size and grid for intensity modulated proton therapy? Simulation of head and neck, prostate and mesothelioma cases. *Med. Phys.* **2012**, *39*, 1298–1308. [[CrossRef](#)] [[PubMed](#)]
21. Langner, U.W.; Eley, J.G.; Dong, L.; Langen, K. Comparison of multi-institutional Varian ProBeam pencil beam scanning proton beam commissioning data. *J. Appl. Clin. Med. Phys.* **2017**, *18*, 96–107. [[CrossRef](#)] [[PubMed](#)]
22. Delaney, A.R.; Tol, J.P.; Dahele, M.; Cuijpers, J.; Slotman, B.J.; Verbakel, W.F.A.R. Effect of dosimetric outliers on the performance of a commercial knowledge-based planning solution. *Int. J. Radiat. Oncol. Biol. Phys.* **2016**, *94*, 469–477. [[CrossRef](#)] [[PubMed](#)]
23. Hussein, M.; South, C.P.; Barry, M.A.; Adams, E.J.; Jordan, T.J.; Stewart, A.J.; Nisbet, A. Clinical validation and benchmarking of knowledge-based IMRT and VMAT treatment planning in pelvic anatomy. *Radiother. Oncol.* **2016**, *120*, 473–479. [[CrossRef](#)] [[PubMed](#)]
24. Hall, D.C.; Trofimov, A.V.; Winey, B.; Paganetti, H. A Geometric knowledge-based model to quickly predict the patient-specific benefits of proton therapy in clival chordoma patients. *Int. J. Radiat. Oncol.* **2016**, *96*, S81–S82. [[CrossRef](#)]
25. Bijman, R.G.; Breedveld, S.; Arts, T.; Astreinidou, E.; de Jong, M.A.; Granton, P.V.; Petit, S.F.; Hoogeman, M.S. Impact of model and dose uncertainty on model-based selection of oropharyngeal cancer patients for proton therapy. *Acta Oncol.* **2017**, *56*, 1444–1450. [[CrossRef](#)] [[PubMed](#)]
26. Hennings, F.; Lomax, A.; Pica, A.; Weber, D.C.; Hrbacek, J. Automated treatment planning system for uveal melanomas treated with proton therapy: A proof of concept analysis. *Int. J. Radiat. Oncol.* **2018**, *5*, 1–16. [[CrossRef](#)] [[PubMed](#)]
27. Liao, Z.; Lee, J.J.; Komaki, R.; Gomez, D.R.; O'Reilly, M.S.; Fossella, F.V.; Blumenschein, G.R.; Heymach, J.V.; Vaporciyan, A.A.; Swisher, S.G.; et al. Bayesian adaptive randomization trial of passive scattering proton therapy and intensity-modulated photon radiotherapy for locally advanced non-small-cell lung cancer. *J. Clin. Oncol.* **2018**, *36*, 1813–1822. [[CrossRef](#)] [[PubMed](#)]
28. Paganetti, H.; Niemierko, A.; Ancukiewicz, M.; Gerweck, L.E.; Goitein, M.; Loeffler, J.S.; Suit, H.D. Relative biological effectiveness ({RBE}) values for proton beam therapy. *Int. J. Rad. Oncol. Biol. Phys.* **2002**, *53*, 407–421. [[CrossRef](#)]
29. Tol, J.P.; Doornaert, P.; Witte, B.I.; Dahele, M.; Slotman, B.J.; Verbakel, W.F.A.R. A longitudinal evaluation of improvements in radiotherapy treatment plan quality for head and neck cancer patients. *Radiother. Oncol.* **2016**, *119*, 337–343. [[CrossRef](#)] [[PubMed](#)]
30. Christianen, M.E.M. C.; Langendijk, J.A.; Westerlaan, H.E.; van de Water, T.A.; Bijl, H.P. Delineation of organs at risk involved in swallowing for radiotherapy treatment planning. *Radiother. Oncol.* **2011**, *101*, 394–402. [[CrossRef](#)] [[PubMed](#)]

31. Barten, D.L.J.; Tol, J.P.; Dahele, M.; Slotman, B.J.; Verbakel, W.F. Comparison of organ-at-risk sparing and plan robustness for spot-scanning proton therapy and volumetric modulated arc photon therapy in head-and-neck cancer. *Med. Phys.* **2015**, *42*, 6589–6598. [[CrossRef](#)] [[PubMed](#)]
32. Tol, J.P.; Dahele, M.; Doornaert, P.; Slotman, B.J.; Verbakel, W.F.A.R. Toward optimal organ at risk sparing in complex volumetric modulated arc therapy: An exponential trade-off with target volume dose homogeneity. *Med. Phys.* **2014**, *41*, 021722. [[CrossRef](#)] [[PubMed](#)]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).