HHS Public Access

Author manuscript

Radiother Oncol. Author manuscript; available in PMC 2024 March 26.

Published in final edited form as:

Radiother Oncol. 2023 July ; 184: 109685. doi:10.1016/j.radonc.2023.109685.

Implementation and Evaluation of an Intelligent Automatic Treatment Planning Robot for Prostate Cancer Stereotactic Body Radiation Therapy

Yin Gao, MS[1,2], **Chenyang Shen, PhD**[1,2], **Xun Jia, PhD**[1,2,*], **Yang Kyun Park, PhD**[2]

[1]Innovative Technology Of Radiotherapy Computations and Hardware (iTORCH) Laboratory, Department of Radiation Oncology, University of Texas Southwestern Medical Center, Dallas, TX, USA

[2]Department of Radiation Oncology, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA

Abstract

Purpose: We previously developed a virtual treatment planner (VTP), an artificial intelligence robot, operating a treatment planning system (TPS). Using deep reinforcement learning guided by human knowledge, we trained the VTP to autonomously adjust relevant parameters in treatment plan optimization, similar to a human planner, to generate high-quality plans for prostate cancer stereotactic body radiation therapy (SBRT). This study describes the clinical implementation and evaluation of VTP.

Materials and methods: We integrate VTP with Eclipse TPS using scripting Application Programming Interface. VTP observes dose-volume histograms of relevant structures, decides how to adjust dosimetric constraints, including doses, volumes, and weighting factors, and applies the adjustments to the TPS interface to launch the optimization engine. This process continues until a high-quality plan is achieved. We evaluated VTP's performance using the prostate SBRT case from the 2016 American Association of Medical Dosimetrist/Radiosurgery Society plan study with its plan scoring system, and compared to human-generated plans submitted to the challenge. Using the same scoring system, we also compared the plan quality of 36 prostate SBRT cases (20 planned with IMRT and 16 planned with VMAT) treated at our institution for both VTP and human-generated plans.

Results: In the plan study case, VTP achieved a score of 142.1/150.0, ranking the third in the competition (median 134.6). For the clinical cases, VTP achieved 110.6±6.5 for 20 IMRT plans and 126.2±4.7 for 16 VMAT plans, similar to scores of human-generated plans with 110.4±7.0 for IMRT plans and 125.4±4.4 for VMAT plans. The workflow, plan quality and planning time of VTP were reviewed to be satisfactory by experienced physicists.

Corresponding Authors: Yangkyun.Park@UTSouthwestern.edu (Yang Kyun, Park), Xun.Jia@UTSouthwestern.edu, xunjia@jhu.edu (Xun Jia).

*Currently at Department of Radiation Oncology and Molecular Radiation Sciences, Johns Hopkins University, Baltimore, MD 21287, USA

Author Responsible for Statistical Analysis: Yin.Gao@UTSouthwestern.edu (Yin Gao)

Conclusion: We successfully implemented VTP to operate a TPS for autonomous human-like treatment planning for prostate SBRT.

1. Introduction

Modern radiation therapy uses advanced delivery techniques, such as Intensity-modulated RT (IMRT) and Volumetric-modulated RT (VMAT) to precisely control a medical linear accelerator (LINAC) to generate a carefully sculptured 3D dose distribution. New technologies successfully maintain highly conformal dose to tumor target shape but with significantly reduced toxicity to preserve organs at risk (OARs)[1]. The success of these novel techniques cannot be achieved without successful patient-specific treatment planning[2,3]. In the current clinical practice, the task of IMRT and VMAT treatment planning is accomplished by a human planner operating a commercial treatment planning system (TPS). Specifically, the planner sets up initial treatment planning objectives for different structures to mathematically define an optimization problem. The optimization engine in TPS is then launched to solve the defined problem and generate a plan. Due to the inherent complexity of finding the optimal solution for this multi-criteria optimization problem, the human planner is required to repeatedly interact with the TPS to adjust the values of treatment planning parameters (TPPs), e.g., dose limit, volume constraint, and priority, in a trial-and-error fashion to steer the plan quality towards desired trade-offs. This iterative process is time-consuming and labor-intensive, and the resulting plan quality is highly dependent on the planner's experience and available planning time[4,5]. Hence, there is a strong desire to develop automated treatment planning approaches that can generate consistent high-quality plans.

Over the years, researchers have devoted tremendous efforts to solve the automated treatment planning problem using various approaches, such as greedy algorithms[6,7,8,9], heuristic approaches[10,11], fuzzy inference[12,13,14], statistics-based methods[15,16], knowledge-based planning[17,18], and multi-criteria optimization[19,20,21]. Deep-learning based methods[22,23,24], particularly reinforcement learning (RL)-based ones have demonstrated great potential[25]. With recent breakthrough of RL, particularly in the deep learning regime, called deep reinforcement learning (DRL), approaches employing this technique show great potential in automated treatment planning[26,27,28,29,30]. A recent study has developed a DRL-based machine parameter optimization approach to directly control LINAC and rapidly generate VMAT plans for prostate cancer[26].

The major challenge in automatic treatment planning is that TPSs can only solve an optimization problem but lack of human-level intelligence to proactively adjust TPPs to improve plan quality. To overcome this problem, we previously proposed an Intelligent Automatic Treatment Planning (IATP) framework[27,28,29,30,31]. In this framework, a virtual treatment planner (VTP), an AI planning robot, was built via end-to-end DRL training, which can interact with a TPS and adjust TPPs during treatment planning for high-quality prostate SBRT plans. Due to the low efficiency of having VTP interacting with Eclipse TPS (Varian Medical Systems, Palo Alto, CA) directly in the training phase, we trained VTP with an in-house TPS which had a similar inverse planning optimization engine to the one in Eclipse TPS. Studies have demonstrated preliminary success in terms of decision making

and autonomously producing high-quality plans of high-dose-rate brachytherapy for cervical cancer [27] and IMRT for prostate cancer[28,29,30].

Based on the previous success, it is of importance to clinically implement our VTP to operate a commercial TPS and evaluate its performance. Hence, in this study, we will report our recent progress designing a workflow to integrate VTP with Eclipse TPS. We will also evaluate the planning performance of VTP by comparing with human-generated clinical plans for prostate SBRT. To our knowledge, this is the first study demonstrating that a DRL-based AI planner can autonomously make decisions like a human to intelligently accomplish SBRT treatment planning in a real-world clinical setting.

2. Materials and Methods

2.1. Overview of VTP

VTP was previously developed to interact with an in-house developed Eclipse-like TPS and adjust TPPs of 13 planning objectives to generate prostate SBRT plans using IMRT. Each planning objective had a set of three TPPs to adjust, i.e., priorities, dose limits, and volume constraints. VTP consisted of three sub-networks: Structure-Net, Parameter-Net, and Action-Net. It took DVHs of 13 relevant structures as input. The three sub-networks were applied sequentially to determine the way of TPP adjustment to improve plan quality. Specifically, Structure-Net first selected a structure, Parameter-Net then chose one TPP of the selected structure, and Action-Net decided on the specific adjustment action of the chosen TPP. Such an architecture was designed to tackle the treatment planning problem via a hierarchical decision-making process like human planners.

Training VTP was achieved via DRL with a Q-learning framework. The purpose of DRL is to establish the relationship between an observed DVH d and the corresponding optimal way of adjusting TPPs, denoted as C, so that repeated applications of TPP adjustments to the TPS lead to a high plan quality as measured by a score function. In the Q-learning framework, this was achieved by establishing the VTP networks as a representation of the so-called optimal action-value function function $Q^*(d, C)$. Once this function was obtained, the proper TPP adjustment C corresponding to the observed DVH d can be inferred as $C = \text{argmax}_{C'} Q^*(d, C')$, i.e., selecting the one that maximizes the $Q^*(d, C)$ function.

To decide $Q^*(d, C)$, we minimized a loss function corresponding to the L2 norm of the residual for the Bellman equation[32], a general property satisfied by the optimal action-value function: $| r + \gamma \max_{a'} Q^*(d', C') - Q^*(d, C) |^2$. Here r is the reward after applying TPP adjustment C to the current plan with DVH d, and d' is the new plan after launching plan optimization with the adjusted TPPs. $\gamma \in [0, 1]$ is a discount factor that determines the importance of future rewards in the learning process. In this study, we incorporated the Proknow scoring system (ProKnow Systems, Sanford, FL, USA) used in the 2016 AAMD/RSS prostate SBRT plan study[33] as the plan quality evaluation criteria, and the improvement in plan quality before and after taking an TPP adjustment as the reward function.

We retrospectively collected 20 prostate cancer patients who were treated with 45 Gy in 5 fraction SBRT at our institution. We randomly selected 10 patients for training, 5 for validation and the remaining 5 for testing. In the training process, we let the VTP networks to interact with the in-house TPS with an ϵ-greedy scheme to generate training data, each being a quadruple $\{d, C, r, d'\}$. The data were than used to minimize the loss function and hence determine $Q^*(d, C)$ with an experience replay strategy. Detailed network structure and training strategies can be found in the previous study[30].

2.2. Interfacing VTP with Eclipse environment

To evaluate VTP in the real clinical environment, a graphical user interface (GUI) was developed to integrate VTP with the clinical Eclipse TPS using its Eclipse Scripting Application Programing Interface (ESAPI). VTP was packaged as a single executable file, so that it can be programmatically called by an ESAPI script with custom parameters. We enabled functions such as plan setup, optimization, dose calculation, data import and export for VTP by the ESAPI script to automate the entire process. We also allowed remote execution of VTP on a GPU cluster to expedite the calculation and minimize the computational burden in local clinical workstations. ProKnow scoring system was integrated in the GUI to automatically report the up-to-date score at each planning iteration. A shared folder was created as the communication channel between VTP and Eclipse to store inputs and outputs for VTP and Eclipse TPS.

Figure 1 shows the schematic diagram of the Eclipse-based VTP workflow. Note that the workflow is fully automated with the options for human planners to intervene in some steps. Specifically, after delineating target and OAR volumes, a human planner can execute ESAPI scripts to call the GUI. The GUI offered options to prepare initial plan setup automatically following the integrated plan template or manually based on their preferences. Plan setup included the selection of isocenter position, dose, fraction, machine parameters (machine, energy, dose rate), fluence model (with or without flattening filter free), field numbers, collimator angle, gantry angle, technique (IMRT or VMAT), and jaw position. After this, the GUI automatically launched Eclipse TPS to start plan optimization and dose calculation. After a plan was generated, DVHs and TPPs were automatically exported into the shared folder and VTP acquired the plan data from the folder and decided an action to adjust planning parameters (dose, volume, and priority of relevant planning structures). VTP stored the adjusted TPPs into the shared folder for Eclipse TPS to pick up and start a new iteration to optimize the plan. The iterative planning process continued until a high-quality plan was achieved, as indicated by the convergence of plan scores. The final plan can be fine-tuned by human planners if deemed necessary.

2.3. Performance evaluation

To evaluate the performance of VTP-enabled treatment planning workflow with Eclipsed TPS, we used it to plan the prostate SBRT case of the 2016 AAMD/RSS Plan Study and compared the final plan score with submitted human-generated plans. ProKnow scoring system quantitatively assesses the plan quality in 15 clinical criteria including

$V_{PTV}[Rx](\%), V_{Prostate}[Rx](\%), D_{PTV}[PTV-0.03\,cc](Gy), Conformation\ Number_{PTV}, V_{Rectum}[36\,Gy].(cc), V_{Bladder}[37\,Gy](cc), D_{Rectum}[40\,\%](Gy), D_{Urethra}[20\,\%](Gy), D_{Bowel}[1\,cc](Gy), D_{Penile\,Bulb}[0.1\,cc](Gy), D_{Neurovascular\,Bundles}[50\,\%](Gy), D_{Right\,Femoral\,Head}[max](Gy), D_{Left\,Femoral\,Head}[max](Gy), D_{Skin}[max](Gy), D_{Testes}[max](Gy)$

The plan can be scored between 0 and 150, with a higher score indicating better plan quality.

Under the approval of the Institutional Review Board of UT Southwestern Medical Center (IRB# STU 082013-008), this study selected a retrospective cohort of 36 males previously treated with prostate SBRT using VMAT between 2017 and 2020 from our clinical database. The age of the patient cohort ranged in between 53 and 88, with a median of 68 years old. The median PTV volume was 68 cm^3 and range was [37, 135] cm^3 . SBRT plans were designed to deliver 4500 cGy in five fractions and all plans normalized to achieve the minimal PTV coverage ($V_{PTV}[Rx](\%) = 95$). To evaluate VTP, which was trained for IMRT treatment planning, 20 cases were re-planned by a dosimetrist (3-year experience). The plans were compared with those generated by VTP. Plan scores and dosimetric metrics were analyzed using non-inferiority tests with a significance level of 0.05. More importantly, we demonstrated VTP's decision-making behaviors by comparing the initial, intermediate, and final plan quality and showing the improvements made by VTP during the planning process. Moreover, although VTP was trained based on IMRT optimization, as it takes DVHs as input to improve plan quality, it can be theoretically applied to different planning techniques. To study the feasibility and generalizability of VTP on VMAT planning, we used it to plan the remaining 16 patients with VMAT and compared with their clinical plans.

3. Results

Our VTP successfully generated a high-quality IMRT plan for the 2016 AAMD/RSS Plan Study case using Eclipse TPS. Fig. 2(a) shows dose distributions in three axial slices, and coronal and sagittal views. Fig. 2(b) illustrates the evolution of scores in the planning process as VTP improved plan score from 115.3 to 142.1 out of 150 in 9 iterations. Fig. 2(c) presents the distribution of human-generated IMRT plans submitted to the planning competition. VTP could have been placed at the 3rd out of all the submitted IMRT plans. Table 1 shows the dosimetric evaluation and scores of the VTP-generated plan.

To compare the VTP-generated and human-generated IMRT plans for the 20 prostate SBRT cases, we report ProKnow scores in Fig. 3(a) and dosimetric metrics in Fig. 3(b). VTP achieved a slightly higher average plan score 110.6±6.5 compared with clinical plan 110.4±7.0 (p=0.00), indicating VTP-generated plans were not inferior to the human-generated plans. Plans generated by VTP and human presented comparable coverage and conformity as $V_{PTV}[45\,Gy]$, $V_{Prostate}[45\,Gy]$, $D_{PTV}[PTV-0.03cc]$ and CI_{PTV}. Non-inferiority test indicated the performance of VTP was not inferior to human on dosimetric metrics including $V_{Rectum}[36\,Gy](cc)$, $D_{Rectum}[40\,\%]$, $D_{Urethra}[20\,\%]$, $D_{Bowel}[1cc]$, and $D_{Skin}[max]$ as p<0.05. For other metrics, $V_{Bladder}[37\,Gy]$, $D_{Penile\,Bulb}[0.1cc]$, $D_{Bowel}[1cc]$, $D_{Right\,Femoral\,Head}[max]$, and $D_{Left\,Femoral\,Head}[max]$, the performance of VTP may be inferior to human as the p>0.05. VTP-generated plans had less mean monitor units (MUs) 3859.4±445.3 than human-generated plans 3942.6±541.0 (p=0.02), indicating the complexity of VTP-generated plans was slightly lower thus an advantage in treatment delivery efficiency. Generally, VTP has comparable planning performance to human planners.

To demonstrate VTP's decision-making behavior, Fig. 4(a) shows the improvements on the plan quality during the planning process for an example case. The plan quality scores were 105.2, 121.9, and 127.6 out of 150, at iteration step 1, 5, and 13, respectively. Iteration 13 gave the best-quality plan with more dose sparing in the urethra and rectum, as pointed by arrows 1 and 3. Dose fall-off was steeper in the inferior region of bladder and the superior lateral region of PTV, as pointed by 2 and 4. Fig. 4(b) presents DVHs of plans at iteration 1 (dot), 5 (square), and 13 (triangle). VTP was able to maintain PTV coverage while effectively increasing the minimum dose to PTV and reducing the dose to OARs. Significant improvements in PTV minimum coverage, max dose to the urethra, high dose in rectum, and max dose to femurs were indicated by 5, 6, 7, and 8, respectively. The VTP-centered treatment planning workflow, plan quality, and decision-making behaviors were reviewed to be satisfactory based on the clinical feedback from dosimetrists and physicists.

To demonstrate the feasibility and generalizability of VTP on different planning techniques, we evaluated VTP on additional 16 prostate SBRT cases. ProKnow scores and dosimetric metrics of the generated plans and their clinically accepted plans are summarized in Fig. 5(a) and (b). Similar to the findings indicated in the compressions for IMRT planning, VTP showed comparable planning performance to human planners with slightly higher averaged ProKonw score (VTP: 126.2±4.7, Human: 125.5±4.4, p=0.01). Fig. 5(c) compares cross-sectional 3D dose distributions and plan scores between the VTP-generated and human-generated plans of one example patient. VTP spared more bladder, rectum, urethra, bowel, and femoral heads. As shown in Fig. 5(c), VTP-generated plan presented better dose fall-off laterally while human-generated plan showed better fall-off anteriorly and posteriorly. VTP spared more dose to urethra but human controlled low dose spread-out in rectum, pointed by 1 and 2, respectively. The VTP-generated plan required 380 more MUs, indicating it was more complex. Overall, both plans present comparable dose distributions for VMAT, proving the feasibility.

4. Discussion

This study described the clinical implementation of an in-house planning robot VTP, as a part of the ongoing development on the IATP framework[27,28,29,30], which was built to overcome the challenge that TPSs often lack human-level intelligence to evaluate plan quality and to steer the optimization process towards a high-quality plan. IATP aimed at employing DRL techniques to achieve an automated treatment planning workflow similar to the current human-centered workflow. To our knowledge, the current study is the first one on implementing a DRL-based AI planner in a clinical TPS and evaluating its performance. Previously, Zhang el al. developed a RL-based planning bot with linear action-value function approximation for optimal action determination for pancreas SBRT using Eclipse TPS[25]. In addition to the use of the "deep" version of reinforcement learning, our study differed from two other aspects. First, the previous study only added or removed one fixed dose-volume constraint with the pre-defined constant priority at each step, because the linear function approximation limited the capability for more complicated tasks. By contrast, with the flexibility offered by deep neural network-based function approximation, our VTP can determine the optimal adjustment on dose, volume and priority in a planning objective. The interpretation about the planning process (Sec 3.3) found that its actions to adjust dose-

volume constraints generally agreed with a human planner's intuition[30]. Second, Zhang el al. carefully selected fixed types of constraints for the planning bot to add/remove, ahead of training. Yet, using the same type of dose-volume constraints with constant priorities for all patients may limit the space of optimal solutions and bias the actions. In this study, our VTP had the flexibility to adjust all TPPs that the TPS exposed to users, and the training process learnt the optimal policy of TPP adjustment via an end-to-end DRL training scheme. This potentially allowed VTP to discover unseen action-value space to gain knowledge without being biased to the known experience.

Computation time is one factor affecting the practicality of a planning tool. Currently, VTP makes TPP adjustments within 30 seconds of each iteration, and it takes ~10-15 iterations to finish the planning of case. Eclipse TPS costs about 5 minutes to solve IMRT optimization per iteration, making the total time 1 to 2 hours. The relatively long planning time were ascribed to two factors, calling for future improvements. First, majority of time was spent on plan optimization and dose calculation. A human planner could pause the optimization and bypass the final dose calculation step, before making a decision on TPP adjustment. However, for our system, the Eclipse interface did not allow us to interfere the optimization and final dose calculation step. Hence, VTP had to wait this to finish to adjust a TPP. Second, the plan quality was improved, but slowly at the last a few iterations. This was mainly caused by the fact that VTP was trained to change one TPP at a time, and hence the impact of changing plan quality at each step was relatively small. In contrast, a human planner may decide to simultaneously modify multiple TPPs in one step to improve the plan from different perspectives. It is our future work to further improve VTP to enable this simultaneous decision-making capability.

A potential application of the VTP clinical workflow is adaptive radiation therapy (ART). Although online ART treatment planning has been driven by many AI models, offline ART planning is still achieved by human planners to manually generate new plans. Not only does this increase the workload, but also potentially results in treatment delay, thus deteriorating treatment outcomes. With the full clinical implementation, a high-quality adaptive plan may be quickly generated by VTP in an automated fashion. The human planner can refine the plan for better quality if needed. Another area that VTP can lead to clinical impacts is at large-scale clinical trial studies investigating effectiveness of radiotherapeutic approaches. Inter- and intra- institutional plan quality variation[34,35] is one factor affecting outcomes of those trial studies. Using VTP to generate treatment plans is expected to reduce quality variations related to planner factors. Note that the plans are actually produced by a commercial TPS under VTP operation, and hence deliverability of the plans is not of a major concern. VTP may also be used as a quality assurance tool for clinical trials to ensure plan quality consistency.

Our study has the following major limitations. First, VTP can only change TPPs for the planning objectives of 13 pre-defined structures but not proactively creating auxiliary structures to guide optimization, which is a common skill of human planners to sculpture dose distribution. Although VTP has achieved similar planning performance as human for prostate SBRT, VTP may struggle with more challenging tumor sites, such as head-and-neck, as it involves larger treatment regions and more overlapping with OARs. One simple

solution may be creating optimization structures manually in advance and including them in the VTP model training along with other planning structures, following the same framework in this study. However, we do hope to build a truly intelligent VTP that can determine the shape and location of optimization structures as needed to ensure the flexibility of using VTP in different scenarios.

Second, the current VTP was trained to pursue a high plan quality in the form of ProKnow score. We made this choice to allow benchmarking VTP in the AAMD/RSS case. However, acceptability of a plan is determined by the attending physician, who may evaluate plan quality with more criteria than the pre-defined metrics in the ProKnow system. A recent study has successfully developed a deep learning-based Virtual Physician Network (VPN). Using adversarial learning based on clinical-approved plans, VPN was trained to model the physicians' preference on plan approval for prostate SBRT[31]. Based on an input plan seeking for evaluation, VPN outputs a probability for the plan to be accepted and highlights areas require improvements, should the plan be rejected. We expect the plan approval probability output by VPN can serve as the reward function to guide the VTP's training and planning process. It is our ongoing work to connect VTP and VPN and jointly train them in an end-to-end process to emulate the physician-planner partnership to cooperatively accomplish treatment planning task with generated plans acceptable to the physicians.

Third, the current VTP was trained to only adjust TPPs in the inverse planning optimization. There are other parameters that critically affect resulting plan quality. For instance, in IMRT, beam angles are known to affect dose distribution[36]. Collimator angle may play a role when shaping dose fall-offs around the target to spare dose to OARs. The current study is only the initial step towards building the IATP framework, whereas extensive subsequent studies are ahead to improve the intelligence level of the virtual planner.

5. Conclusions

We successfully implemented VTP to operate a commercial TPS for autonomous human-like treatment planning for prostate SBRT in the clinical environment. The fully automated workflow and outstanding planning performance demonstrated feasibility and effectiveness of the IATP framework.

Disclosures:

The authors have no conflicts of interests to report. This work is supported by a National Institutes of Health grants R37CA214639, R01CA237269 and R01CA254377.

Data sharing statement:

Research data are not available at this time.

References

1. Marur S, Forastiere AA. Head and neck squamous cell carcinoma: Update on epidemiology, diagnosis, and treatment. Mayo Clinic Proceedings. 2016;91(3):386–396. doi:10.1016/j.mayocp.2015.12.017 [PubMed: 26944243]

2. Oelfke U, Bortfeld T. Inverse planning for photon and proton beams. Medical Dosimetry. 2001;26(2):113–124. doi:10.1016/s0958-3947(01)00057-7 [PubMed: 11444513]

3. Webb S. The physical basis of IMRT and inverse planning. The British Journal of Radiology. 2003;76(910):678–689. doi:10.1259/bjr/65676879 [PubMed: 14512327]

4. Nelms BE, Robinson G, Markham J, et al. Variation in external beam treatment plan quality: An inter-institutional study of planners and Planning Systems. Practical Radiation Oncology. 2012;2(4):296–305. doi:10.1016/j.prro.2011.11.012 [PubMed: 24674168]

5. Das IJ, Cheng C-W, Chopra KL, Mitra RK, Srivastava SP, Glatstein E. Intensity-modulated radiation therapy dose prescription, recording, and delivery: Patterns of variability among institutions and Treatment Planning Systems. JNCI: Journal of the National Cancer Institute. 2008;100(5):300–307. doi:10.1093/jnci/djn020 [PubMed: 18314476]

6. Xing L, Li JG, Donaldson S, Le QT, Boyer AL. Optimization of importance factors in inverse planning. Physics in Medicine and Biology. 1999;44(10):2525–2536. doi:10.1088/0031-9155/44/10/311 [PubMed: 10533926]

7. Wu X, Zhu Y. An optimization method for importance factors and beam weights based on genetic algorithms for radiotherapy treatment planning. Physics in Medicine and Biology. 2001;46(4):1085–1099. doi:10.1088/0031-9155/46/4/313 [PubMed: 11324953]

8. Lu R, Radke RJ, Happersett L, et al. Reduced-order parameter optimization for simplifying prostate IMRT planning. Physics in Medicine and Biology. 2007;52(3):849–870. doi:10.1088/0031-9155/52/3/022 [PubMed: 17228125]

9. Wang H, Dong P, Liu H, Xing L. Development of an autonomous treatment planning strategy for radiation therapy with effective use of population-based prior data. Medical Physics. 2017;44(2):389–396. doi:10.1002/mp.12058 [PubMed: 28133746]

10. Yan H, Yin FF. Application of distance transformation on parameter optimization of inverse planning in intensity-modulated radiation therapy. Journal of Applied Clinical Medical Physics. 2008;9(2):30–45. doi:10.1120/jacmp.v9i2.2750 [PubMed: 18714279]

11. Wahl N, Bangert M, Kamerling CP, et al. Physically constrained voxel-based penalty adaptation for ultra-fast IMRT planning. Journal of Applied Clinical Medical Physics. 2016;17(4):172–189. doi:10.1120/jacmp.v17i4.6117 [PubMed: 27455484]

12. Yan H, Yin FF, Guan H, Kim JH. Fuzzy logic guided inverse treatment planning. Medical Physics. 2003;30(10):2675–2685. doi:10.1118/1.1600739 [PubMed: 14596304]

13. Holdsworth C, Kim M, Liao J, Phillips MH. A hierarchical evolutionary algorithm for multiobjective optimization in IMRT. Medical Physics. 2010;37(9):4986–4997. doi:10.1118/1.3478276 [PubMed: 20964218]

14. Holdsworth C, Kim M, Liao J, Phillips M. The use of a multiobjective evolutionary algorithm to increase flexibility in the search for better IMRT plans. Medical Physics. 2012;39(4):2261–2274. doi:10.1118/1.3697535 [PubMed: 22482647]

15. Lee T, Hammad M, Chan TCY, Craig T, Sharpe MB. Predicting objective function weights from patient anatomy in prostate IMRT treatment planning. Medical Physics. 2013;40(12):121706. doi:10.1118/1.4828841 [PubMed: 24320492]

16. Boutilier JJ, Lee T, Craig T, Sharpe MB, Chan TCY. Models for predicting objective function weights in prostate cancer IMRT. Medical Physics. 2015;42(4):1586–1595. doi:10.1118/1.4914140 [PubMed: 25832049]

17. Ge Y, Wu QJ. Knowledge-based planning for intensity-modulated radiation therapy: A review of data-driven approaches. Medical Physics. 2019;46(6):2760–2775. doi:10.1002/mp.13526 [PubMed: 30963580]

18. Li X, Zhang J, Sheng Y, et al. Automatic IMRT planning via Static Field Fluence Prediction (AIP-SFFP): A deep learning algorithm for real-time prostate treatment planning. Physics in Medicine & Biology. 2020;65(17):175014. doi:10.1088/1361-6560/aba5eb [PubMed: 32663813]

19. Craft DL, Hong TS, Shih HA, Bortfeld TR. Improved planning time and plan quality through multicriteria optimization for intensity-modulated radiotherapy. International Journal of Radiation Oncology*Biology*Physics. 2012;82(1). doi:10.1016/j.ijrobp.2010.12.007

20. Biston M-C, Costea M, Gassa F, et al. Evaluation of fully automated a priori MCO treatment planning in VMAT for head-and-neck cancer. Physica Medica. 2021;87:31–38. doi:10.1016/j.ejmp.2021.05.037 [PubMed: 34116315]
21. Hong TS, Craft DL, Carlsson F, Bortfeld TR. Multicriteria optimization in intensity-modulated radiation therapy treatment planning for locally advanced cancer of the pancreatic head. International Journal of Radiation Oncology*Biology*Physics. 2008;72(4):1208–1214. doi:10.1016/j.ijrobp.2008.07.015
22. Nguyen D, Jia X, Sher D, et al. 3D radiotherapy dose prediction on head and neck cancer patients with a hierarchically densely connected U-net deep learning architecture. Physics in Medicine & Biology. 2019;64(6):065020. doi:10.1088/1361-6560/ab039b [PubMed: 30703760]
23. Shen C, Nguyen D, Zhou Z, Jiang SB, Dong B, Jia X. An introduction to deep learning in medical physics: advantages, potential, and challenges. Physics in Medicine & Biology. 2020;65(5):05TR01. doi:10.1088/1361-6560/ab6f51
24. Nguyen D, McBeth R, Sadeghnejad Barkousaraie A, et al. Incorporating human and learned domain knowledge into training deep neural networks: A differentiable dose-volume histogram and adversarial inspired framework for generating Pareto optimal dose distributions in radiation therapy. Medical Physics. 2019;47(3):837–849. doi:10.1002/mp.13955 [PubMed: 31821577]
25. Zhang J, Wang C, Sheng Y, et al. An interpretable planning bot for pancreas stereotactic body radiation therapy. International Journal of Radiation Oncology*Biology*Physics. 2021;109(4):1076–1085. doi:10.1016/j.ijrobp.2020.10.019
26. Hrinivich WT, Lee J. Artificial intelligence-based radiotherapy machine parameter optimization using reinforcement learning. Medical Physics. 2020;47(12):6140–6150. doi:10.1002/mp.14544 [PubMed: 33070336]
27. Shen C, Gonzalez Y, Klages P, et al. Intelligent inverse treatment planning via deep reinforcement learning, a proof-of-principle study in high dose-rate brachytherapy for cervical cancer. Physics in Medicine and Biology. 2019;64(11):115013. Published 2019 May 29. doi:10.1088/1361-6560/ab18bf [PubMed: 30978709]
28. Shen C, Nguyen D, Chen L, et al. Operating a treatment planning system using a deep-reinforcement learning-based virtual treatment planner for prostate cancer intensity-modulated radiation therapy treatment planning. Medical Physics. 2020;47(6):2329–2336. doi:10.1002/mp.14114 [PubMed: 32141086]
29. Shen C, Chen L, Gonzalez Y, Jia X. Improving efficiency of training a virtual treatment planner network via knowledge-guided deep reinforcement learning for intelligent automatic treatment planning of radiotherapy. Medical Physics. 2021;48(4):1909–1920. doi:10.1002/mp.14712 [PubMed: 33432646]
30. Shen C, Chen L, Jia X. A hierarchical deep reinforcement learning framework for intelligent automatic treatment planning of prostate cancer intensity modulated radiation therapy. Physics in Medicine & Biology. 2021;66(13):134002. doi:10.1088/1361-6560/ac09a2
31. Gao Y, Shen C, Gonzalez Y, Jia X. Modeling physician's preference in treatment plan approval of stereotactic body radiation therapy of prostate cancer. Physics in Medicine and Biology. 2022;67(11):115012. doi:10.1088/1361-6560/ac6d9e
32. Bellman R, Karush R. Dynamic Programming a Bibliography of Theory and Application. Santa Monica, Calif: The Rand Corp; 1964.
33. Nelms B. Published June 2016. https://blog.proknowsystems.com/wp-content/uploads/2016/06/2016-AAMD-RSS-Plan-Study.pdf
34. Berry SL, Boczkowski A, Ma R, Mechalakos J, Hunt M. Interobserver variability in radiation therapy plan output: Results of a single-institution study. Pract Radiat Oncol. 2016;6(6):442–449. doi:10.1016/j.prro.2016.04.005 [PubMed: 27374191]
35. Williams MJ, Bailey M, Forstner D, Metcalfe PE. Multicentre quality assurance of intensity-modulated radiation therapy plans: a precursor to clinical trials. Australas Radiol. 2007;51(5):472–479. doi:10.1111/j.1440-1673.2007.01873.x [PubMed: 17803801]
36. Jia X, Men C, Lou Y, Jiang SB. Beam orientation optimization for intensity modulated radiation therapy using adaptive l(2,1)-minimization. Physics in Medicine and Biology. 2011;56(19):6205–6222. doi:10.1088/0031-9155/56/19/004 [PubMed: 21891848]

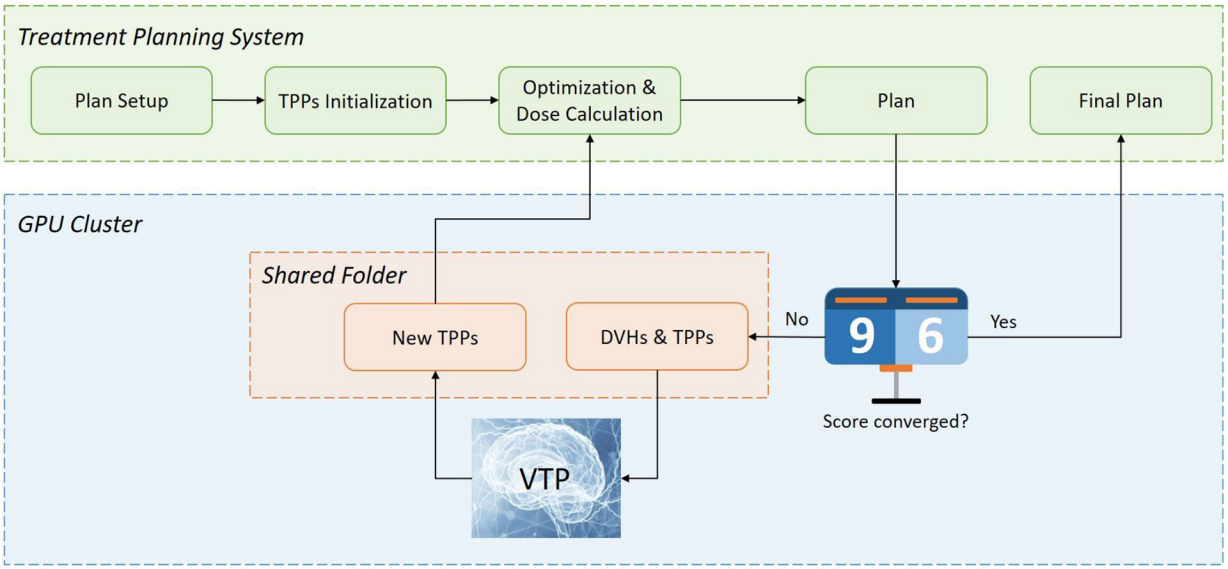

Figure 1.
Fully automated workflow of VTP to operate Eclipse TPS.

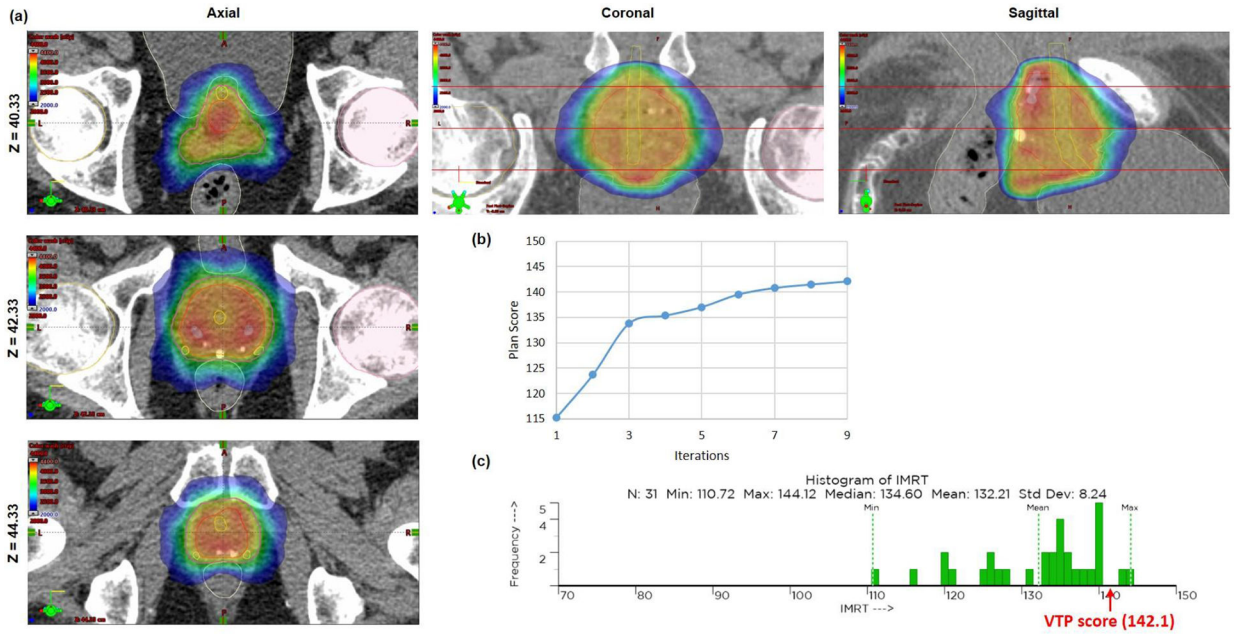

Figure 2.

(a) Dose distribution of the VTP generated plan for the 2016 AAMD/RSS Plan Study case. Horizontal lines in the coronal and sagittal views indicating locations of the axial slides. Dose color wash is shown between 44 Gy and 20 Gy. (b) Evolution of plan scores in the planning process of VTP. (c) Score distribution of human-generated IMRT plans submitted to the 2016 AAMD/RSS prostate SBRT plan study. The score of the VTP-generated plan ranked the third, as pointed by the arrow.

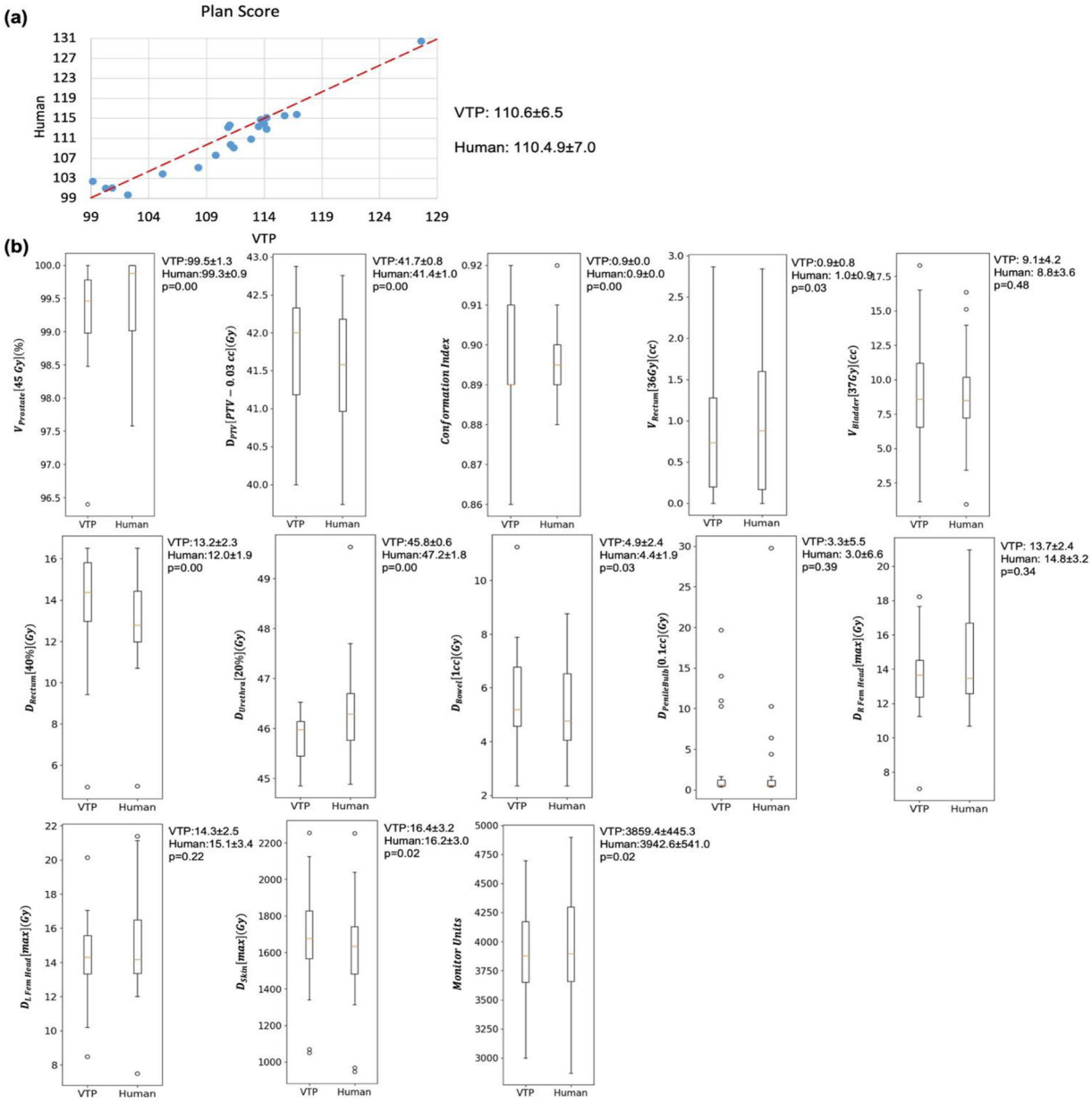

Figure 3.

(a) Scores of IMRT plans generated by VTP and human planner for 20 prostate SBRT patients. Dash line indicates equal plan quality. (b) Boxplots of dosimetric metrics and MUs to compare VTP-generated with human-generated plans of the 20 cases. The edges of boxes are the upper and lower quartiles. The horizontal lines inside boxes represent the median values. Mean ± standard deviation and *p*-value of non-inferior testing for each metric are calculated and displayed in the right upper corner of the boxplot.

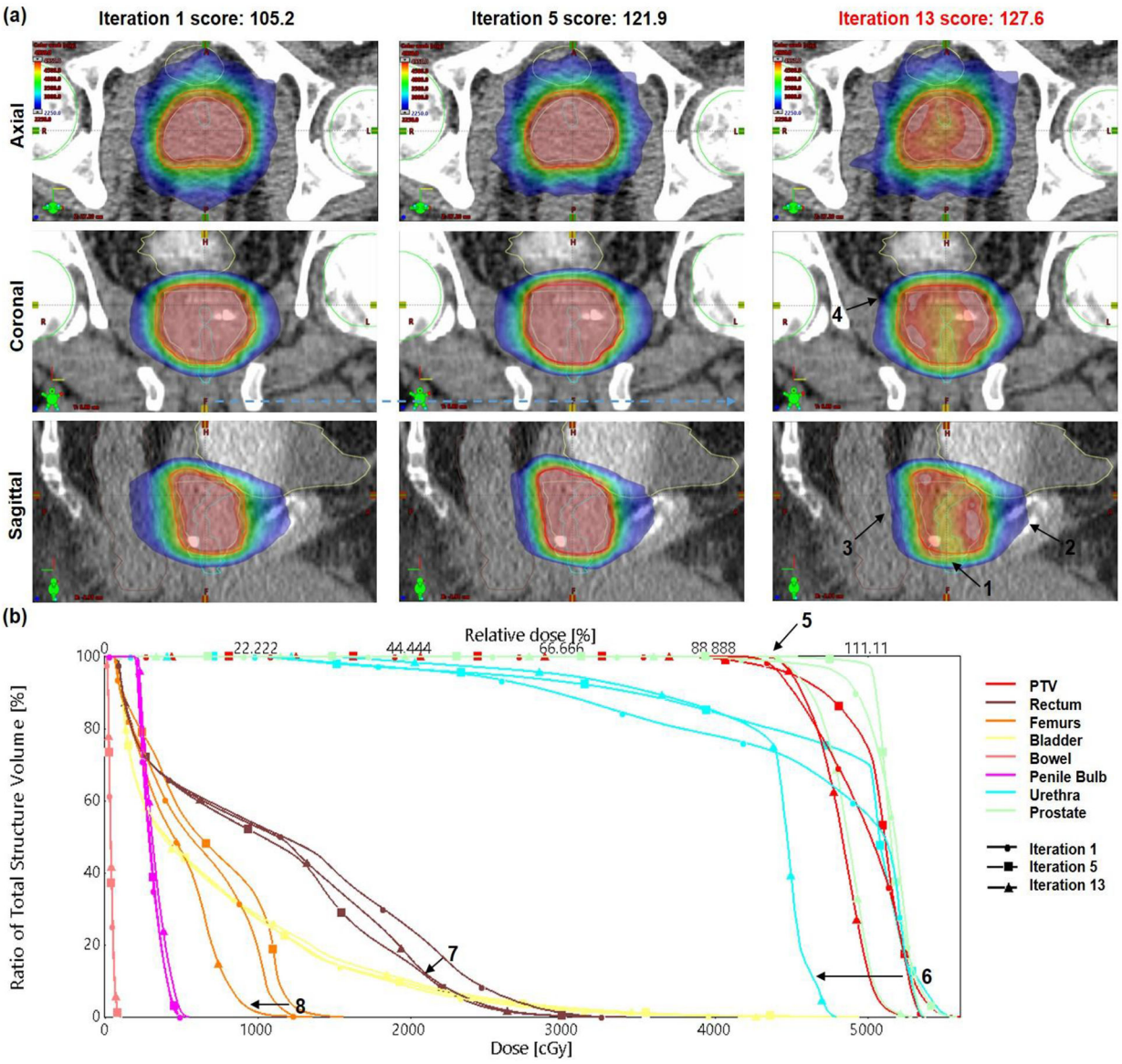

Figure 4.

(a) Evolution of cross-sectional dose distributions and plan scores of plans at iteration 1, 5, and 13 generated by VTP. Dose improvements on the high dose to urethra and low dose spread out in rectum are pointed by 1 and 3. Steeper dose gradient is highlighted by 2 and 4. Dose color wash is shown between 49.5 Gy and 22.5 Gy. (b) Comparison on DVHs of plans at iteration 1 (dot), 5 (square), and 13 (triangle). Improvements on PTV coverage, dose to urethra, rectum, and femurs are pointed by 5, 6, 7, and 8.

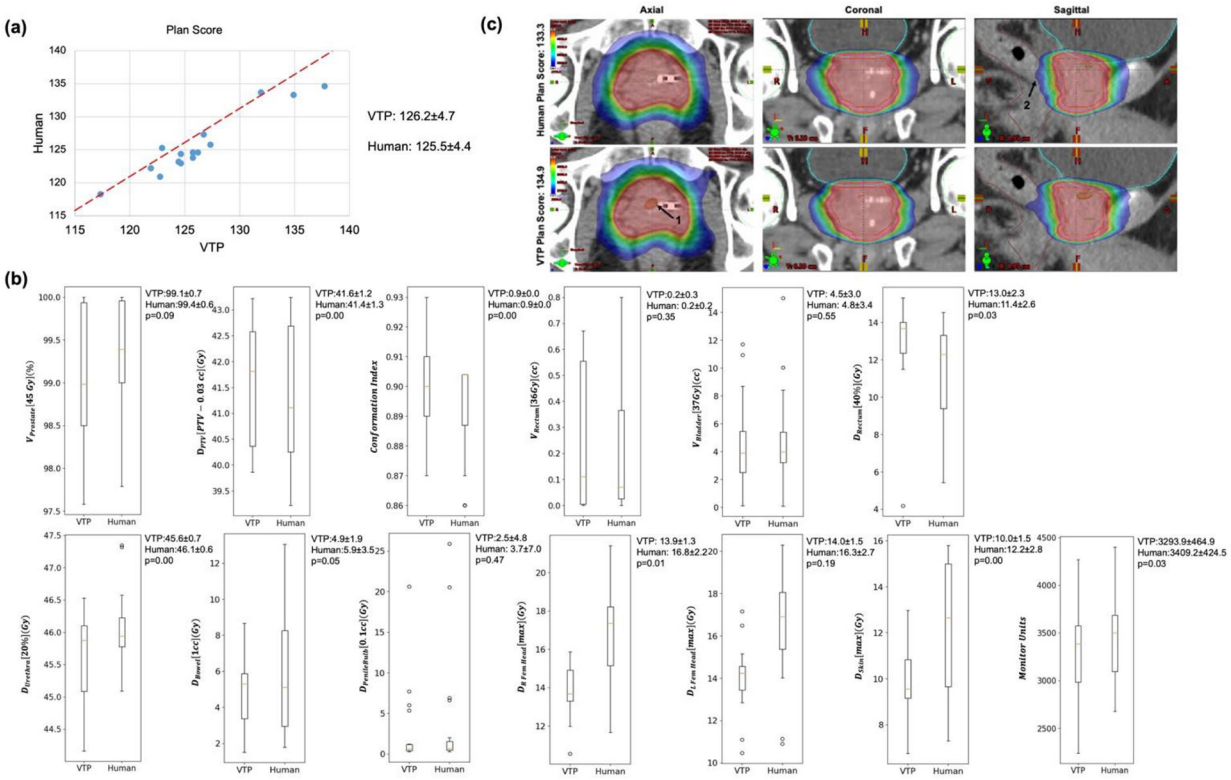

Figure 5.

(a) Scores of VMAT plans generated by VTP and human planner for 16 prostate SBRT patients. Dash line indicates equal plan quality. (b) Boxplots of dosimetric metrics and MUs to compare VTP-generated with human-generated plans of the 16 cases. Mean ± standard deviation and *p*-value of non-inferior testing for each metric are calculated and displayed in the right upper corner of the boxplot. (c) Comparisons on cross-sectional 3D dose distributions and scores. Dose color wash is shown between 45 Gy and 22.5 Gy. PTV, prostate, urethra, bladder, rectum, and femurs are segmented in magenta, red, dark green, cyan, brown, and yellow lines.

Table 1.

Dosimetric evaluation and scores of the 2016 AAMD/RSS prostate SBRT plan study case.

METRIC		Scored
Volume (%) of PTV covered by Rx dose (Gy)	95.0	*35.0 / 35.0*
Volume (%) of Prostate covered by Rx dose (Gy)	97.5	*19.2 / 20.0*
Dose (Gy) covering whole PTV minus 0.03 (cc)	34.0	*8.8 / 10.0*
Conformation Number [36.25 (Gy), PTV]	0.9	*8.6 / 10.0*
Volume (cc) of Rectum covered by 36 (Gy)	0.1	*15.0 / 15.0*
Volume (cc) of Bladder covered by 37 (Gy)	0.0	*15.0 / 15.0*
Dose (Gy) covering 40 (%) of Rectum	13.0	*10.6 / 12.0*
Dose (Gy) covering 20 (%) of Urethra	40.3	*9.3 / 10.0*
Dose (Gy) covering 1 (cc) of Bowel	1.8	*4.8 / 5.0*
Dose (Gy) covering 0.1 (cc) of Penile Bulb	2.1	*3.0 / 3.0*
Dose (Gy) covering 50 (%) of Neurovascular Bundles	38.6	*1.5 / 3.0*
Maximum dose (Gy) inside Right Femoral Head	14.0	*2.8 / 3.0*
Maximum dose (Gy) inside Left Femoral Head	13.1	*2.9 / 3.0*
Maximum dose (Gy) inside Skin	10.0	*3.0 / 3.0*
Maximum dose (Gy) inside Testes	0.3	*2.6 / 3.0*
Total Plan Score		*142.1 / 150.0*