

Are Auto-Generated Organ-at-Risk Contours Good Enough in Prostate Radiotherapy?

Han Liu^{*}, Benjamin Sintay, David Wiant

Department of Radiation Oncology, Cone Health Cancer Center, Greensboro, NC, USA

Email: *han.liu@conehealth.com

How to cite this paper: Liu, H., Sintay, B. and Wiant, D. (2025) Are Auto-Generated Organ-at-Risk Contours Good Enough in Prostate Radiotherapy? *International Journal of Medical Physics, Clinical Engineering and Radiation Oncology*, **14**, 63-73.
<https://doi.org/10.4236/ijmpcero.2025.142005>

Received: April 8, 2025

Accepted: May 10, 2025

Published: May 13, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Purpose: Manual contouring of organs-at-risk (OARs) in radiotherapy planning is a time-consuming and labor-intensive task that is prone to inter- and intra-observer variabilities. Additionally, contours drawn prior to treatment may not accurately reflect the patient's anatomy throughout the entire treatment course, meaning that doses planned based on pre-treatment contours may not correspond to the actual delivered doses. This study seeks to explore the impact of OAR contour accuracy on prostate radiotherapy outcomes. **Methods:** OARs (bladder and rectum) were manually delineated on planning CT and daily CBCT images for 20 patients. Atlas-based segmentation algorithms were used to automatically generate OAR contours on the planning CT scans. Both low-risk (LRP, CTV = Prostate) and intermediate-risk (IRP, CTV = Prostate+SV) patient groups were simulated. Treatment plans were created via a novel automated planning application that utilized a knowledge-based planning solution for both auto-generated OARs (aOAR-plan) and manual delineated OARs (mOAR-plan). Planned doses were transferred from CT to CBCTs based on clinical shifts, and contour-based deformable registrations were applied to calculate the cumulative doses. Cumulative dose differences for OARs were used to assess the agreement between aOAR- and mOAR-plans. Additionally, five patients were selected to compare the cumulative dose differences between the transferred and re-calculated doses on CBCT. **Results:** The Overlap index/Dice similarity coefficient between auto- and manual-contours was $0.89 \pm 0.10/0.77 \pm 0.11$ for the bladder and $0.76 \pm 0.15/0.70 \pm 0.07$ for the rectum. Both LRP and IRP groups demonstrated good dosimetric agreements between aOAR- and mOAR-plans. No significant OAR dose differences between the transferred and re-calculated doses were observed. **Conclusion:** Contours manually drawn prior to treatment are being used to represent patient anatomy that can fluctuate by several millimeters daily, which raises questions about the time and effort involved. Automated contouring

tools enhance contour consistency and deliver acceptable doses, providing a more reliable alternative to manual contours in prostate radiotherapy.

Keywords

Auto-Generated Contours, Knowledge-Based Planning

1. Introduction

Prostate cancer is a major health issue and one of the most commonly diagnosed cancers among men, as well as one of the leading causes of death globally [1] [2]. Treatment options for prostate cancer depend on factors such as cancer stage, grade, patient health and preferences. These options typically include active monitoring, surgery, radiotherapy, hormone therapy, chemotherapy or a combination of these approaches [3]. Among these options, external beam radiotherapy (EBRT) has been shown to be one of the most effective modalities for localized prostate cancer [4].

It has long been believed that accurate delineation of the target and organs-at-risk (OARs) is the foundation of successful radiotherapy, ensuring effective treatment while protecting surrounding normal tissue and critical organs from unnecessary radiation damage. Suboptimal organ segmentation could have a significant negative impact on treatment outcome [5]. Large contouring errors, such as incorrect organ identification, improper site delineation, or incorrect expansions, are ranked as the second most common failure mode in the AAPM TG report #100, and could lead to planning or optimization failures [6].

Manual contouring in radiation treatment planning presents several challenges that can impact both the efficiency and accuracy of treatment. This process is highly time-consuming and labor-intensive, which can lead to treatment delays, especially for rapidly progressing tumors. Additionally, manual contouring is subject to inter- and intra-observer variability of up to 10% - 20%, as different clinicians may interpret anatomy differently, or the same clinician may make inconsistent judgments over time [7] [8]. These uncertainties can result in suboptimal treatment plans. Various geometric metrics, such as the overlap index and dice similarity coefficient, have been proposed in the literature to quantify the contour variability and to serve as an indirect indication of treatment planning quality [9] [10]. However, these metrics do not correlate well with treatment outcomes [11]. Much effort has been made to evaluate the impact of contour uncertainties (inter- and intra-observer uncertainties, automatic segmentation, etc.) on target and OAR dosimetry for different treatment sites [8] [11] [12]. However, these attempts have primarily been confined to the treatment planning phase.

Since a prostate treatment course typically spans multiple fractions over several weeks, a key concern is that the contours drawn on the planning CT prior to radiation treatment may not accurately represent the patient's anatomy throughout

the entire treatment course. This is due to inter-fraction and intra-fraction motion of the prostate and OARs, meaning that doses evaluated based on pre-treatment contours may not reflect the actual delivered doses to the patient. If not considered properly, those uncertainties could result in an under-treatment of the tumor and overexposure of surrounding healthy tissues and critical organs.

Rectal toxicity is considered one of the limiting factors in EBRT planning due to the anatomic proximity of the rectum to the prostate. Unnecessary overdosing of the rectum can result in acute effects, as well as long-term complications. Currently, volumetric modulated arc therapy (VMAT) has become the preferred delivery technique due to its steep dose fall-off gradients around the target (thus improved OAR and normal tissue sparing), high dose conformality, and increased efficiency in delivery speed compared to intensity modulated radiotherapy (IMRT) and three-dimensional conformal radiotherapy (3D-CRT) techniques [13] [14]. Additionally, a new technique, injectable hydrogel spacer (SpaceOAR, Boston Scientific Corporation, Marlborough, MA), has achieved quick and widespread clinical acceptance as the standard of practice in prostate EBRT due to its capability to reduce the rectum toxicity by increasing the distance between the prostate and rectum. Considering all the factors listed above, we want to investigate whether OAR contours need to be accurate in the planning phase of prostate radiotherapy.

In recent years, semi-automatic and/or full-automated contouring tools have attracted increasing research interests and have started to gain popularity in routine clinical practice. Automated contouring tools have the potential to offer improvements in contour consistency through reduced inter-observer variability and provide time-savings from consult to treatment. However, not all centers have adopted automated tools into clinical practice due to several challenges and limitations. Besides the high initial cost and regulatory or legal considerations, one major barrier to adopting an auto-contouring tool in treatment planning is the contour accuracy as a limiting factor. In this work, we investigate the impact of the auto-generated OAR contours on the treatment outcome due to the anatomic changes during the course of prostate radiotherapy.

2. Methods and Materials

For this retrospective study, 20 prostate cancer patients were selected, each having undergone planning computed tomography (CT) and 40 daily setup cone beam CT (CBCT) scans. All patients underwent CT simulation in the supine position, utilizing Vac-Lok immobilization for the lower legs. The CT scans were acquired with a slice thickness of 2 mm, covering the region from the mid-abdomen to the mid-thigh.

The bladder and rectum were manually delineated on both the planning CT and daily CBCT images. All CBCTs were rigidly registered to the planning CT using the recorded clinical shifts for subsequent data analysis. Variable margins were applied to expand the planning CT contours, and the expansions required to

encompass 95% of the subsequent CBCT contours were recorded. This measurement reflects the precision of manual contouring on the planning CT in representing the actual tissue position over the course of radiation treatment.

Bladder and rectum contours were automatically generated using MIM Software's atlas-based segmentation algorithms (MIM Software Inc., Cleveland, OH) on the planning CT images. The contour atlas was developed using data from 70 previously treated prostate cancer patients who were representative of our patient population, encompassing a range of target and organ-at-risk (bladder and rectum) sizes and shapes, specifically the bladder and rectum. During the auto-contour generation workflow, the number of match cases used to run auto segmentation was set to 4 to ensure optimal contour accuracy. To quantify the agreement between manual and auto-generated contours, the Overlap index (*OI*) and dice similarity coefficient (*DSC*) were calculated. The prescription dose was 1.95 Gy per fraction over 40 fractions, totaling 78 Gy. Treatment plans were created through a novel automated planning (AP) application based on the Eclipse Scripting Application Programming Interface (ESAPI), which accesses a knowledge-based planning (KBP) solution for two sets of OAR structure sets: manual contours (mOAR-plan) and auto-generated contours (aOAR-plan). The AP workflow begins by importing a valid CT image and associated RT structure set into the Eclipse treatment planning system (TPS), followed by script initiation to create a treatment course and plan. Two VMAT fields are generated with the isocenter placed at the geometric center of the target, using 358-degree arcs and collimator rotations of 355 and 85 degrees. The script generates or utilizes all appropriate names, reference points, prescription details, beam calculation model, and dose calculation setting. The AP routine then uses a site-specific KBP model, trained with data from 80 previously treated prostate patients, to generate initial DVH estimates for plan optimization. Outliers were identified and excluded to ensure the robustness of the model. Once optimized, the plan undergoes final dose calculation and is normalized so that 100% of the prescribed dose (78 Gy) is delivered to 98% of the planning target volume (PTV). The AP script automatically evaluates the plan quality against departmental dosimetric guidelines to ensure adequate OAR sparing. If any constraints fail, targeted optimization is applied by the AP script. The final treatment plan is generated once all criteria are met. In this manner, the AP routine generates unbiased treatment plans for both manual and auto-generated contour sets without requiring user intervention, post calculation adjustments, or interactive modifications typical of manual planning.

This retrospective study investigated both low-risk patient (LRP) and intermediate-risk (IRP) patient groups. The clinical target volume (CTV) for LRP was the prostate, while for IRP it includes the prostate and seminal vesicle. The CTV-to-PTV margins were non-isotropic, with a 6 mm posterior expansion and an 8 mm expansion in all other directions. The calculated planning doses were transferred from the planning CT to daily CBCT images based on recorded clinical shifts. Contour-based deformable registrations were then applied to obtain the cumula-

tive doses for evaluation. Cumulative dose comparisons were made between the aOAR-plans and the mOAR-plans for the bladder and rectum. Dosimetric indices used for the plan evaluation were D_{1cc} (a representation of maximum dose), D_{mean} (mean dose), V_{XXGy} (percentage of volume receiving XX Gy) for the OAR. To ensure a fair comparison, all generated plans were normalized identically. Although treatment plans were created with different sets of contours, the dosimetric evaluation was conducted using the manually created contour sets.

In this study, the cumulative doses were derived from the transferred doses between the planning CT and daily CBCT images. To assess the impact of using transferred doses versus re-calculated doses on CBCT images, five patients were selected for comparison.

3. Results

The average CTV volume was 56.2 ± 18.8 cc, ranging from 28.4 cc to 88.4 cc, for LRP, and 71.4 ± 23.5 cc, ranging from 34.2 cc to 110.5 cc, for IRP. The application of non-uniform margins resulted in average PTV volumes of 138.1 ± 34.3 cc (range: 80.9 - 191.3 cc) for LRP and 183.6 ± 43.8 cc (range: 109.4 - 262.7 cc) for IRP. The mean rectal volume was 86.6 ± 35.9 cc with a range of 45.8 to 172.5 cc, while the mean bladder volume was 175.4 ± 77.1 cc, ranging from 83.0 to 386.1 cc.

Our data revealed significant inter-fraction variations in the bladder and rectum. On average, the planning contours required expansion of 5.0 ± 5.2 mm for the bladder, and 5.1 ± 3.8 mm for the rectum to encompass 95% of the daily CBCT contours across all patients. **Figure 1** illustrates a representative comparison of bladder and rectum contours between the planning CT and daily CBCT images for one patient.

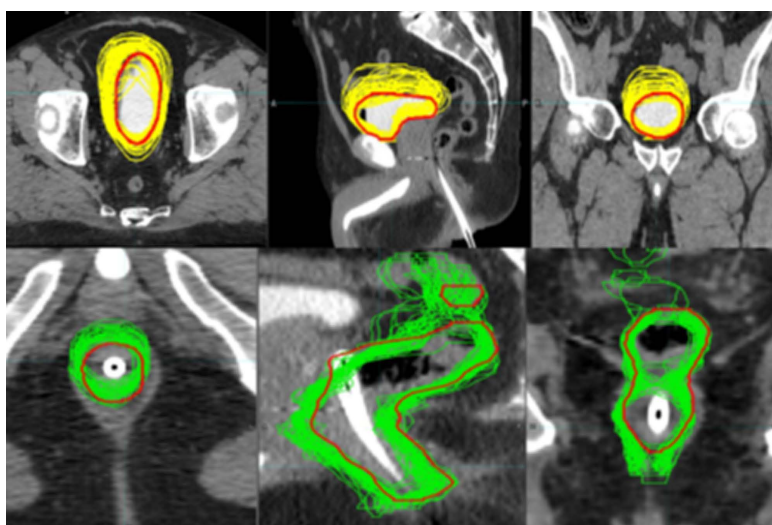


Figure 1. Comparisons of contours between the planning CT and daily CBCTs for the bladder (upper) and rectum (lower) in one representative patient. The contours are displayed on the axial (left), sagittal (middle) and coronal (right) slices of the planning CT scan. The planning CT contours are depicted in solid red, while the daily CBCT contours are shown in solid yellow for the bladder and solid green for the rectum.

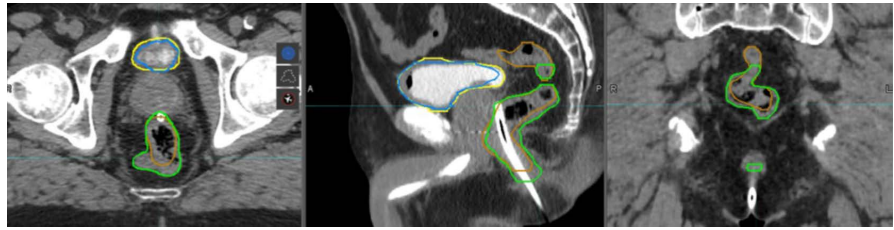


Figure 2. Comparisons of OAR contours between manual delineations and auto-generated contours using an atlas-based segmentation algorithm for one representative patient.

Figure 2 presents a comparison of bladder and rectum contours between manual delineation and auto-generated contours produced using the atlas-based segmentation algorithm for a representative patient, shown on axial, sagittal and coronal slices. **Figure 3** illustrates the *OI* and *DSC* between manual and auto-generated contours on the planning CT images for all 20 patients in this study. The average *OI* and *DSC* values for the bladder were 0.89 ± 0.11 and 0.79 ± 0.12 , respectively, while for the rectum, they were 0.75 ± 0.13 and 0.72 ± 0.06 , respectively.

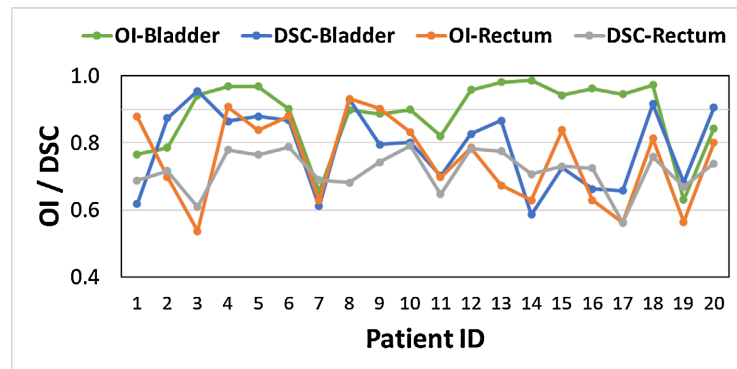


Figure 3. The overlap index (*OI*) and dice similarity coefficient (*DSC*) between the manual and auto-generated contours on the treatment planning CT.

Figure 4 and **Figure 5** display the cumulative dose comparisons for the bladder and rectum between mOAR-plans and aOAR-plans in the LRP and IRP groups, respectively. Both patient groups demonstrated good dosimetric agreement between the aOAR-plans and mOAR-plans.

Differences in cumulative doses to the OARs were used as an indicator of agreement between aOAR-plans and mOAR-plans. **Figure 6** illustrates the percentage of patients with cumulative dose differences exceeding a predefined criterion between plans generated using manual and automated OAR contours for both LRP and IRP groups. The figure demonstrates that good agreement was achieved between manual and automated contours for the majority of patients in the study.

In this study, cumulative doses were calculated using two methods: transferred doses from planning CT to daily CBCT and re-calculated doses on CBCT images. Five patients were selected to evaluate differences between these methods. As shown in **Figure 7**, no significant dose differences were observed for the bladder and rectum between the transferred and re-calculated doses, except for patient #2

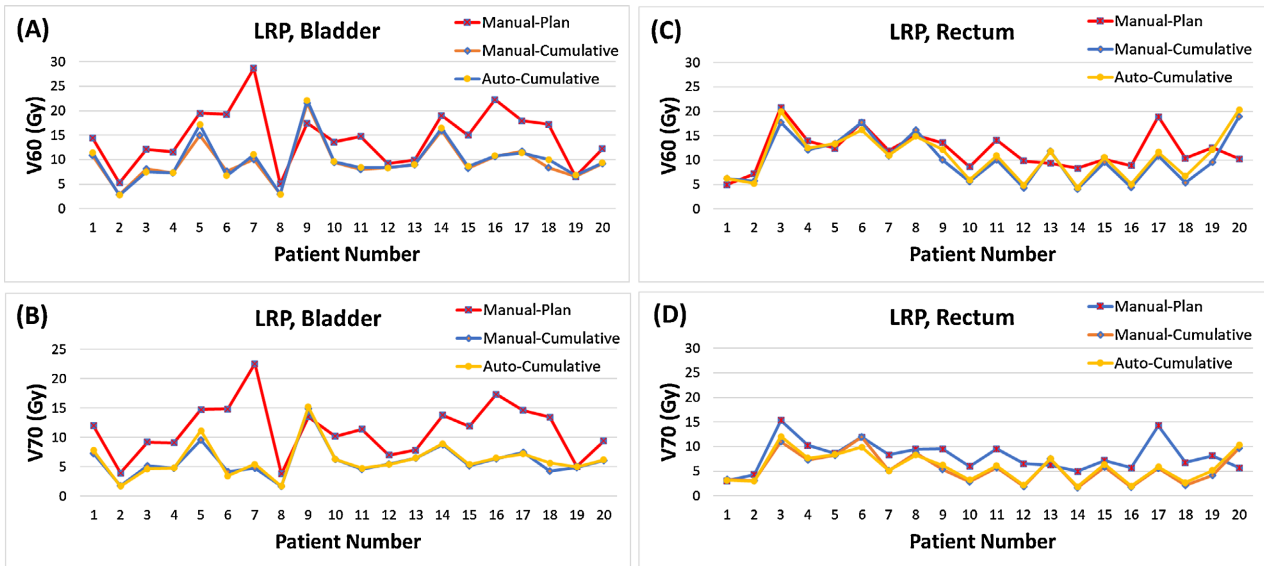


Figure 4. Comparison of cumulative doses (V_{70} and V_{60}) for the bladder and rectum between plans generated from manual contours and atlas-based auto-contours in the low-risk patient group.

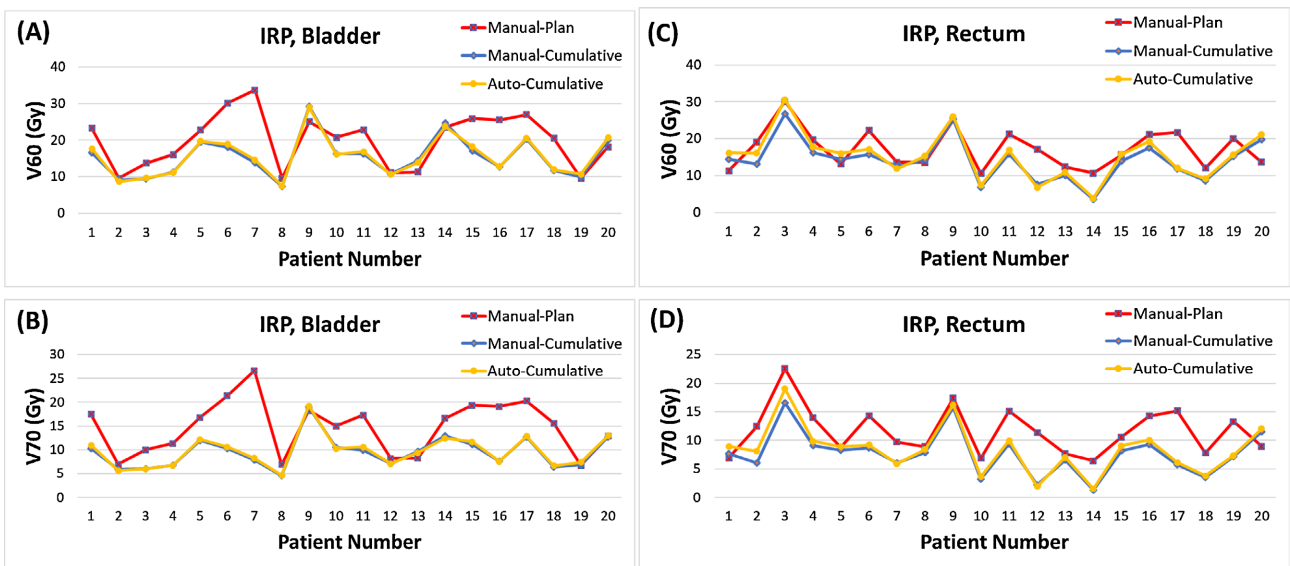
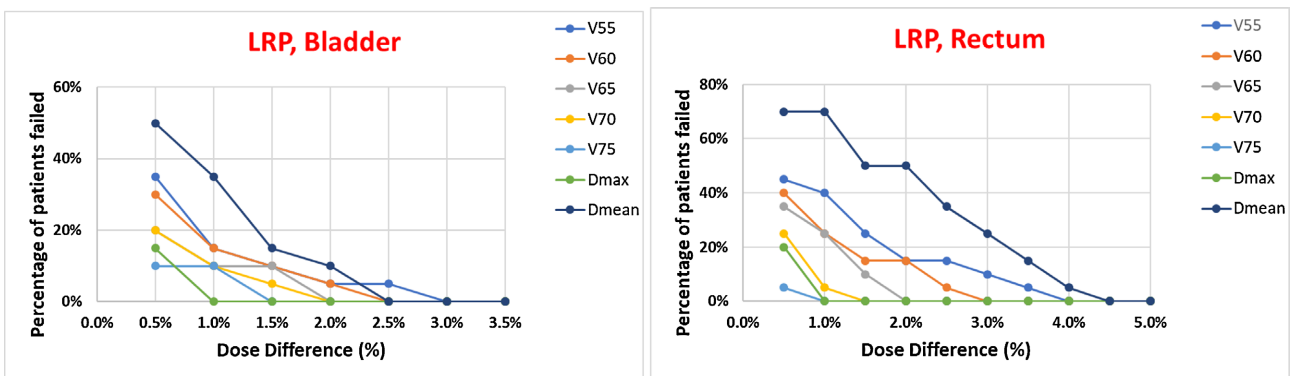


Figure 5. Comparison of cumulative doses (V_{70} and V_{60}) for the bladder and rectum between plans generated from manual contours and atlas-based auto-contours in the intermediate-risk patient group.



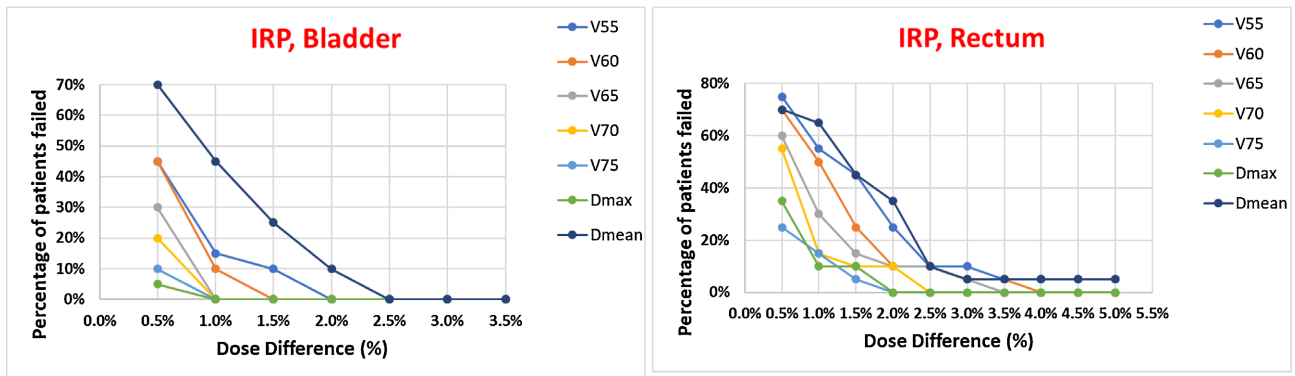


Figure 6. Breakdown of cumulative dose differences between plans generated using manual and automated OAR contours for the low-risk and intermediate-risk patient groups.

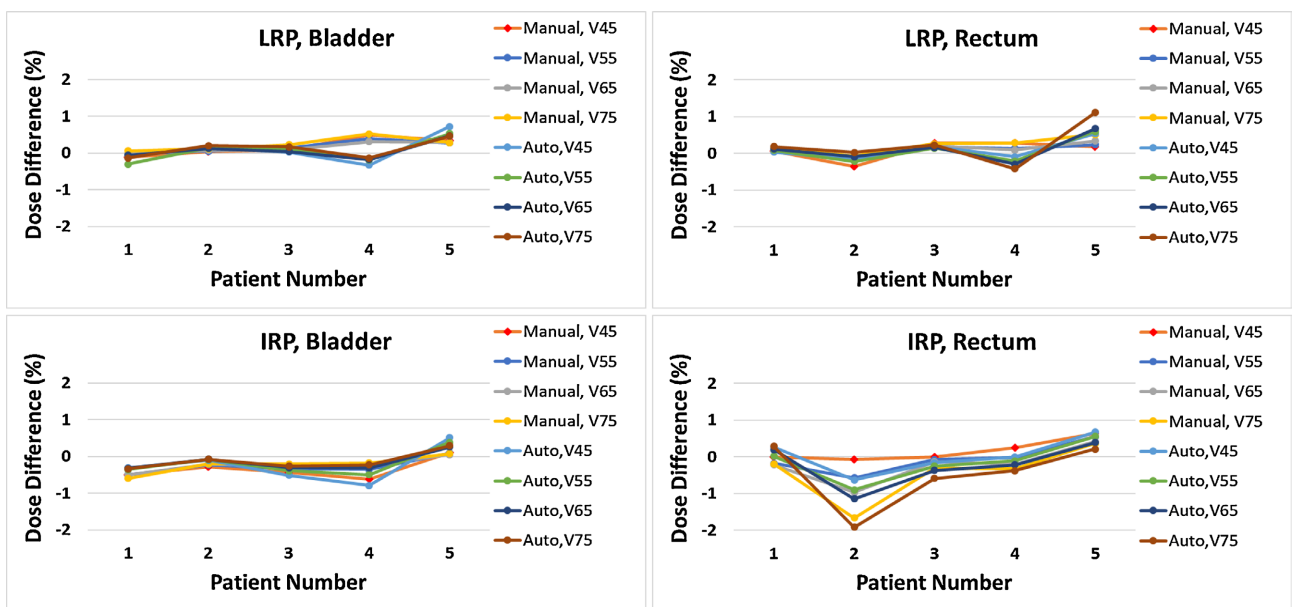


Figure 7. Comparison of cumulative dose differences between transferred and recalculated doses for plans generated with manual and auto-OAR contours in the low-risk and intermediate-risk patient groups.

(IRP group), where the difference remained within 2%.

4. Discussion

Precise delineation of the target and organs-at-risk during the treatment planning process is crucial for the accuracy and effectiveness of the radiation therapy. However, contouring is a time-consuming and labor-intensive task that is often subjective, depending heavily on the clinician’s expertise. In prostate EBRT, despite accounting for inter- and inter-observer contour variability, target structures such as the prostate and seminal vesicles still exhibit daily deformation, while the bladder and rectum undergo both deformation and displacement relative to the prostate. Consequently, manually drawn contours on planning CT images prior to radiation treatment may not reflect the patient’s anatomy throughout the treatment course. This was corroborated in our study, where the average expansion required

to encompass 95% of the daily CBCT contours over all patients was 5.0 ± 5.2 mm for the bladder and 5.1 ± 3.8 mm for the rectum. Considering this, it may be feasible to use slightly less “accurate” organ delineation to model the actual delivered dose given the inherent anatomical variations over the treatment period.

The primary objective of treatment planning in radiotherapy is to ensure adequate target coverage while minimizing exposure to surrounding normal tissue and critical organs. Studies have emphasized that target delineation accuracy is essential for achieving optimal target dosimetric coverage [8]. In this study, we focused on evaluating the OARs and found reasonable agreement between manual and auto-generated contours with average *OII/DSC* values of $0.89 \pm 0.11/0.79 \pm 0.12$ for the bladder, and $0.75 \pm 0.13/0.72 \pm 0.06$ for the rectum. It is important to note that atlas-based segmentation was employed to generate the auto-contours in this study. With the ongoing rapid advancements in artificial intelligence-based tools, the accuracy of auto-generated contours is expected to improve, potentially surpassing the results observed here. Additionally, the use of an auto-planning routine played a critical role in this evaluation, enabling the consistent generation of un-biased high-quality plans based on the provided structure sets.

In this study, differences in the cumulative doses to the OARs were used as an indicator of the agreement between aOAR-plans and mOAR-plans, with contour-based deformable registration employed to calculate the cumulative doses. For both LRP and IRP groups, good dosimetric agreements between the aOAR- and mOAR-plans were observed for both the bladder and rectum, as shown in **Figure 5** and **Figure 6**. **Figure 6** highlights the percentage of patient failures based on specific cumulative dose difference criteria between the aOAR- and mOAR-plans. For both LRP and IRP groups, only a small percentage of patients failed to meet the 2% dose difference criteria for the bladder and rectum.

Two methods were used to calculate the cumulative doses: transferred doses from planning CT to daily CBCT and re-calculated doses on CBCT images. No significant dose differences for the bladder and rectum were observed between the two methods except for patient #2 in the IRP group. However, the dose difference for this patient remained within 2%. Further investigation of this patient’s CBCT images revealed that the larger dose differences were due to significant gas filling variation in the rectum.

We want to point out that this is a retrospective study evaluating whether auto-generated OAR contours are sufficiently accurate for prostate treatment planning by simulating the cumulative dose differences between plans created by using manual and auto-generated OAR contours. The impact on clinical outcome, which is of greater clinical relevance, falls outside the scope of this study and could be addressed in future clinical trials.

This work is a feasibility study focused on prostate radiotherapy, and caution should be exercised when extrapolating these findings to more complex disease sites. In contrast to areas such as head and neck, prostate treatment planning involves relatively fewer critical OARs, such as the bladder and rectum, which are

parallel organs where the focus is on reducing mid to higher dose levels. Additionally, the geometric relationship between the target and these OARs tends to be relatively consistent across prostate patients, which may explain why manual contouring of OARs has minimal impact on plan quality. However, more complex anatomy, critical structures, and the geometric relationship between targets and OARs, especially in cases involving serial organs, may result in different dosimetric and geometric interactions compared to those observed in this study.

5. Conclusion

Manually drawn contours on planning CT scans are being used to represent anatomy that can fluctuate by several millimeters on a daily basis, raising questions about the time and effort involved. Additionally, reliance on manual contouring increases the likelihood of errors, particularly in high-pressure clinical environments, which could compromise treatment outcomes. Automated contouring tools offer improvement in contour consistency through reduced inter-observer variability, provide time-savings from consultation to treatment, and deliver acceptable doses over the treatment course compared to manual contours. It is important to note that these results are specific to prostate radiotherapy and require further investigation for more complex treatment sites.

Funding

This research was supported by a grant from Varian Medical Systems, Palo Alto, CA.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Siegel, R.L., Miller, K.D. and Jemal, A. (2019) Cancer Statistics, 2019. *CA: A Cancer Journal for Clinicians*, **69**, 7-34. <https://doi.org/10.3322/caac.21551>
- [2] Mattiuzzi, C. and Lippi, G. (2019) Current Cancer Epidemiology. *Journal of Epidemiology and Global Health*, **9**, 217-222. <https://doi.org/10.2991/jegh.k.191008.001>
- [3] Hamdy, F.C., Donovan, J.L., Lane, J.A., Mason, M., Metcalfe, C., Holding, P., *et al.* (2016) 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. *New England Journal of Medicine*, **375**, 1415-1424. <https://doi.org/10.1056/nejmoa1606220>
- [4] Podder, T.K., Fredman, E.T. and Ellis, R.J. (2018) Advances in Radiotherapy for Prostate Cancer Treatment. In: Schatten, H., Ed., *Advances in Experimental Medicine and Biology*, Springer International Publishing, 31-47. https://doi.org/10.1007/978-3-319-99286-0_2
- [5] Fiorino, C., Reni, M., Bolognesi, A., Cattaneo, G.M. and Calandrino, R. (1998) Intra- and Inter-Observer Variability in Contouring Prostate and Seminal Vesicles: Implications for Conformal Treatment Planning. *Radiotherapy and Oncology*, **47**, 285-292. [https://doi.org/10.1016/s0167-8140\(98\)00021-8](https://doi.org/10.1016/s0167-8140(98)00021-8)
- [6] Huq, M.S., Fraass, B.A., Dunscombe, P.B., Gibbons, J.P., Ibbott, G.S., Mundt, A.J., *et*

- al.* (2016) The Report of Task Group 100 of the AAPM: Application of Risk Analysis Methods to Radiation Therapy Quality Management. *Medical Physics*, **43**, 4209-4262. <https://doi.org/10.1118/1.4947547>
- [7] Alasti, H., Cho, Y., Catton, C., Berlin, A., Chung, P., Bayley, A., *et al.* (2017) Evaluation of High Dose Volumetric CT to Reduce Inter-Observer Delineation Variability and PTV Margins for Prostate Cancer Radiotherapy. *Radiotherapy and Oncology*, **125**, 118-123. <https://doi.org/10.1016/j.radonc.2017.08.012>
- [8] Liu, H., Amaloo, C., Sintay, B. and Wiant, D. (2021) Dosimetric Effects Due to Inter-Observer Variability of Organ Contouring When Utilizing a Knowledge-Based Planning System for Prostate Cancer. *International Journal of Medical Physics, Clinical Engineering and Radiation Oncology*, **10**, 47-58. <https://doi.org/10.4236/ijmpcero.2021.102005>
- [9] Lee, W.R., Roach, M., Michalski, J., Moran, B. and Beyer, D. (2002) Interobserver Variability Leads to Significant Differences in Quantifiers of Prostate Implant Adequacy. *International Journal of Radiation Oncology, Biology, Physics*, **54**, 457-461. [https://doi.org/10.1016/s0360-3016\(02\)02950-4](https://doi.org/10.1016/s0360-3016(02)02950-4)
- [10] Vinod, S.K., Min, M., Jameson, M.G. and Holloway, L.C. (2016) A Review of Interventions to Reduce Inter-Observer Variability in Volume Delineation in Radiation Oncology. *Journal of Medical Imaging and Radiation Oncology*, **60**, 393-406. <https://doi.org/10.1111/1754-9485.12462>
- [11] Lim, T.Y., Gillespie, E., Murphy, J. and Moore, K.L. (2019) Clinically Oriented Contour Evaluation Using Dosimetric Indices Generated from Automated Knowledge-Based Planning. *International Journal of Radiation Oncology, Biology, Physics*, **103**, 1251-1260. <https://doi.org/10.1016/j.ijrobp.2018.11.048>
- [12] Tsuji, S.Y., Hwang, A., Weinberg, V., Yom, S.S., Quivey, J.M. and Xia, P. (2010) Dosimetric Evaluation of Automatic Segmentation for Adaptive IMRT for Head-and-Neck Cancer. *International Journal of Radiation Oncology, Biology, Physics*, **77**, 707-714. <https://doi.org/10.1016/j.ijrobp.2009.06.012>
- [13] Otto, K. (2007) Volumetric Modulated Arc Therapy: IMRT in a Single Gantry Arc. *Medical Physics*, **35**, 310-317. <https://doi.org/10.1118/1.2818738>
- [14] Yoo, S., Wu, Q.J., Lee, W.R. and Yin, F. (2010) Radiotherapy Treatment Plans with Rapidarc for Prostate Cancer Involving Seminal Vesicles and Lymph Nodes. *International Journal of Radiation Oncology, Biology, Physics*, **76**, 935-942. <https://doi.org/10.1016/j.ijrobp.2009.07.1677>