



Clinical Validation of Commercial AI Prostate SBRT Planning: Dosimetric Equivalence with Deficiencies Isolated to Institution-Specific Organ-at-Risk Objectives

Nayoon J. Lee, University of Pennsylvania; Steven Philbrook, MS; William R. Green, MD; Eva Berlin, MD; John Chojnowski, MD, PhD; Nicholas Molina; Jonathan Cheng; Sara Arias-Obando; Rafe McBeth, PhD

Introduction/Background

Commercial AI planning systems achieve clinically acceptable quality while reducing planning time, but adoption is limited by a “last-mile” challenge: these systems are trained on generalized objectives that often miss institution-specific organ-at-risk (OAR) priorities and contouring conventions. We hypothesized these deficiencies reflect missing institutional knowledge rather than algorithmic limitations, and tested this by validating a commercial AI system against manual plans and assessing whether a lightweight institutional model could resolve identified gaps.

Methods/Intervention

23 prostate SBRT cases (40 Gy/5) were replanned with a commercial AI system trained on a 36.25 Gy/5 regimen, using identical CT and structure sets as approved clinical plans, normalized to D95%=100%. Metrics included PTV D95%/D99%, Paddick Conformity Index, and institutional OAR constraints (urethra, bladder, rectum, femoral heads, penile bulb, skin). A blinded A/B preference test on 20 cases was reviewed by 3 physicians. Five AI plans underwent gamma-based QA (2%/2mm). A separate model trained on 21 institutional cases with institution-prioritized bowel constraints was benchmarked against it.

Results/Outcome

AI plans met PTV D95% in all cases with comparable conformality, using consistently 2 arcs (vs. 2–3 manual) and averaging 438 fewer monitor units, in under 10 minutes. Excluding urethra, 99.2% of DVH objectives were met (390/393); including urethra, 93.5%, with violations concentrated in urethral D0.035cc (mean excess $\approx$ 101 cGy) due to exclusion from vendor training (Fig. 1). Across 60 blinded evaluations, no plan was rated unacceptable; manual plans were favored in 86% of evaluations with a stated preference (42% no preference), localizing the gap to the omitted urethral objective (Fig. 2). QA achieved 100% gamma pass rates. The institution-specific model matched target coverage (CTV D99% within 84 cGy) with superior bowel sparing versus the commercial model (Fig. 3).

Conclusion

Commercial AI-based prostate SBRT planning generalizes safely across regimens and approaches clinical equivalence with manual planning. Residual deficiencies are concentrated in omitted institutional objectives, not the optimization approach. A lightweight model trained on 21 cases resolved this gap, confirming institutional data as the limiting factor for clinical readiness.

Statement of Impact

This establishes a data-efficient framework for adapting commercial AI planning systems to local protocols without full retraining, reconciling vendor models with local priorities.

Average Dosimetric Difference by Structure Objective

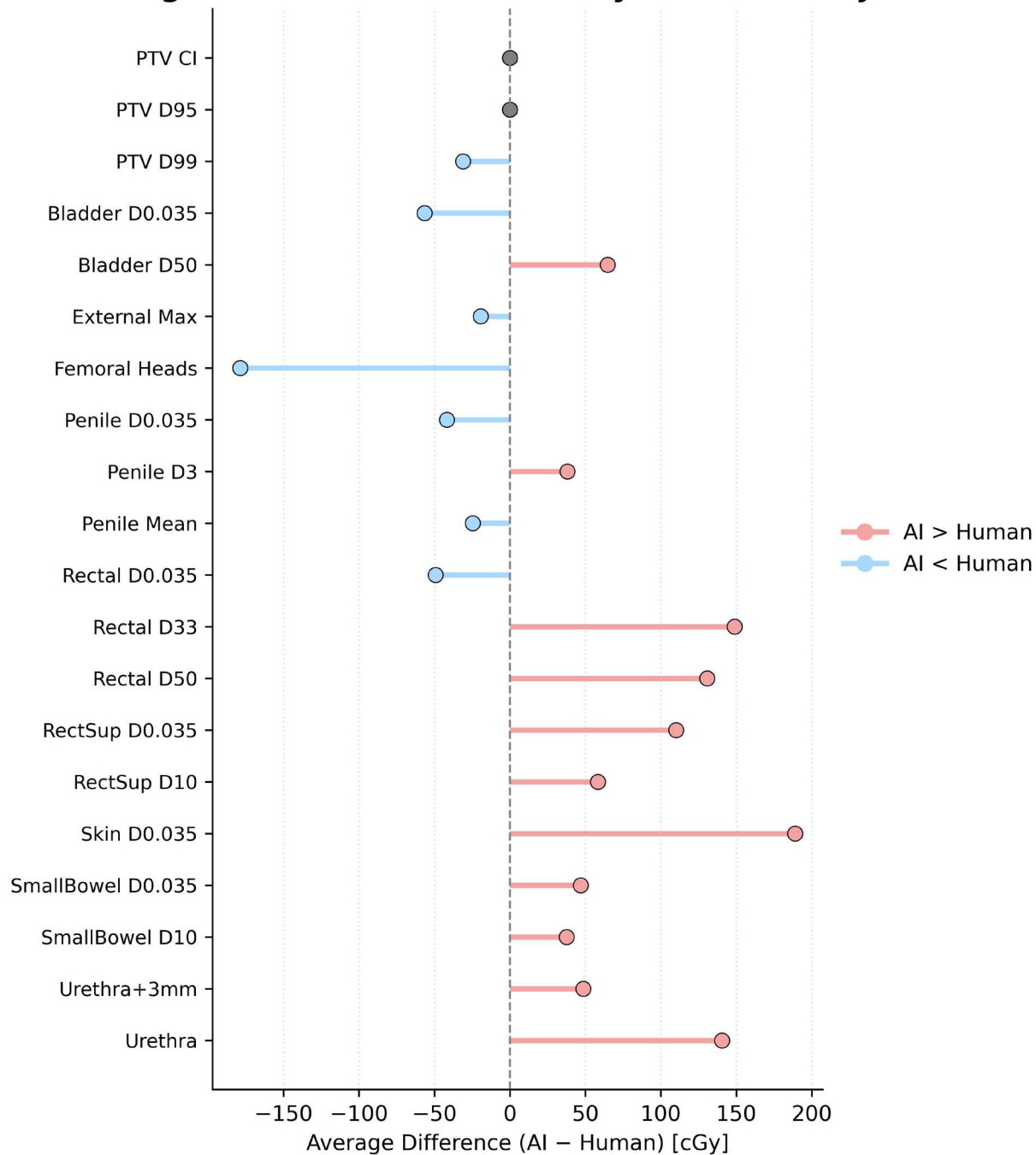


FIGURE 1: Average Dosimetric Difference by Structure Objective between Human Expert and Vendor Plans.

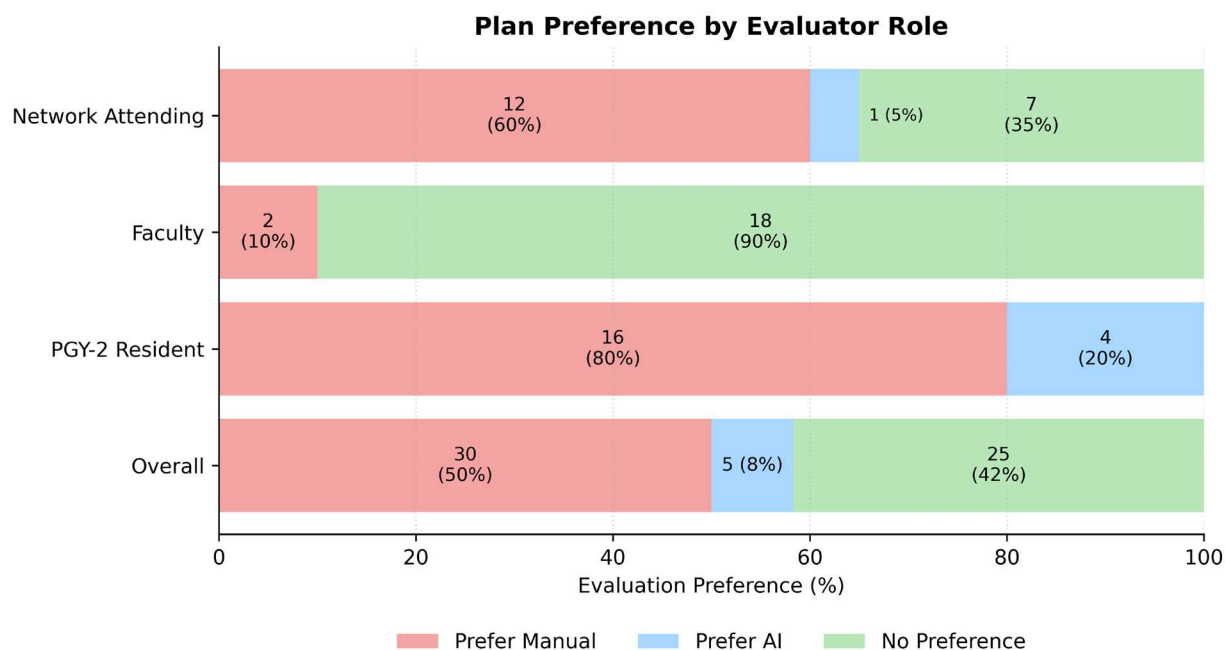


FIGURE 2: Blinded Plan Preference by Evaluator (Network Attending, Faculty, PGY-2 Resident, and Overall).

Structure	Objective	Clinical	In-house Validation	Vendor
PTV	D97% (cGy)	6985.0 cGy	6985.0 cGy	7020.2 cGy
	D99% (cGy)	6880.9 cGy	6890.9 cGy	6906.7 cGy
	D99% (cGy)	7074.7 cGy	7023.0 cGy	7044.4 cGy
Rectum	V36 Gy (%)	25.05 %	25.00 %	26.44 %
	V40 Gy (%)	21.22 %	21.12 %	22.87 %
	V53 Gy (%)	11.40 %	11.93 %	13.83 %
	V61 Gy (%)	6.71 %	7.29 %	8.83 %
	D1cc (cGy)	7060.8 cGy	7123.9 cGy	7180.3 cGy
Bladder	V40 Gy (%)	15.63 %	15.66 %	15.42 %
	V44 Gy (%)	13.81 %	13.91 %	13.77 %
	V57 Gy (%)	8.92 %	9.09 %	9.02 %
	V61 Gy (%)	7.73 %	7.91 %	7.85 %
Bowel – Small	V32 Gy (%)	0.00 %	0.00 %	0.00 %
	V36 Gy (%)	0.00 %	0.00 %	0.00 %
	D1cc (cGy)	97.4 cGy	96.4 cGy	100.1 cGy
Bowel – Large	D1cc (cGy)	458.9 cGy	446.0 cGy	464.8 cGy
Femur Head	V40 Gy (%) L	0.00 %	0.00 %	0.00 %
	V40 Gy (%) R	0.00 %	0.00 %	0.00 %
Body	Max (cGy)	7391.5 cGy	7477.5 cGy	7536.5 cGy
External Patient	Max (cGy)	7391.5 cGy	7477.5 cGy	7536.5 cGy

FIGURE 3: External Beam Radiation Therapy Dose Metrics Comparison between Clinical, In-House Model, and Commercial Vendor Model.

Keywords

Prostate Stereotactic Body Radiation Therapy (SBRT); AI Treatment Planning; Dosimetric Validation; Organ-at-risk Optimization Institutional Model Adaption; Clinical AI Adoption