Clinical Evaluation of the MVision AI Commercial Contouring Solution for Head and Neck (H&N) and Prostate Radiotherapy Treatments

Kartik Kumar[1], Kevin Tran[1], Supun Hettige[1] and Christopher Wong[1]
[1]Alfred Health Radiation Oncology, Bayside health

2026 JOINT AAPM | COMP MEETING
JULY 19–22 | VANCOUVER, BC

Introduction

Deep-learning auto-segmentation is increasingly used in radiotherapy to reduce contouring time and inter-observer variability but must be validated before clinical adoption. This study validates a commercial system (MVision AI) against clinical contours across head-and-neck and prostate cohorts using geometric and dosimetric endpoints and examines the suitability of different performance metrics for different contours.

Method

MVision AI auto-contours were compared against approved clinical contours across two sites. Cohort, metrics, and acceptance thresholds are given in **Table 1**. Dosimetric agreement was assessed by recomputing dose on the AI and clinical contours from the same clinical plan, with no re-optimisation, so differences reflect only contour geometry.

Table 1. Cohort characteristics, geometric and dosimetric metrics, and clinical acceptance thresholds.

	H&N	Prostate
Patients	15	15
Slice thickness	1.25-3.0mm	1.25–3.0 mm
Metric	**Type**	**Threshold**
DSC	Volume overlap	≥ 0.80
s-DSC (τ = 3 mm)[1]	Boundary	≥ 0.85
Mean, D2, D95, D98	Dose difference	—

Results

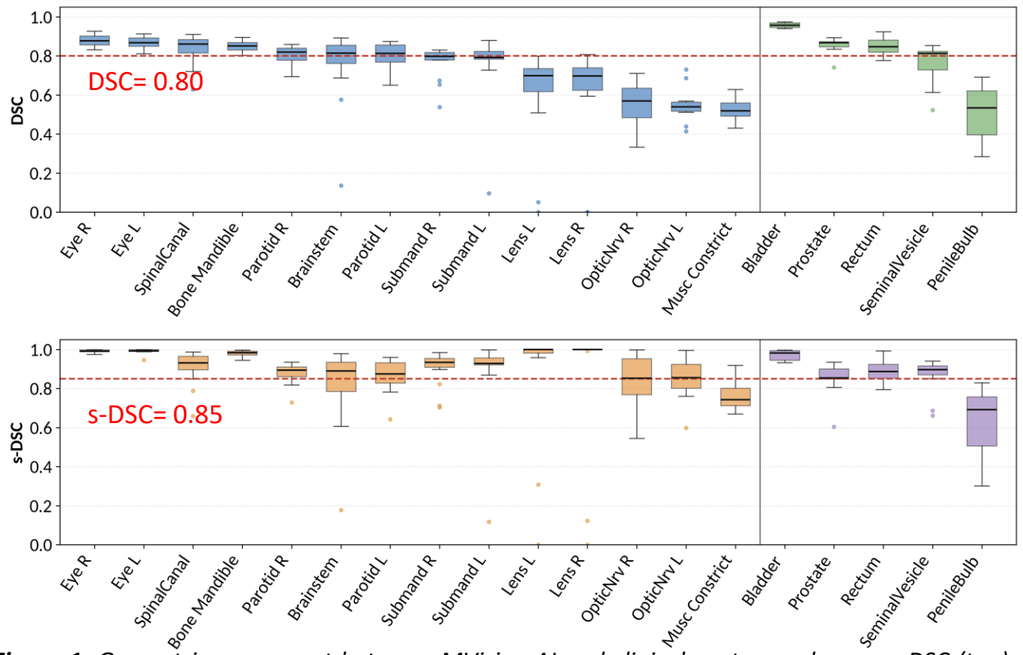

Figure 1. *Geometric agreement between MVision AI and clinical contours, shown as DSC (top) and surface DSC (τ = 3 mm, bottom). Small structures (lens, optic nerves) fall below the DSC threshold while remaining above the surface-DSC threshold, illustrating that the choice of metric changes the acceptance outcome for these structures.*

Take Home Message

1. **MVision is clinically acceptable for most OARs.** Median DSC 0.52–0.96; s-DSC ≥ 0.85 for all structures except muscle constrictor and penile bulb ***(Fig 1)***.
2. **The two metrics diverge for small organs.** Disagreements concentrate in contours < 10 cc (median 1.3 cc); larger organs agree ***(Fig 2)***.
3. **DSC's rejection has no dosimetric basis.** Contours DSC rejected but s-DSC accepted differ by only 0.15 Gy (2.8%) mean / 0.09 Gy (1.1%) near-max ***(Table 2)***.
4. **Use s-DSC for small OARs** — DSC's volume-overlap is disproportionately penalised at low volume.

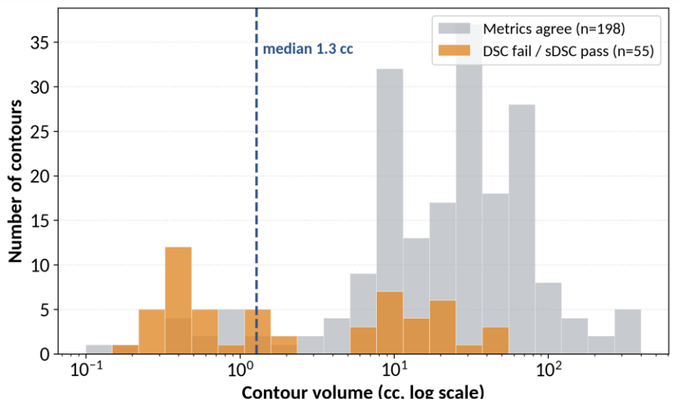

Figure 2 *Contour volume for cases where DSC and surface DSC disagreed compared with cases where the two metrics agreed .The two metrics diverge specifically for small structures.*

Table 2. Dosimetric impact of contours rejected by DSC (< 0.80) but accepted by surface DSC (≥ 0.85) — n = 55, volume < 10 cc.

Dosimetric endpoint	Median	IQR
Mean dose difference	0.15 Gy (2.8%)	0.01–0.99 Gy
Near-max (D2) difference	0.09 Gy (1.1%)	0.02–0.93 Gy

Conclusion

MVision AI produced clinically acceptable contours for most organs at risk across both head-and-neck and prostate sites. DSC disproportionately penalises small-volume contours, rejecting geometry that is dosimetrically acceptable. Surface DSC should therefore be the preferred acceptance metric for small OARs [2].

Kartik Kumar
Alfred Health Radiation Oncology, Melbourne , Australia
Email: ka.kumar@alfred.org.au

References
1. Nikolov S, Blackwell S, Zverovitch A, Mendes R, Livne M, De Fauw J, et al. Clinically applicable segmentation of head and neck anatomy for radiotherapy: deep learning algorithm development and validation study. J Med Internet Res. 2021;23(7):e26151.
2. Vaassen F, Hazelaar C, Vaniqui A, Gooding M, van der Heyden B, Canters R, et al. Evaluation of measures for assessing time-saving of automatic organ-at-risk segmentation in radiotherapy. Phys Imaging Radiat Oncol. 2020;13:1–6.