Abstract Overview

Multicentric Comparative Analysis of 8 Commercial Deep Learning-Based Autocontouring Solutions for the Delineation of Head and Neck Organs-at-Risk

Authors: M. Westermayer, N. Delaby, A. Sottiaux, S. Kara Mohammed, MC. Biston, C. Khamphan, S. Huger, S. Chiavassa, I. Bessieres, I. Achag , M. Stadler, S. Farzam, C. Millardet-Martin , T. Lacornerie, C. Fontbonne, JM Fontbonne, L. Madec, R. Castanet, A. Coutte, J. Biau, Y. Pointreau, A.A. Serre, J. Castelli, A. Barateau, A. Badey

Purpose

- The study aims to compare the performance of 8 different commercial DL-based AS solutions for H&N on CT images.

Mat&Methods

- Data: 20 HN patients.
- AI solutions: MiradaDLCExpert, MVision , RayStation, ART-Plan, Syngo.Via, Limbus, MIM, AutoContour.
- OARs: brainstem, oral cavity, spinal canal, parotid glands, mandible, optic nerves, eyes, and submandibular glands.
- Manual contours: done by 5 rad Oncs .
- Geometric evaluation: DICE and HD95%.
- Dosimetric impact: by projecting the reference treatment plan (optimized using manual reference contours) onto the autocontours and comparing DVH points.
- Statistical investigation of relation between geometric results and dosimetric differences.

Results

- The mean vDSC values exceeded 0.8 across almost all structures and solutions.
- TheraPanacea had the highest vDICE results for all the OARs except Oral Cavity (with highest DICE was for syngo.via and MVision).
- Dosimetric differences had very few strong corellations with geometric indices, and in particular for organs located closer to the target volumes.

Conclusion

•All commercial solution provided satisfying geometric results, however since a weak correlation with dosimetric impact, case specific manual correction and approval by a rad onc are still required.

•Leading performance vs. other commercial solutions as a result of High geometric accuracy across OARs.
Achieving the highest DICE for most OARs, demonstrating competitive performance among 8 evaluated tools.
ART-Plan is among the solutions achieving mean DICE scores > 0.8 for most structures in Head and neck cancer.

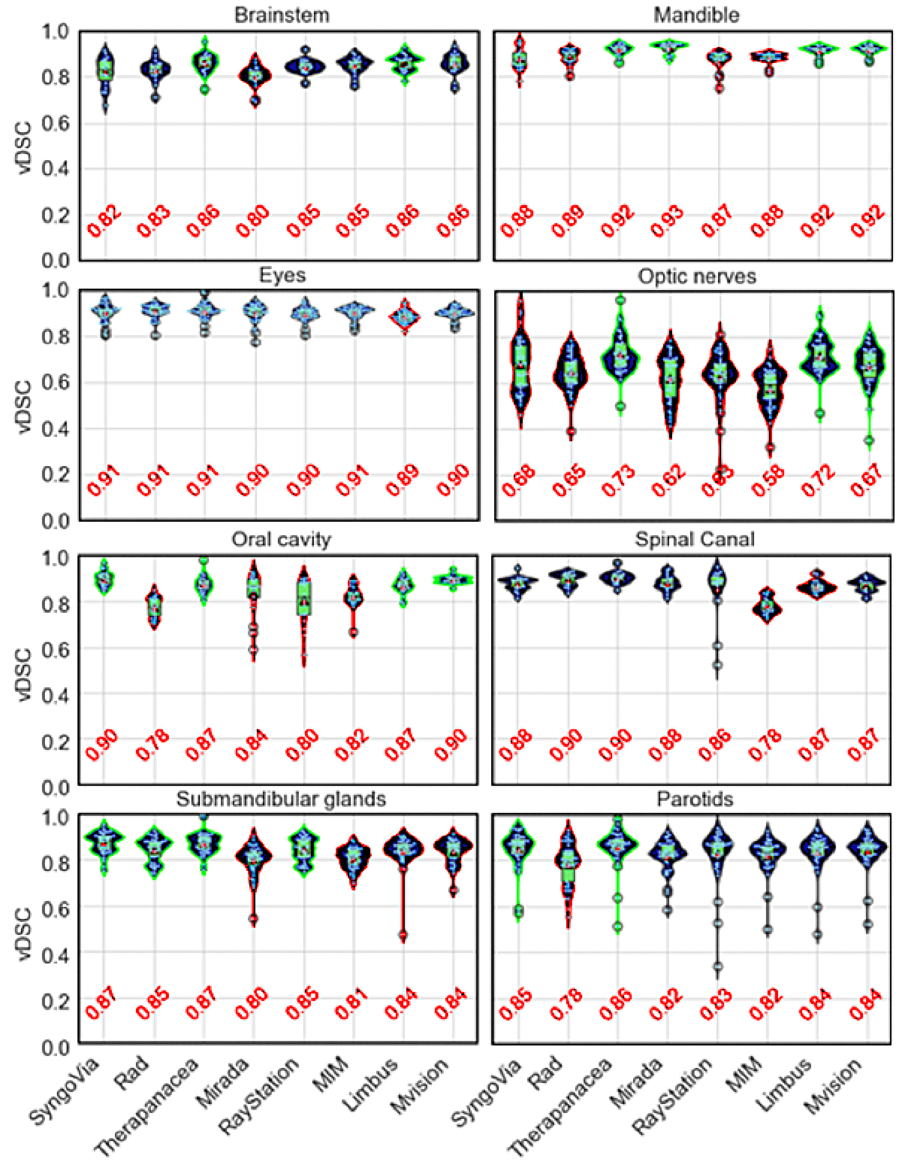

Fig. 1: vDSC scores obtained by the comparison of the reference contours with the individual auto segmentation solutions contours. Solutions that differ significantly (p < 0.05) from more than three other solutions are highlighted with either a green or red frame , indicating superior or inferior performance, respectively. Mean values are indicated as red numbers.

GUSTAVE ROUSSY CANCER CAMPUS GRAND PARIS
CENTRE DE LUTTE CONTRE LE CANCER LEON BERARD
Penn UNIVERSITY of PENNSYLVANIA
THERAPANACEA

Abstract Overview

Performance Evaluation of an AI-Driven Auto-Segmentation Tool for MR-Based Cervical Cancer Brachytherapy Planning

Authors: Elaine Limkin, Teo Kevin, Sophie Espenel, Isabelle Dumas, Ines Chaffai-Nouri, Alexandru Iancu, Anne-Agathe Serre, Frédéric Gassa, Thomas Theodoridis, Sami Romdhani, Lhassa Macke, Madalina-Liana Costea, Nikos Paragios, Charlotte Robert

Purpose

- Manual delineation of organs-at-risk (OARs) for cervical brachytherapy is labor intensive, time consuming, and subject to inter-observer variability.

Mat&Methods

- Structures: Bladder, rectum, sigmoid, bowel.
- Training N=327 MR T2 images with varied brachytherapy applicators (Vienna, Geneva, Ring, Venezia or personalized mold) from a single center in Europe.
- Testing N=42 patients from 2 centers (EU and US).
- Evaluation Metric : DICE, and HD95 & qualitative evaluation from 2 rad Oncs.

Results

- DSC: Bladder = 0.9, Rectum= 0.86, Sigmoid = 0.75, bowel = 0.8.
- Clinical acceptability after Physicis Review (with minor corrections):
  - Bladder 94%, Rectum 100%, Sigmod 86% and Bowel 92%.

Conclusion

- This tool may shorten the critical gap between MRI acquisition and treatment delivery, optimizing the most time intensive step of the treatment workflow.

- High geometric accuracy and clinical acceptability across datasets and applicator types.
  The AI tool achieves strong performance for OAR delineation in Cervical cancer Brachytherapy, with DSC up to 0.9 and clinical acceptability rates up to 100% after minor corrections. (Trained on diverse MRI data and validated on independent multicenter cohorts, showing consistent results across different clinical settings.)
- Reduces a critical workflow bottleneck.
  Automates a labor-intensive step, helping shorten the time between MRI acquisition and treatment delivery.

	Quantitative metrics (Manual vs AI contours)						Qualitative Evaluation of the AI Contours
	Combined Datasets		Dataset 1 (N=25)		Dataset 2 (N=17)		
OAR	Mean DSC	Mean HD95 (mm)	Mean DSC	Mean HD95 (mm)	Mean DSC	Mean HD95 (mm)	Mean (A%+B%)
Bladder	0.9±0.04	4.4±4.0	0.91±0.1	4.8±4.8	0.9±0.1	4.2±3.5	94%
Rectum	0.86±0.1	7.2±7.7	0.85±0.1	8.5±8.1	0.86±0.1	6.2±7.5	100%
Sigmoid	0.75±0.1	29.7±26.8	0.79±0.1	28.3±31.3	0.72±0.2	30.6±23.9	86%
Bowel	0.8±0.1	22.8±19.2	0.83±0.1	21.9±24.3	0.78±0.1	23.3±15.4	92%

Table 1: Quantitative and qualitative evaluation results.

GUSTAVE ROUSSY CANCER CAMPUS GRAND PARIS

CENTRE DE LUTTE CONTRE LE CANCER LEON BERARD

Penn University of Pennsylvania

THERAPANACEA

Abstract Overview

Automated VMAT Planning for Ultra-Hypofractionated Prostate Radiotherapy

Authors: Pauline Maury, Willliam Green, Rafe McBeth, Enis Ozyar, Gorkem Gungor, Madalina-Liana Costea, Baris Ungun, Remi Vauclin, Elie Mengin, Nikos Paragios, Mario Terlizzi

Purpose

- To clinically validate a DL based automated planning (AP) pipeline solution for localized prostate (36.25 Gy / 5 fx).

Mat&Methods

- The AP pipeline was trained on 138 VMAT manual plans.
- Clinical validation was done on 20 patients from 2 centers in the following configuration:
  - 11 plans VersaHD with Agility MLC with Raystation TPS.
  - 9 plans TrueBeam with HDMLC with Eclipse TPS.
  - Quantiative evaluation was done between Manual and AP plans based on GETUC criteria.
  - Qualitative review was done by 3 senior Rad Oncs.
  - A subset of cases underwent patient specific plan QA measurements.

Results

- The full AP process required approximately 10 minutes.
- AP plans demonstrated superior target coverage compared to MPs.
- During physician review, the mean normalization adjustment was 2.9±1.9%, balancing CTV coverage and rectal sparing.
- All AP plans except one (minor sigmoid deviation) were considered clinically acceptable.
- Patient-specific QA confirmed excellent deliverability, with gamma pass rates above 99% (3%/2mm).

Conclusion

- The proposed AP pipeline enables fast, consistent, and clinically acceptable VMAT plan generation for ultra-hypofractionated prostate radiotherapy across different TPS and linac models.
- Integration of MRI-based urethral delineation is expected to further enhance OAR sparing.

- Fast generation of clinically acceptable VMAT plans.
  The automated pipeline produces plans in ~10 minutes (≈20 minutes including review) for Prostate cancer Radiotherapy, with nearly all plans deemed clinically acceptable.
- Improved target coverage with consistent plan quality.
  Demonstrates superior CTV coverage compared to manual plans, achieving target coverage in 20/20 cases under standardized criteria.
- Robust and clinically deliverable across systems.
  Validated across multiple TPS and linac models, with patient-specific QA confirming high deliverability (gamma pass rates >99%).

Structure	Dose Criteria	Manual Plans (mean + SD)	SmartPlan Plans (mean + SD)
PTV	D95% ≥ 34.4 Gy	35.48 ± 0.49	36.77 ± 0.18 *
	Dmax ≤ 120 %	37.93 ± 1.02	38.95 ± 0.71
CTV	D99% ≥ 36.25 Gy	35.98 ± 0.67	38.95 ± 0.71 *
Rectum	D50% < 8.1 Gy	9.49 ± 2.62	10.22 ± 1.69
	D20% < 29 Gy	19.19 ± 3.11	18.96 ± 2.7
	D1cc < 36.00 Gy	35.63 ± 0.67	37.41 ± 0.93 *
Bladder	D40% < 18.1 Gy	5.1 ± 4.2	6.59 ± 5.01 *
	D10% < 36.25 Gy	20.68 ± 6.68	22.32 ± 7.37
Left Femoral Head	D5% < 14.5 Gy	11.74 ± 1.52	11.96 ± 1.72
Right Femoral Head	D5% < 14.5 Gy	11.87 ± 1.49	11.71 ± 1.55

Table 1: Quantitative results Manual vs AutoPlan doses; with * are marked the statistically significant differences .

GUSTAVE ROUSSY CANCER CAMPUS GRAND PARIS · CENTRE DE LUTTE CONTRE LE CANCER LEON BERARD · Penn UNIVERSITY of PENNSYLVANIA · THERAPANACEA

Abstract Overview

Clinical Validation of AI Prostate SBRT Plans: Comparison to Manual Plans and Generalization From 36.25 Gy to 40 Gy

Authors: Nayoon J Lee, Anran Zhao, , Ledi Wang , Steven Philbrook , William R Green , Madalina-Liana Costea, Baris Ungun, Remi Vauclin, Elie Mengin, Nikos Paragyos, Rafe A McBeth

Purpose

- To evaluate generalization from the vendor-trained 36.25 Gy/5 regimen to the local 40 Gy/5 standard at UPENN hospital.

Mat&Methods

- Nr of patients: 23 treated on Halcyon Machine.
- Model training: prescription 36.25 Gy/5fx without urethra sparing.
- Quantitative criteira : institutional SBRT constraints including urethra sparing, bladder wall, rectal wall , femoral heads and penile bulb.

Results

- The AI plans were generated in less than 10 min and consistently used 2 arcs and less MUs, while the manual plans balistics varied between 2 and 3 VMAT arcs.
- The AI plans generalized to 40Gy regimen only through normalization, preserving target coverage and favorable doses to most of OARs.
- Constraints deviations were in structures not included in the model training: Urethra D0.035 cc limit at 40 Gy by a mean of ~101 cGy.

Conclusion

- SmartPlan produced clinically competitive plans with successful generalization across institutional practice variations, demonstrating improved planning efficiency. Two targeted areas for improvement have been identified: The integration of MR-based urethral contours into the planning workflow, and vendor model updates that explicitly optimize urethral constraints.

- Improves planning efficiency and consistency.
  Generating high-quality SBRT plans in <10 minutes with standardized delivery and fewer monitor units.
- Enables use across institutional practices.
  Demonstrating successful generalization from 36.25 Gy/5 to 40 Gy/5 while preserving target coverage.
- Supports clinical implementation with minimal customization.
  Achieving high constraint adherence (~93.5%) and identifying clear, addressable areas for improvement.

Plan Performance Summary Statistics

Metric	Constraint_Value	Plan Type / Constraint_Type	Mean_Dose AI (TheraAuto)	Mean_Dose Human (Initial)	Median_Dose AI (TheraAuto)	Median_Dose Human (Initial)	Min_Dose AI (TheraAuto)	Min_Dose Human (Initial)	Max_Dose AI (TheraAuto)	Max_Dose Human (Initial)	Average Constraint Delta AI (TheraAuto)	Average Constraint Delta Human (Initial)	Average Difference (AI - Human)	Pass_Rate AI (TheraAuto)	Pass_Rate Human (Initial)
PTV_4000_CI_Paddick_4000cGy	1	Min Dose (>=)	0.9	0.9	0.9	0.9	0.8	0.8	0.9	0.9	-0.1	-0.1	0.0	100.0%	100.0%
PTV_4000_D95%[cGy]	4000	Min Dose (>=)	4000.0	4000.0	4000.0	4000.0	4000.0	4000.0	4000.0	4000.0	0.0	0.0	0.0	100.0%	100.0%
PTV_4000_D99%[cGy]	3600	Min Dose (>=)	3885.7	3916.7	3889.7	3919.8	3750.4	3836.4	3932.4	3954.7	-285.7	-316.7	-31.0	100.0%	100.0%
BladderWall_D0.035cc[cGy]	4200	Max Dose (<=)	4089.8	4146.3	4090.0	4157.2	4046.8	4042.2	4122.9	4194.2	+110.2	+53.7	-56.5	100.0%	100.0%
BladderWall_D50cc[cGy]	1825	Max Dose (<=)	267.1	202.3	202.2	132.4	91.1	43.7	762.5	835.6	+1557.9	+1622.7	+64.8	100.0%	100.0%
External_Pat_Max[cGy]	4280	Max Dose (<=)	4240.0	4259.3	4240.0	4259.3	4240.0	4259.3	4240.0	4259.3	+40.0	+20.7	-19.3	100.0%	100.0%
FemoralHeads_D10cc[cGy]	3000	Max Dose (<=)	1273.0	1451.6	1231.4	1453.6	1095.3	1074.2	1578.8	2067.0	+1727.0	+1548.4	-178.6	100.0%	100.0%
PenileBulb_D0.035cc[cGy]	3150	Max Dose (<=)	1319.1	1360.9	732.4	873.1	231.4	174.5	3685.8	3128.2	+1830.9	+1789.1	-41.8	87.0%	100.0%
PenileBulb_D3cc[cGy]	2500	Max Dose (<=)	275.0	236.9	259.3	215.2	128.0	76.3	448.1	509.2	+2225.0	+2263.1	+38.1	100.0%	100.0%
PenileBulb_Mean[cGy]	1000	Max Dose (<=)	400.7	376.1	345.5	323.4	154.7	99.1	889.2	898.5	+599.3	+623.9	+24.6	100.0%	100.0%
RectalWall_D0.035cc[cGy]	4200	Max Dose (<=)	3934.7	3983.9	3971.4	4058.5	3709.0	3287.4	4079.0	4161.5	+265.3	+216.1	-49.2	100.0%	100.0%
RectalWall_D33%[cGy]	3900	Max Dose (<=)	1585.9	1437.1	1589.8	1340.3	1270.1	911.5	1996.5	2135.4	+2314.1	+2462.9	+148.8	100.0%	100.0%
RectalWall_D50%[cGy]	2400	Max Dose (<=)	1146.5	1015.7	1148.9	945.5	857.3	482.8	1424.5	1572.5	+1253.5	+1384.3	+130.7	100.0%	100.0%
RectumSuperior_D0.035cc[cGy]	3000	Max Dose (<=)	412.8	302.7	388.7	285.2	151.6	101.0	1480.0	771.5	+2587.2	+2697.3	+110.1	100.0%	100.0%
RectumSuperior_D10cc[cGy]	2500	Max Dose (<=)	152.9	94.6	160.2	98.1	78.6	41.4	234.6	164.5	+2347.1	+2405.4	+58.3	100.0%	100.0%
Skin_D0.035cc[cGy]	2000	Max Dose (<=)	1139.2	950.2	1117.1	951.5	973.8	671.5	1325.2	1090.6	+860.8	+1049.8	+188.9	100.0%	100.0%
SmallBowel_D0.035cc[cGy]	2900	Max Dose (<=)	123.1	76.2	94.5	52.2	34.9	15.2	447.1	300.3	+2776.9	+2823.8	+46.9	100.0%	100.0%
SmallBowel_D10cc[cGy]	1950	Max Dose (<=)	88.0	50.5	73.3	39.2	29.3	11.2	240.4	159.2	+1862.0	+1899.5	+37.5	100.0%	100.0%
Urethra+3mm_D0.035cc[cGy]	4200	Max Dose (<=)	4180.9	4132.3	4161.7	4135.1	4099.6	4066.4	4353.5	4189.0	+19.1	+67.7	+48.6	82.6%	100.0%
Urethra_D0.035cc[cGy]	4050	Max Dose (<=)	4150.8	4010.2	4132.9	4014.6	4092.3	3959.4	4326.6	4048.8	-100.8	+39.8	+140.6	0.0%	100.0%

Abstract Overview

Automated VMAT Planning for Prostate Radiotherapy on Halcyon System

Authors: Gorkem Gungor, Ana-Maria Gardareanu, Madalina-Liana Costea, Baris Ungun, Remi Vauclin, Elie Mengin, Nikos Paragios, Gilles Moliner

Purpose

- To test a commercially available, fully automated, end-to-end VMAT treatment planning pipeline for prostate radiotherapy on the Halcyon system, across multiple indications (6 dose prescriptions).

Mat&Methods

- Machine learning models were trained using data from multiple institutions in Europe.
- External validation was conducted on 30 patients across six different regimens, ranging from normo-fractionated to ultra-hypofractionated schedules (Configuration: 2 arc balistic for all the VMAT AP plans).
- Quantitaive evaluation was done comparing AP doses vs the manual plans.
- Qualitative review was done by one senior Rad Onc.
- A subset of plans underwent planQA measurements.

Results

- AP plans were created within 10–20 minutes, including TPS export and dose calculation.
- Using only two arcs, the AP pipeline consistently produced high-quality results, whereas clinical practice often employs additional arcs to meet objectives.
- 29/ 30 AP plans were deemed clinically acceptable; one was rejected due to a hotspot.
- Excellent deliverability with gamma pass rates >99% (3%/3 mm) and >95% (2%/2 mm).

Conclusion

- This study demonstrates the feasibility of a fully automated planning pipeline that generates high-quality, machine-deliverable VMAT plans for prostate cancer on the Halcyon system,with minimal human intervention.

- Delivers ready-to-treat VMAT plans, significantly reducing planning time and enabling a streamlined, low-touch workflow.
  10-20 minutes for the plan generation & 10 minutes for plan review by physician¤ total ~20–30 minutes.
- Ensures consistent, high-quality plans across multiple prescriptions.
  With 29/30 cases clinically acceptable and comparable to manual planning.
- Produces machine-deliverable treatments with high reliability.
  Demonstrating excellent QA results (>99% gamma pass) using a standardized two-arc approach.

	Number of plans that met the mandatory quantitative critera			
	Manual Planning		AutoPlanning	
Prescription	PTV criteria	OAR criteria	PTV criteria	OAR criteria
80 Gy SV in 40 Fx	8/8	8/8	8/8	8/8
80 Gy LN in 40 Fx	8/8	8/8	8/8	8/8
76 Gy in 38 Fx	4/4	4/4	4/4	4/4
74 Gy in 37 Fx	2/2	2/2	2/2	2/2
60 Gy in 20 Fx	3/4	4/4	4/4	4/4
36.25 Gy in 5 fx	4/4	4/4	4/4	4/4
Overall 30 cases	29/30	30/30	30/30	30/30

Table 1: Comparison of manual vs autoplanning in terms of number of cases that met the protocol criteria for each prescription.

GUSTAVE ROUSSY CANCER CAMPUS GRAND PARIS | CENTRE DE LUTTE CONTRE LE CANCER LEON BERARD | Penn UNIVERSITY of PENNSYLVANIA | THERAPANACEA