AI-Assisted Dose and Dose Rate Optimization for Deliverable Proton Conformal Flash Radiotherapy

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

Yin Gao[1], Andrew Friberg[1], Arnaud Pin[2], Rasmus Nilsson[3], Lingshu Yin[1], Constantinos Koumenis[1], Eric S. Diffenderfer[1], Jeffrey D. Bradley[1], Alexander Lin[1], Jennifer Wei Zou[1], Kevin Teo[1],
[1]Department of Radiation Oncology, University of Pennsylvania, Philadelphia, PA, [2]Ion Beam Applications SA, Louvain-La-Neuve, Belgium, [3]RaySearch Laboratories AB, Stockholm, Sweden

Abstract

Purpose: Proton conformal FLASH radiotherapy must simultaneously satisfy clinical dosimetric goals and deliver ultra-high-dose-rates under machine constraints. This highly coupled optimization problem is currently addressed through manual trial-and-error, limiting plan optimality and efficiency, and hindering clinical translation. This study aims to develop an AI-assisted Dose And dose-Rate opTimization (DART) framework to autonomously generate FLASH plans simultaneously meeting the dose and dose-rate criteria.

Methods: Fifteen head-and-neck reirradiation patients (40 Gy in 5 fractions) were retrospectively selected under IRB approval. Proton conformal FLASH planning was performed in RayStation-FLASH treatment planning system configured for an IBA ConformalFLASH delivery system. The DART framework integrated a Deep Reinforcement Learning (DRL)-based auto-planning agent and post-planning spot sequence optimization. The planning agent interfaced with RayStation-FLASH to autonomously guide inverse optimization. It received plan information, including dose-volume histograms and dose-rate-volume histograms, and output continuous parameter updates governing dose and dose-rate delivery. Training was conducted using a policy-based DRL approach guided by a composite reward function that jointly evaluated dosimetric and dose-rate performance. Following plan generation, raster scanning sequences were optimized using a genetic algorithm to maximize dose-rate. Plan integrity was evaluated via log-file-based dose and dose-rate reconstruction.

Results: The trained DRL agent autonomously operated RayStation-FLASH to generate clinically acceptable FLASH plans with clinically acceptable dose quality, achieving organs-at-risk sparing and reduced hotspots comparable to manual planning while substantially increasing high dose-rate coverage (>40GyRBE/s). Across all organs-at-risk, the V40GyRBE/s, V60GyRBE/s, V80GyRBE/s, and V100GyRBE/s improved by 12.40%, 9.40%, 8.47%, and 5.10% on average. The largest dose-rate gains were observed in the parotids, submandibular glands, and larynx, with V40GyRBE/s increases of 19.30%, 17.23%, and 10.21%. Post-planning spot sequence optimization further enhanced dose-rates without altering dose quality.

Conclusion: We developed and validated the AI-driven DART for proton conformal FLASH head-and-neck reirradiation, generating high-quality and deliverable plans with optimized dose and dose-rate.

Contact

Yin Gao, PhD
Department of Radiation Oncology,
University of Pennsylvania
Email: yin.gao@pennmedicine.upenn.edu

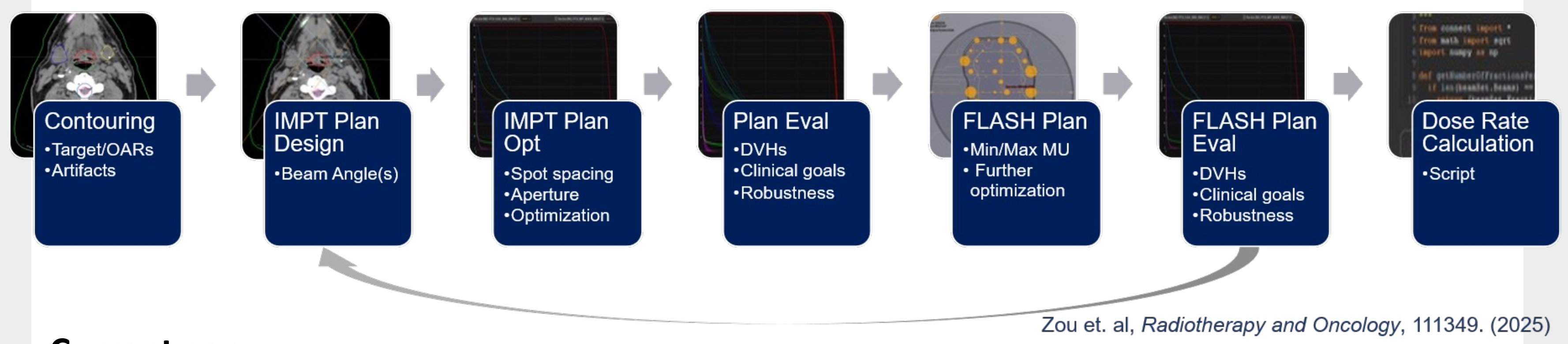
INTRODUCTION

Clinical benefit of proton conformal FLASH

- Combines conformal proton dose shaping with ultra high dose rates to potentially reduce normal-tissue toxicity.
- Especially valuable for head-and-neck (H&N) reirradiation, where targets often lie near previously irradiated critical organs.

Challenges to FLASH planning

- Dose quality and dose rate must be improved together; optimizing one may compromise the other.
- Multiple coupled parameters create a large, patient-specific solution space.
- Current planning requires repeated optimization, FLASH conversion, evaluation, and dose-rate calculation.
- Final solutions must also satisfy realistic machine and delivery constraints.



Zou et. al, *Radiotherapy and Oncology*, 111349. (2025)

Current gap

- Existing auto-planning methods mainly optimize or predict dose distributions.
- They cannot efficiently navigate the coupled dose and dose-rate–deliverability solution space.

Study objective

- Develop Dose And dose-Rate opTimization (DART), in which a deep reinforcement learning (DRL) agent learns through direct interaction with the TPS to replace manual trial-and-error and identify effective strategies for deliverable proton conformal FLASH planning.

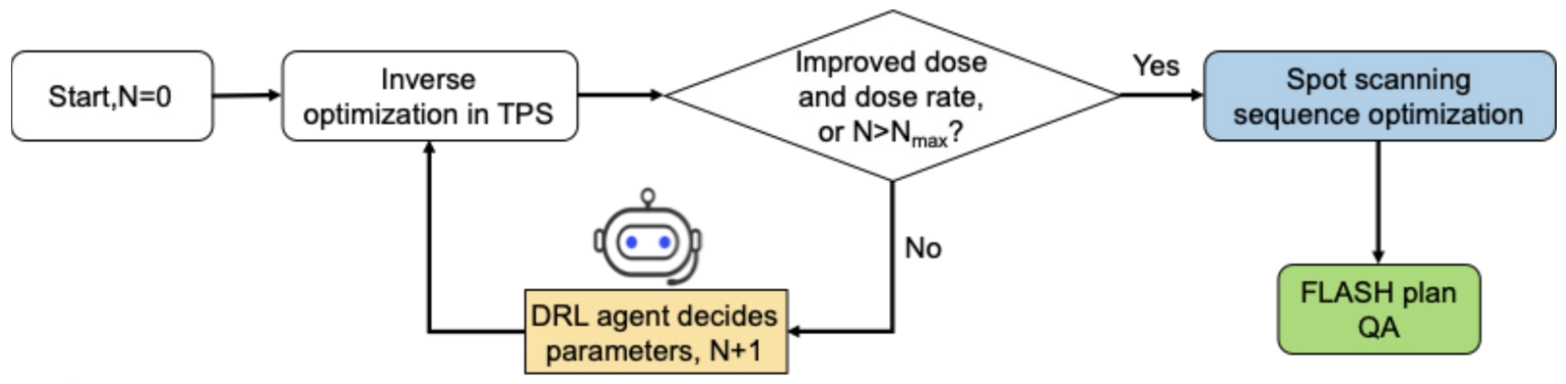
METHODS AND MATERIALS

Patient cohort and FLASH planning

- 15 H&N reirradiation patients, prescribed 40 Gy in 5 fractions.10 patients for training; 5 for testing.
- Clinical proton plans served as benchmarks. FLASH plans were generated in RayStation-FLASH for an IBA ConformalFLASH system using matched beam angles.

DART framework

- The DRL agent iteratively adjusted planning parameters through direct TPS interaction.
- After convergence, spot order was optimized and plan deliverability was evaluated.



DRL Planning Agent

At each iteration, the agent observed the current plan and proposed continuous parameter updates:

$$s_t = [\mathrm{DVH}_t,\ \mathrm{DRVH}_t,\ \mathbf{p}_t] \qquad \mathbf{p}_{t+1} = \mathbf{p}_t + \Delta\mathbf{p}_t$$

- **Input:** DVHs, DRVHs, and current planning parameters
- **Output:** Objective-weight and FLASH delivery-parameter updates
- The TPS reoptimizes the plan and returns updated dose and dose-rate feedback.

Dose and dose-rate reward

At each iteration, the agent observes the current plan and proposes continuous parameter updates:

$$S_t = S_{\mathrm{dose}} + \lambda S_{\mathrm{DR}} - P_{\mathrm{violation}}$$

$$r_t = S_{t+1} - S_t$$

- S_{dose}: target coverage and OAR sparing; S_{DR}: weighted DRVH coverage, emphasizing dose rates ≥40 GyRBE/s; $P_{violation}$: penalties for unacceptable dose metrics
- The agent is rewarded only when the new plan improves the overall dose and dose-rate trade-off.

Soft Actor–Critic learning

SAC balances exploitation of effective planning strategies with exploration of new parameter combinations:

$$\pi^* = \arg\max_{\pi} \mathbb{E}[\min(Q_1, Q_2) + \alpha\mathcal{H}(\pi)]$$

- **Actor network :** proposes continuous planning-parameter updates
- **Twin critics networks:** estimate their expected long-term benefit
- **Entropy term:** encourages exploration of the large solution space
- **Replay buffer and target networks:** improve training efficiency and stability

RESULTS

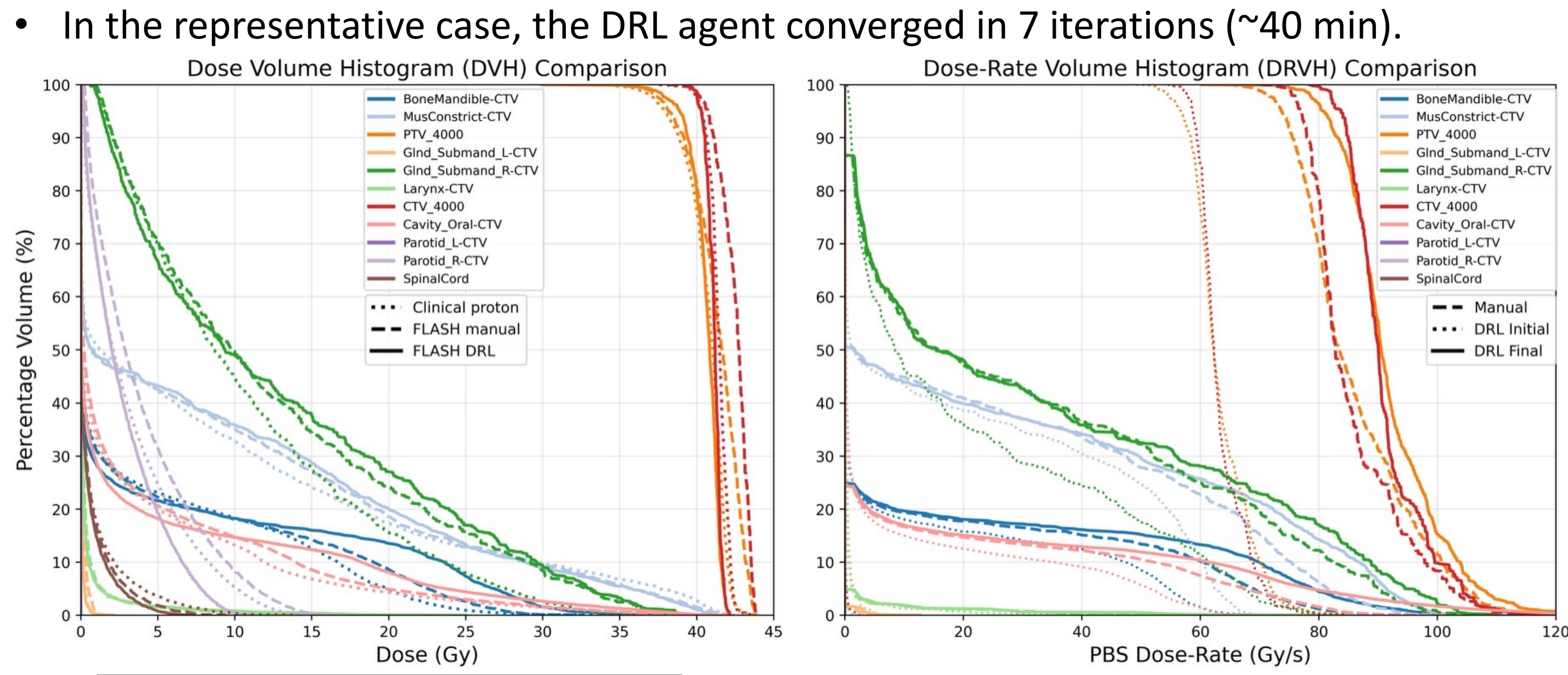
DART improved high dose-rate coverage

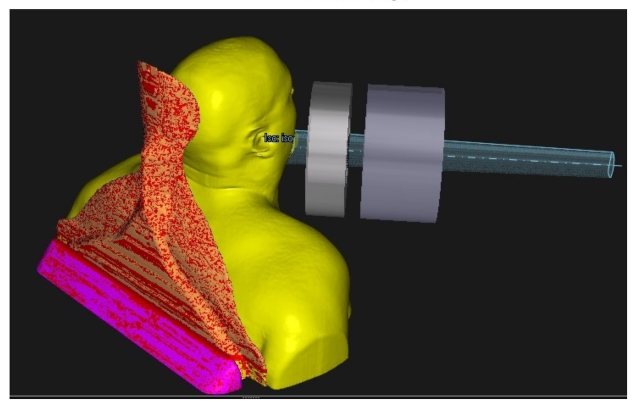
- Across all OARs, mean improvements were 12.40% for V40, 9.40% for V60, 8.47% for V80, and 5.10% for V100 GyRBE/s.
- The largest V40 gains occurred in the parotids (19.32%) and submandibular glands (17.24%).
- Notable V80 gains were observed in the larynx (18.00%) and constrictors (11.25%).

	$V_{40GyRBE/s}$ (%)	$V_{60GyRBE/s}$ (%)	$V_{80GyRBE/s}$ (%)	$V_{100GyRBE/s}$ (%)
Parotids	19.32	8.00	NA	NA
Bone Mandible	6.08	8.63	4.33	NA
Oral Cavity	3.69	5.27	7.00	1.00
Constrictor	8.22	11.14	11.25	2.00
Larynx	10.21	5.67	18.00	NA
Submandibular	17.24	12.18	6.24	7.50
All OARs (Mean)	12.40	9.40	8.47	5.10

DRL generated clinically acceptable FLASH plans

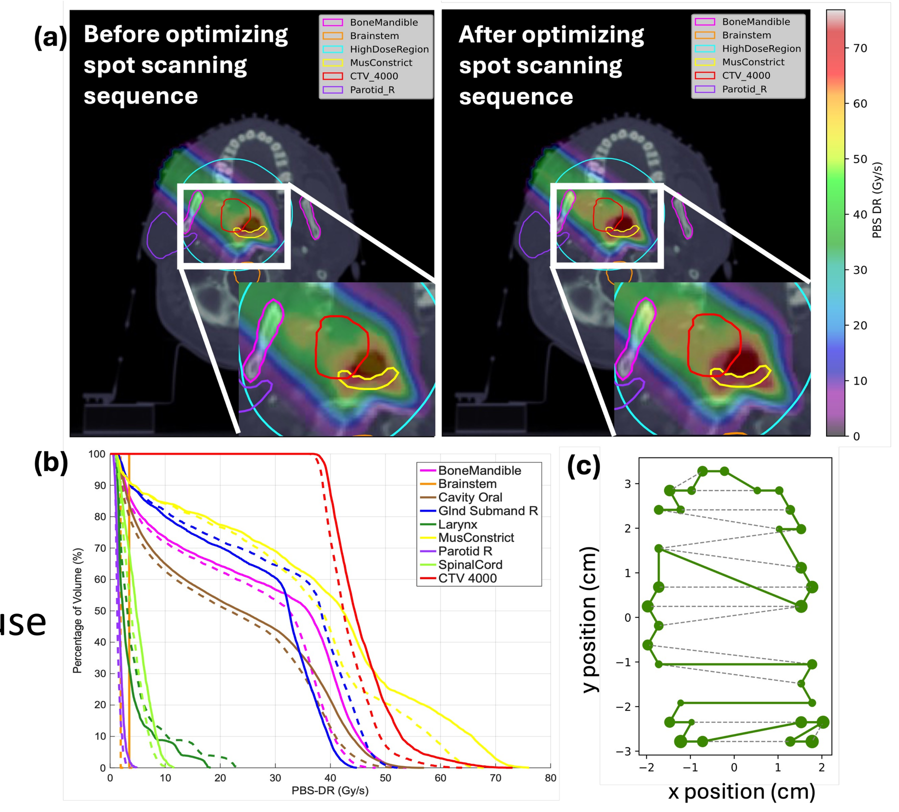
- Final plans met clinical dose goals with OAR sparing comparable to manual plans.
- In the representative case, the DRL agent converged in 7 iterations (~40 min).





Spot-sequencing added further gains

- Reordering the raster scanning path increased dose rate in nearby OARs.
- Dose distribution was preserved because spot weights remained unchanged.



CONCLUSIONS

Clinical Impact

- Developed for automated, deliverable proton conformal FLASH planning
- Maintains clinical dose quality while maximizing high-dose-rate

Technical Innovation

- First DRL-based framework to directly operate a TPS for joint dose & dose-rate optimization
- Spot sequence optimization further enhances delivery-level dose-rate