Feasibility of a Reinforcement-Learning Based Framework for Patient Specific Needle Placement and Dwell Time Optimization in Gynecologic HDR Brachytherapy

Gianna Hatheway, MS, Yang Sheng, PhD, Ryan Sanford, MS, Diandra Ayala-Peacock, MD, Junzo Chino, MD and Julie A. Raffi, PhD
Duke University Medical Center

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

ABSTRACT

Introduction:
Gynecologic high-dose-rate (HDR) brachytherapy is an effective treatment for locally advanced cervical and other gynecologic cancers, but treatment planning and interstitial needle placement remain technically complex and highly dependent on physician experience. Advances in computational optimization and improved training tools may help improve procedural consistency and planning efficiency. This work explores an approach to support hybrid intracavitary–interstitial gynecologic brachytherapy: the development of a reinforcement learning (RL) framework for automated needle placement and dwell time optimization

Materials and Methods:
A reinforcement learning framework was implemented in Python using an OpenAI Gym–compatible environment and the Proximal Policy Optimization (PPO) algorithm. Patient-specific anatomical information was extracted from CT imaging and RTSTRUCT contour data to generate anatomically feasible tandem and interstitial needle trajectories while avoiding organs at risk (OARs). Dose deposition was modeled using the TG-43 formalism for an ^{192}Ir HDR source. A continuous-action policy was used to optimize dwell time distributions using clinically relevant dose–volume metrics, including CTV_{HR} coverage (D90, D98), OAR sparing (D2cc for rectum, bladder, sigmoid, bowel, and vagina), and high-dose region control (V200). The framework was trained on four patient datasets and evaluated on three additional cases with a prescription dose of 600 cGy per fraction.

Results:
The reinforcement learning framework successfully generated anatomically feasible tandem and needle trajectories along with corresponding dwell time distributions for all evaluated patient cases. CTV_{HR} coverage was moderate across cases, with D90 values ranging from 457.5 to 740.1 cGy. High-dose regions remained controlled, and OAR sparing constraints were variable across all patients.
The reinforcement learning generated plans demonstrated a consistent attempt to maximize CTV_{HR} target coverage, with mean CTV_{D90} values of 519.7 ± 119.3 cGy. Organ-at-risk (OAR) doses showed moderate interpatient variability, with generally acceptable rectum, sigmoid, and bowel doses, but elevated bladder and vaginal doses in several cases. Reward values were heterogeneous (mean: −10,703 ± 16,767), with lower-performing cases corresponding to reduced high-dose coverage and increased constraint violations.

Conclusion:
This work demonstrates the feasibility of a reinforcement learning–based framework for patient-specific needle placement and dwell time optimization in gynecologic HDR brachytherapy The development of this RL framework provides a potential tool to support advancements in TRUS-guided gynecologic brachytherapy, offering a foundation for future research aimed at improving treatment planning automation, and procedural guidance.

Gianna Hatheway
Duke University Medical Physics Program
Gianna.Hatheway@duke.edu
330-316-8865

INTRODUCTION

- **Gynecologic high-dose-rate (HDR) brachytherapy** is an effective treatment for locally advanced cervical and other gynecologic cancers.
- **Treatment planning and interstitial needle placement** are technically complex and highly dependent on physician experience.
- **Computational optimization and improved training tools** have the potential to improve procedural consistency and planning efficiency.
- **This work explores** a reinforcement learning (RL) framework for automated interstitial needle placement and dwell time optimization in hybrid intracavitary–interstitial gynecologic brachytherapy.

METHODS AND MATERIALS

- **Developed** a reinforcement learning (RL) framework in Python using an OpenAI Gym–compatible environment and the **Proximal Policy Optimization (PPO)** algorithm
  - Proximal Policy Optimization (PPO) used to optimize HDR brachytherapy treatment plans through iterative trial-and-error learning
  - **Two custom OpenAI Gymnasium environments** developed:
    - **Tandem Environment:** optimizes tandem angle and tandem/ovoid dwell times.
    - **Needle Environment:** optimizes interstitial needle selection and dwell time adjustments
    - **Observations** include dosimetric metrics (CTV_{HR} D90, OAR doses), dose coverage metrics, dwell statistics, and isodose values
- **Generated patient-specific anatomy** from CT images and RTSTRUCT contours to create anatomically feasible tandem and interstitial needle trajectories while avoiding organs at risk (OARs).
  - This was done by defining an entry zone and a target zone (figure 2 below)
    - The **entry zone** was created from the vaginal lumen on a plane 2 cm inferior to the CTV_{HR} while the **target zone** was defined on a parallel plane shifted 20 mm superior to the CTV_{HR} with a 30% larger radius, allowing generation of anatomically feasible needle trajectories (≤45° from the vaginal canal axis and ≥5 mm from the tandem)
- **Modeled dose deposition** using the **TG-43** formalism for an **Ir-192 HDR** brachytherapy source.
- **Optimized dwell time distributions** with a continuous-action policy using clinically relevant dose–volume metrics, including **CTV D90** and **OAR D2cc**
- **Ran two separate environments with different reward functions**
  - **Training 1:** Reward function based on **predicted tandem and ovoid placement**; trained on **4 patient datasets** and evaluated on **3 independent cases**.
  - **Training 2:** Reward function based on **ground-truth tandem and ovoid placement**; trained on **17 patient datasets** and evaluated on **7 independent cases**.
  - **Hierarchical reward function** prioritized maximizing CTV_{HR}D90 and target coverage first, followed by minimizing dose to organs at risk (bladder, rectum, sigmoid, bowel, and vagina)

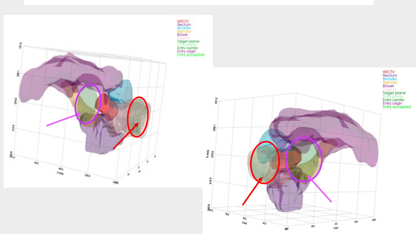

Figure 1 (above): Visual example of entry and target zone definition on a real patient anatomy from two different perspectives. The entry zone is circled in red, and the target zone is circled in pink

RESULTS

- **Training 1 – Predicted tandem + needle optimization (4 patients):** Achieved clinical **CTV_{HR}D90 (>600 cGy)** in **2/4 patients**, with a maximum D90 of **859 cGy**.
  - **Rectum, sigmoid, and bowel** constraints were met in all cases; **bladder** exceeded constraints in **2/4** patients and **vaginal** dose in **1/4**.
- **Training 2 – Ground-truth tandem + needle optimization (7 patients):** Using the **clinically placed tandem and ovoids** as the starting geometry, the RL agent optimized needle placement and dwell times for all applicators
  - Mean **CTV_{HR} D90 = 520 ± 119 cGy**, with an average **162.6 cGy improvement** over the baseline tandem-only plan.
  - OAR sparing remained variable, with **bladder and vaginal D2cc** most frequently exceeding planning objectives.
- **Comparison with clinical treatment plans:** Clinical plans achieved the **CTV_{HR}D90 goal in all 7 patients**, whereas the RL-generated plans met the target in **2/7** ground-truth cases.
- The RL framework selected **patient-specific needle numbers (1–6)** and tandem angles, demonstrating planning behavior similar to clinical decision-making

Parameter	Clinical Goals	Patient 1	Patient 2	Patient 3	Patient 4
CTV_{HR} D_{90}	> 600 cGy	859.26	661.40	585.96	457.47
Rectum D_{2cc}	< 390 cGy	273.65	146.16	208.60	289.91
Sigmoid D_{2cc}	< 390 cGy	388.01	353.31	109.12	177.44
Bladder D_{2cc}	< 440 cGy	470.57	472.02	371.12	382.19
Bowel D_{2cc}	< 340 cGy	153.97	193.34	267.87	67.127
Vagina D_{2cc}	< 680 cGy	489.625	641.258	556.56	759.18
Tandem Angle	15, 30, 45	15	15	45	15
# Needles	0-20	6	1	1	4

Table 1 (above): Dosimetric results of needle placement based on predicted T&O placement and reward function 1

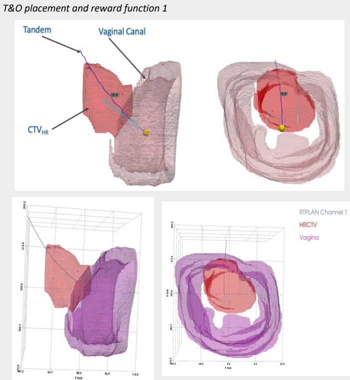

Figure 2 (above): Visual comparison of predicted tandem location (first row) to actual tandem location (second row)

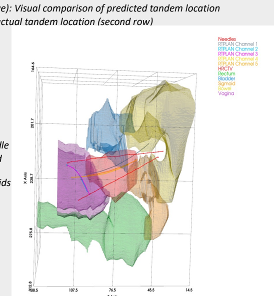

Figure 3 (right): Visualization of Patient #20 needle placement based on ground-truth tandem and ovoids (T&O)

RESULTS

Parameter	Clinical Goals	Patient 19	Patient 20	Patient 21	Patient 22	Patient 23	Patient 24	Patient 25
CTV_{HR} D_{90}	> 600 cGy	410.4	661.9	553.1	531.1	512.7	641.6	327.3
Rectum D_{2cc}	< 390 cGy	335.8	267.5	310.6	213.4	350.6	542.5	379.9
Sigmoid D_{2cc}	< 390 cGy	127.9	220	527.7	329.4	369.9	475.8	330.7
Bladder D_{2cc}	< 440 cGy	541.6	519.6	528.8	456.3	511.3	680.2	397.2
Bowel D_{2cc}	< 340 cGy	196.8	322.3	444.3	340.9	83.7	562.5	368.4
Vagina D_{2cc}	< 680 cGy	478.2	624.5	847.7	543.1	701.5	941.4	887.7
Reward	-	-21579.33	26.2	-1331.31	-2855.02	-4579.78	29.58	-44634.54
# Needles	0-20	2	1	2	1	4	1	1
# Steps	-	120	120	120	120	120	1	120

Table 2 (above): Dosimetric results of needle placement based on ground-truth T&O placement and reward function

Parameter	Clinical Goals	Patient 19	Patient 20	Patient 21	Patient 22	Patient 23	Patient 24	Patient 25
CTV_{D90}	> 600 cGy	602.6	739.6	605.9	703.4	716.6	888	609.6
Rectum D_{2cc}	< 390 cGy	259	405.9	309.9	261.6	307	244.6	462.3
Sigmoid D_{2cc}	< 390 cGy	187.9	166.3	398.3	327.3	298.3	380.5	468.3
Bladder D_{2cc}	< 440 cGy	487.6	441.1	508.4	446.6	524.1	542.1	443
Bowel D_{2cc}	< 340 cGy	277.2	250.1	364.2	319.8	82	313.2	541.7
Vagina D_{2cc}	< 680 cGy	489	369.6	665.9	545.4	664.6	1060.2	791.7
# Needles	-	2	2	2	2	2	2	4
Tandem Angle	15, 30, 45	15	30	30	30	30	45	30

Table 3 (above): Clinical Results for Comparison

DISCUSSION

Training 1 – Predicted tandem + needle optimization:
Demonstrated the feasibility of using reinforcement learning to **jointly predict applicator geometry, needle placement, and dwell times** from patient anatomy.
Successfully generated treatment plans while meeting **most** OAR dose constraints, **although target coverage and bladder sparing remained inconsistent**.

Training 2 – Ground-truth tandem + needle optimization:
Using the **clinically placed tandem and ovoids** to isolate the performance of the needle optimization algorithm.
RL improved CTV_{HR} D90 over the baseline tandem-only plans but **did not consistently achieve the level of target coverage obtained in clinical treatment plans**.
Results suggest the **current reward function does not fully capture the clinical priority of maximizing target coverage before optimizing OAR sparing.**

Comparison with clinical plans:
Clinical plans achieved more consistent target coverage while balancing OAR dose through expert-guided optimization.
The RL framework demonstrated adaptive, patient-specific selection of tandem angles and needle configurations, indicating a primitive learning of anatomy-dependent planning strategies.
Further improvements to the **reward function, episode termination criteria, and training dataset size** are expected to improve coverage consistency and overall clinical performance.

CONCLUSIONS

This work presents the development and implementation of a reinforcement learning framework for patient-specific needle placement and dwell time optimization in gynecologic HDR brachytherapy. **While clinical objectives were achieved with highly variable success, the framework demonstrates the feasibility of generating anatomically informed needle configurations using only RTSTRUCT-derived data. Additionally, the model shows promise in predicting tandem placement without explicit knowledge of the cervical os location.**
These findings highlight the potential of reinforcement learning as a tool to enhance TRUS-guided gynecologic brachytherapy. This approach establishes a foundation for future work focused on improving treatment planning automation, real-time procedural guidance, and physician training.