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

Automate Radiotherapy Patient Scheduling with Large Language Model

Eric Zhang[1,2], Wen Li[1], Annette Souranis[1], Georgia Paparoidamis[1], Michael B.Roumeliotis[1], Xun Jia[1]

[1] Department of Radiation Oncology and Molecular Radiation Sciences, Johns Hopkins University School of Medicine, Baltimore, MD
[2] Westlake High School, Westlake Village, CA

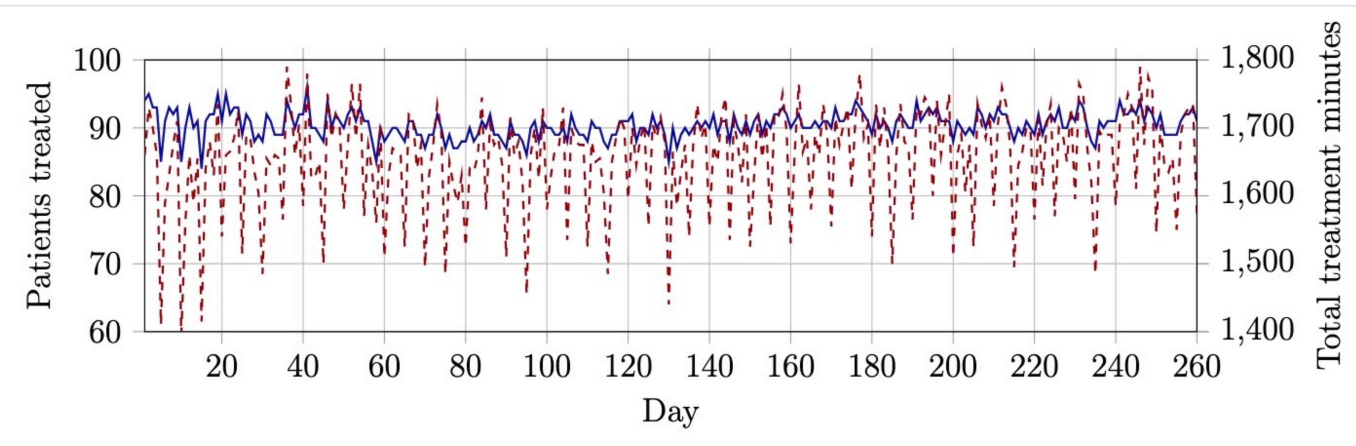
JOHNS HOPKINS MEDICINE
RADIATION ONCOLOGY & MOLECULAR RADIATION SCIENCES

Background and Study Goal

- RT scheduling must coordinate multi-fraction continuity, machine eligibility, first-fraction workflows, BID spacing, anesthesia timing, and clinic capacity.
- Manual scheduling is labor-intensive and difficult to adapt when patient volume or operational priorities change.
- Goal: evaluate whether an LLM can generate feasible RT schedules under realistic clinical and operational constraints.

Simulated Clinical Environment

- Three LINACs; 7:00 AM–5:00 PM operating window.
- One-year simulation with 260 treatment days and 1,400 synthetic new patients generated using clinically derived category proportions.
- 12 treatment categories with category-specific scheduling requirements.
- Daily workload matched typical practice: 84–96 patients/day; mean total treatment time ≈1,655 min/day.



Stable simulated patient volume and total scheduled treatment time across the 260-day study period

ChatGPT-Based Scheduling Framework

- Reusable prompt encoded hard rules, feasibility handling, scheduling objectives, and output requirements.
- Day-specific prompt supplied current schedule state, new patients, scheduling horizon, and scenario objective.
- ChatGPT 5.4 generated daily machine schedules and patient-level schedules without access to an external solver or code execution.
- Outputs were checked by deterministic Python validation and manual spot review.

Clinical rules → Daily inputs → ChatGPT schedule → Validation

Evaluation Scenarios

S1: weekly time consistency using a soft 60-minute window.

S2: add LINAC continuity to reduce unnecessary machine switching.

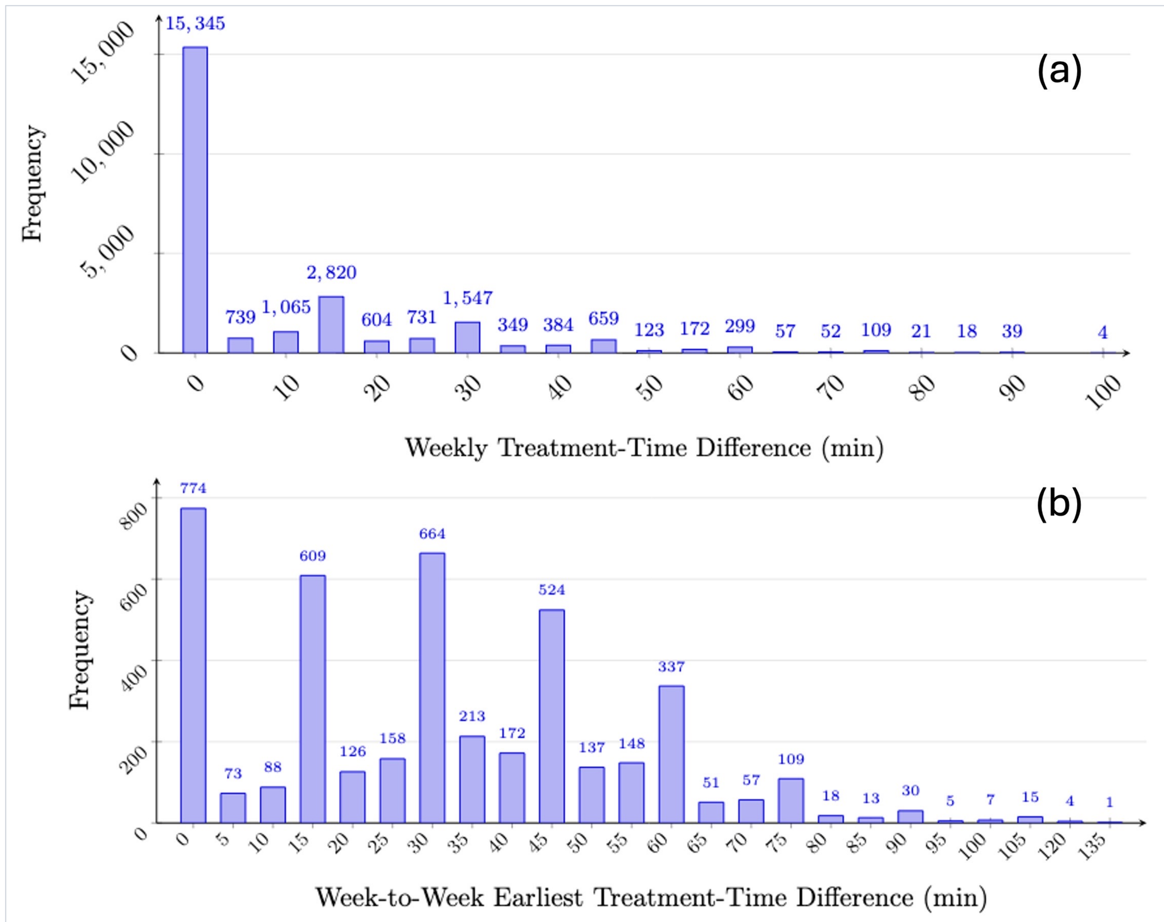
S3: relax time consistency only when preserving it creates >20-minute idle gaps.

S4: test infeasible requests and conflict reporting.

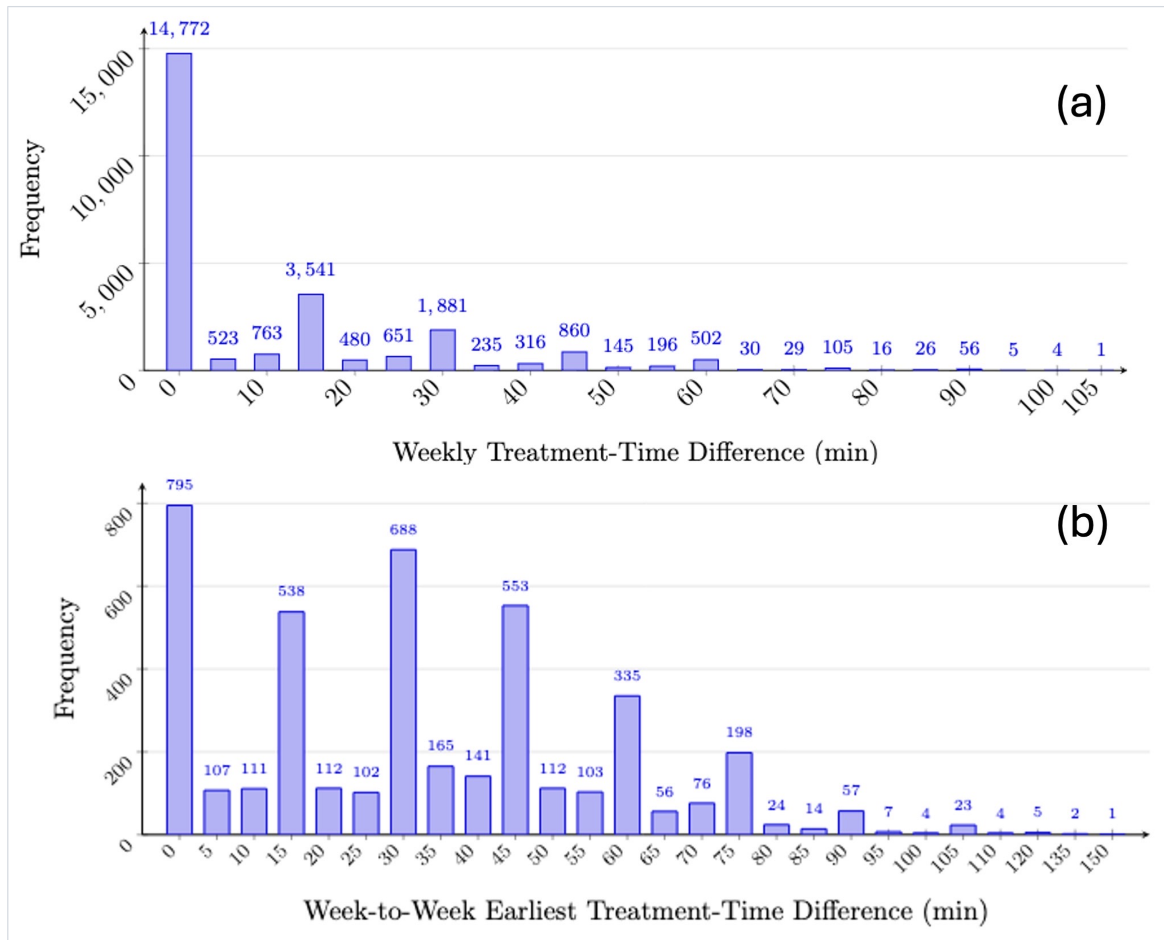
Results

≈99% fractions within the preferred 60-min weekly window	**54.6% → 10.1%** machine switching after adding LINAC continuity	**169.5 → 89.2 min** Friday mean idle gap after gap-constrained relaxation	**10/10** infeasible request scenarios correctly detected

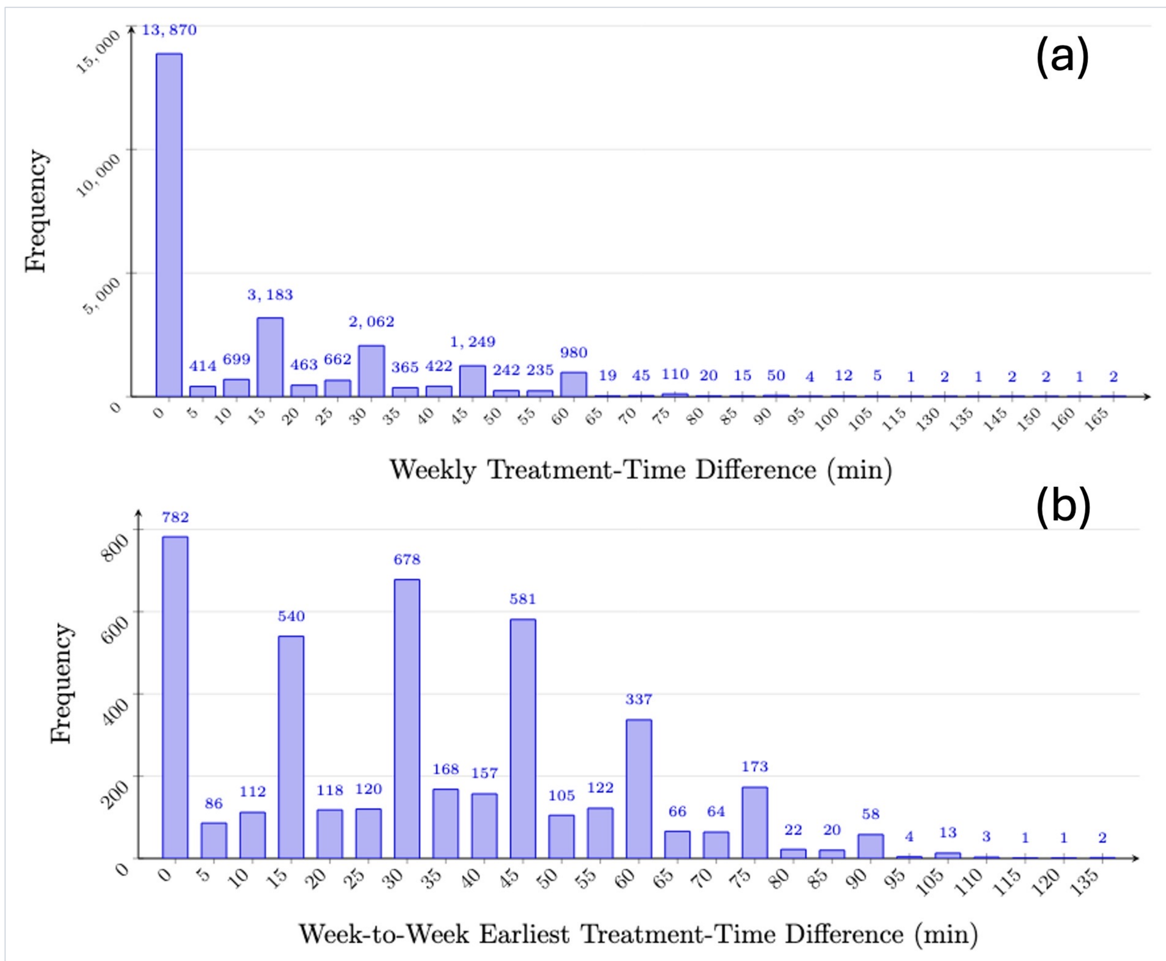
- Across all feasible scenarios, the LLM-generated schedules satisfied the predefined feasibility rules.
- Baseline weekly time consistency was strong: 24,837 of 25,137 fractions (98.8%) remained within the preferred 60-minute weekly window.
- Adding a LINAC-continuity objective reduced the machine switching rate from 54.6% to 10.1% with minimal effect on time consistency.
- Allowing gap-constrained temporal relaxation improved Friday compactness, decreasing the Friday mean idle gap from 169.5 to 89.2 minutes while maintaining strong overall consistency.



Baseline weekly time consistency



After adding LINAC continuity



Gap-constrained temporal relaxation

Performance Summary

Metric	Baseline	+ continuity	+ relaxation
≤60-min weekly window	98.8%	98.9%	98.8%
week-to-week ≤60 min	92.8%	89.1%	90.1%
machine switching	54.6%	10.1%	12.0%
Friday mean gap	159.6 min	169.5 min	89.2 min
overtime days	3	2	3

Summary & Conclusion

- The ChatGPT-based scheduling framework generated feasible candidate radiotherapy schedules while adapting to different patient- and clinic-centered objectives.
- Structured natural-language prompting allowed clinic priorities to be modified without redesigning the core framework.
- LLM-assisted scheduling appears feasible as a flexible, human-in-the-loop decision-support approach.
- The study supports continued investigation of LLM-based RT scheduling under realistic clinical constraints.