Evaluating Vision Language Model for Autonomous Offline Adaptive Radiotherapy Decision Support in Head and Neck Cancers

K Verdecchia, B Luo, C Li, H Nusrat, A Doemer, B Movsas, F Siddiqui, K Thind
Henry Ford Cancer Institute, Detroit, Michigan

AAPM 2025
JULY 27–30 | WASHINGTON, DC
67TH ANNUAL MEETING & EXHIBITION
COMING TOGETHER TO FORGE AHEAD IN MEDICAL PHYSICS

IMPACT/INNOVATION

Implementation of Vision-Language Models (VLMs) within the clinic is advantageous due to their potential to guide clinical decision-making without requiring model training.

In this work, an internally developed framework was utilized to automatically query clinical data and prompt a VLM requesting an assessment for the need for treatment adaptation. Retrospectively collected data included daily treatment localization CBCT images and patient weight variations from 60 head-and-neck cancer patients who had undergone a 35-fraction radiation therapy treatment course.

For each treatment fraction, multiple CBCT slices were compared to corresponding slices from the first-fraction CBCT. These imagesets were provided to the VLM alongside a structured prompt that incorporates additional clinical data, describes the input images, and requests a decision regarding whether offline adaptation should be initiated based on observed anatomical changes.

This framework enables a "plug-and-play" approach, allowing integration of advancing VLMs to improve robustness over time. Early results demonstrate promising potential for autonomous guidance of adaptive treatment decisions.

CONTACT

Kyle Verdecchia
Henry Ford Health
2799 W. Grand Blvd
Detroit, MI 48202
kverdec1@hfhs.org

HENRY FORD HEALTH

INTRODUCTION

Current triggers for automated offline adaptive radiotherapy (ART) often function as "black boxes," providing limited transparency into the reasoning behind a decision. Vision-Language Models (VLMs) offer a promising solution by enabling a clearer rationale regarding the decision.

Offline ART for head-and-neck (HN) patients has been shown to provide dosimetric benefits through mid-course treatment replanning[1,2]. Additionally, artificial intelligence approaches demonstrated value for identifying anatomical deformations within HN patient cohorts[3, 4].

Building on this, we propose that a VLM integrated within an internally developed automated framework can analyze anatomical changes observed in daily cone-beam computed tomography (CBCT) images, alongside patient weight variations, to flag patients who may benefit from offline adaptive treatment planning. A key advantage of incorporating a VLM is that these models do not require training and can be readily deployed as a plug-and-play solution within our framework, with performance expected to improve as underlying VLMs continue to advance.

Ultimately, this study investigates an autonomous framework that utilizes a VLM to flag candidates for offline ART in HN cancer.

METHODS

An autonomous framework that queries the oncology information system for volumetric CBCT images and longitudinal patient specific clinical covariates (weekly weight and image reconstruction diameter). For every treatment three axial slices are extracted from the daily pretreatment CBCT and the baseline first CBCT. The three slices were spatially isolated at isocenter and at ±18 mm longitudinally.

These clinical data inputs were synthesized into a templated prompt that instructs the VLM model (Llama4-Maverick) to quantify external anatomical changes and generate a binary decision for adaptation (Figure 1); a sampling temperature of 0.5 was used; ranging between 0 and 1 where lower values are more deterministic.

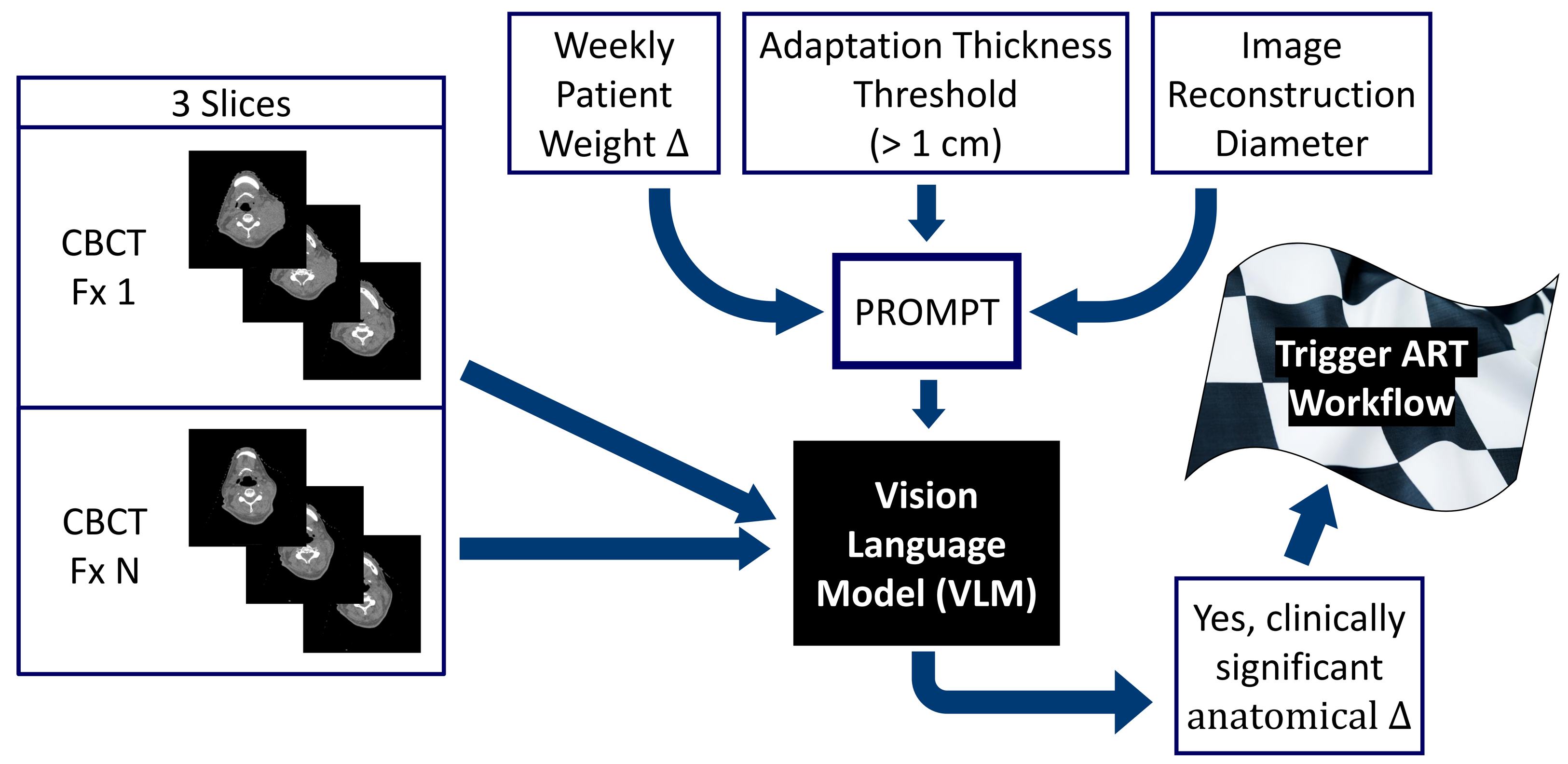

Figure 1: Illustrates a flowchart of the designed framework that automatically retrieves CBCT image slices and patient clinical context to feed the VLM, requesting a decision regarding offline-ART.

The VLM was instructed to flag ART when thicknesses exceeded 1 cm. Mean adaptation trigger point determined by the model was compared to the historical clinical mean using a paired two-tailed t-test.

The templated prompt includes categories and text paraphrased here:

- Image Description

Brief description of the acquired images fed into the VLM; relative patient weight included here.

- Goal to Achieve

Instructions for the VLM to evaluate and calculate external contour differences.

- Focused Image Context

Quantitative dimensions of the images fed into the VLM, with emphasis directing the focus of the VLM to external contour differences.

- Confidence Table for Adapting (Table 1)

Confidence [%]	External Contour Differences	Effected Quadrants	Magnitude [cm]	Adapt?
0	None	All	<0.4	No
10	Very small	1 or 2	0.4 to 0.8	No
20	Small	1 or 2	0.8 to 1.2	No
30	Very small	3 or 4	0.4 to 0.8	No
40	Small	3 or 4	0.8 to 1.2	No
50	Medium	1 or 2	1.2 to 1.6	No
60	Medium	3 or 4	1.2 to 1.6	No
70	Large	1 or 2	1.6 to 2.0	No
80	Large	3 or 4	1.6 to 2.0	No
90	Very large	1 or 2	>2.0	Yes
100	Very large	3 or 4	>2.0	Yes

Table 1: Confidence Table for Adaptation Provided in the VLM Prompt

METHODS (cont.)

- Output Summary

Requesting succinct output by the VLM in tabular format regarding the following:

1) Additional comments worth noting.
2) Average magnitude of tissue difference in cm, where negative represents a decrease and positive represents an increase.
3) What is your confidence for adapting?
4) Do we adapt (yes or no)?

RESULTS

To optimize the VLM, robustness tests were completed for the following inputted parameters; minor differences in output was observed as these parameters were varied (data not shown):

- longitudinal slice resolution (18 mm)
- window width (500 HU)
- sampling temperature hyperparameter (0.5)

N = 60 patients treated by a 35-fraction radiotherapy treatment course.

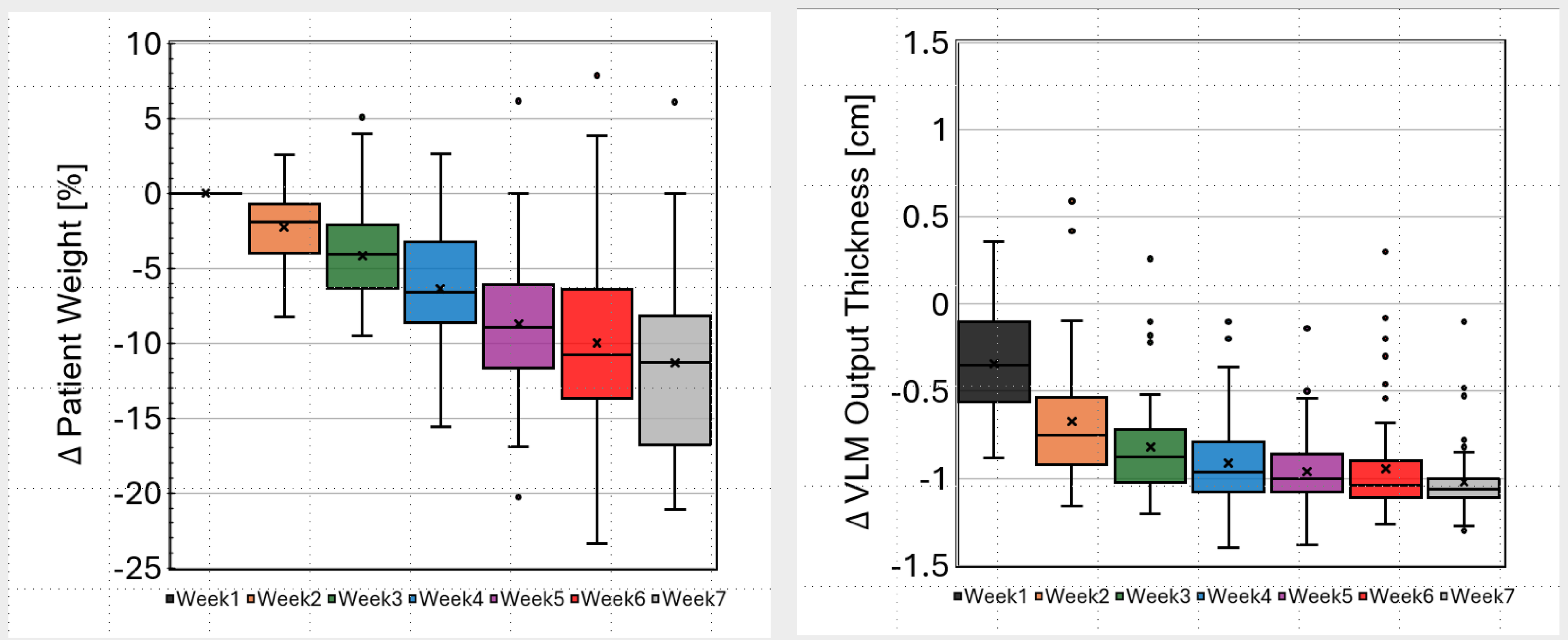

Figure 2: Weekly averages of the 60 patients are presented for (A) the change in external contour thickness estimated by the VLM and (B) the relative weight change percentage recorded in their chart. The bar represents the median and the cross represents the average, with outliers included as circles.

- Adaptation was flagged after 18.4 ± 1.3 fractions and significantly differed ($p<0.01$) from that performed clinically 23.8 ± 0.4 fractions.
- Seven patients were not flagged for adaptation by the VLM that were clinically adapted.
- The max thickness at adaptation was manually measured to be -1.5 ± 0.1 cm.

DISCUSSION

- Artificial intelligence has the capability to guide on-treatment clinical decisions within radiation oncology.
- As expected, on average, a decrease in patient weight were observed weekly, and decreases in the external CBCT images were observed weekly. Also, progressive decreases to the CBCT external contour were identified by the system as the treatment course progressed. Currently, the prompt focusses mostly on image analysis and external changes but increasing emphasis within the prompt on the weight change may be worthwhile investigating further.
- A limitation with the current system/prompt is flagging for internal tissue changes, as some of the patients that were adapted for clinically were not flagged by the VLM; further individualized analysis is required.
- Including HN patients that were not clinically adapted, and expanding the cohort beyond HN patients will also be essential to refine system performance and robustly characterize sensitivity and specificity across disease sites.
- To summarize, further investigations on an individualized basis, prompt refinement, and system robustness are required prior to clinical implementation.

CONCLUSION

This study establishes the feasibility of a VLM-based framework as an explainable, automated surveillance tool for offline ART. The proposed system demonstrates the ability to provide autonomous and interpretable decision support, with results showing early identification for offline adaptation compared to historical clinical scenarios.

REFERENCES

1. Schwartz DL, Garden AS, Thomas J, et al. Adaptive radiotherapy for head-and-neck cancer: Initial clinical outcomes from a prospective trial. Int. J. Radiat. Oncol. Biol. Phys. 2012; 83(3): 986–993.
2. Schwartz DL, Garden AS, Shah SJ, et al. Adaptive radiotherapy for head and neck cancer - Dosimetric results from a prospective clinical trial. Radiother. Oncol. 2013; 106(1): 80–84.
3. Guidi G, Maffei N, Vecchi C, et al. A support vector machine tool for adaptive tomotherapy treatments: Prediction of head and neck patients criticalities. Phys. Medica 2015; 31(5): 442–451.
4. Guidi G, Maffei N, Meduri B, et al. A machine learning tool for re-planning and adaptive RT: A multicenter cohort investigation. Phys. Medica 2016; 32(12): 1659–1666.