Language-Guided Agentic AI Framework for Image Registration in Prostate SBRT Using Model Context Protocol

Anran Zhao[1] and Rafe McBeth[1]
[1]Department of Radiation Oncology, University of Pennsylvania

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

ABSTRACT

Purpose: To develop and validate an agentic AI framework using Model Context Protocol (MCP) that enables clinicians to control CT-MR image registration through natural language guidance. The study aims to determine if text-based clinical intent produces predictable, structure-specific alignment outcomes for prostate stereotactic body radiation therapy (SBRT) treatment planning.

Methods: An MCP server was developed providing specialized registration tools for a Large Language Model agent to invoke dynamically. The framework utilizes natural language parsing to identify focus structures. For this study, three prostate SBRT patients with pre-existing prostate, bladder, and rectum contours on both CT and MRI were used. The agentic framework calls MCP tools to perform a sequence of operations: an initial global Mutual Information (MI) alignment for baseline registration, followed by structure-specific Dice Similarity Coefficient (DSC) optimization via a weighted grid search using the available contours. Cases were processed under four language-driven conditions: one with no specific guidance (baseline MI-only), and three instructing the agent to align specifically to the prostate, bladder, or rectum.

Results: Each alignment method consistently produced the highest DSC for its intended target structure across all patients. Target structure DSC improved by 5-11% over baseline MI registration. The largest improvement occurred for bladder-focused alignment, while smaller improvements were noted when baseline registration already showed high overlap. Non-targeted structures showed expected trade-offs in alignment, consistent with rigid registration physics. Structure-specific refinements resulted in a 1-4 mm shift from baseline translation, confirming that different language instructions produce unique registration transforms.

Conclusion: This work demonstrates a consistent relationship between natural language guidance and registration outcomes. The framework allows clinical teams to achieve anatomical alignment goals through conversational interaction rather than manual parameter adjustment. This MCP-based architecture serves as a foundation for more complex autonomous workflows, where similar natural language interfaces could be utilized for radiotherapy treatment planning.

CONTACT

Anran Zhao
Department of Radiation Oncology
University of Pennsylvania
3400 Civic Center Blvd
Philadelphia, PA 19104
anran.zhao@pennmedicine.upenn.edu
(804) 300-1952

INTRODUCTION

- Stereotactic body radiation therapy (SBRT) delivers high radiation doses, needing precise alignment to protect organs at risk
- Prostate SBRT treatment planning requires CT scans for dose calculation and MR scans to see soft tissue
- Image registration is currently performed manually, requiring manual adjustments
- Model Context Protocol (MCP) framework lets LLM agents select and run specialized software tools for image registration [1]
- This study aims to determine if text-based clinical intent produces predictable, structure-specific alignment outcomes

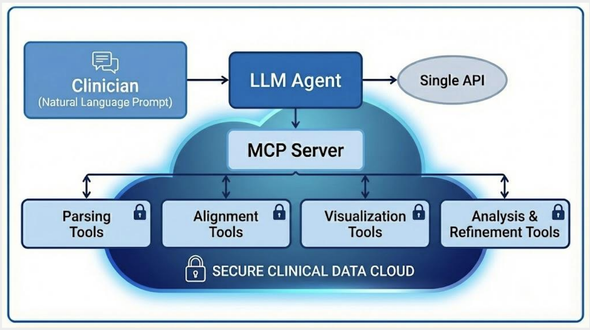

Figure 1. MCP Workflow Architecture Visualization

METHODS AND MATERIALS

- Built an MCP server holding specialized registration tools that an LLM agent can invoke dynamically
- Framework utilizes natural language parsing to automatically identify focus structures from text instructions
- Evaluated using data from three prostate SBRT patients with existing prostate, bladder, and rectum contours
- Performed an initial global Mutual Information (MI) alignment to establish a baseline match
- Followed with structure-specific Dice Similarity Coefficient (DSC) optimization via a weighted grid search using available contours
- Processed cases under four text conditions: baseline MI-only, "Align by Prostate", "Align by Bladder", and "Align by Rectum"

RESULTS

- Each text directive consistently produced the highest DSC for its intended target structure across all patients (Table 1)
- Target structure DSC values improved by 5% to 11% over the baseline global registration
- Structure-focused refinements caused a 1 to 4 mm translation shift from baseline alignment [2]
- Aligning specifically to the prostate boosted its target DSC by 0.057 over baseline [3]
- Bladder-focused alignment yielded a DSC improvement of 0.063
- Rectum-focused alignment improved rectum DSC by 0.038
- Target-specific optimization completed in 130 to 150 seconds

Table 1. Mean DSC results for structures across all patients (n = 3)

	Prostate DSC	Bladder DSC	Rectum DSC
Baseline (MI)	0.768	0.573	0.755
Align by Prostate	**0.825**	0.614	0.757
Align by Bladder	0.809	**0.636**	0.745
Align by Rectum	0.794	0.591	**0.793**

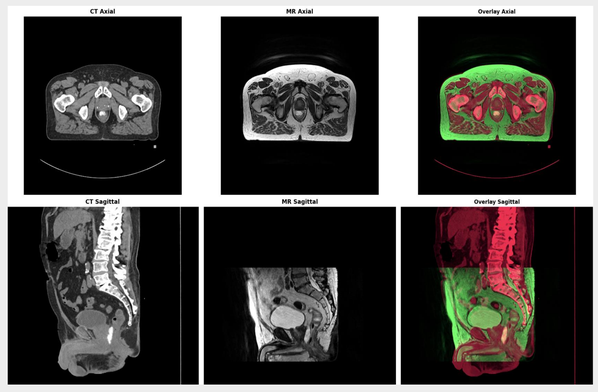

Figure 2. Language-guided mutual information image registration

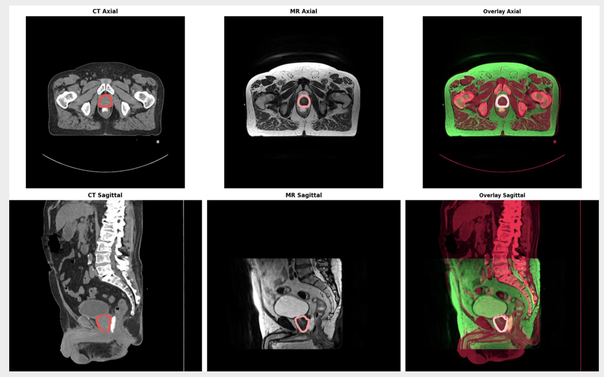

Figure 3. Language-guided prostate-focused image registration

DISCUSSION

- Establishes a foundation where an AI agent can directly execute physician image registration instructions
- Global MI tools naturally align to high-contrast bones, while text commands shift priority to soft-tissue targets
- 3D contour overlays show that focusing on one organ naturally decreases overlap for nearby structures [4]
- The modular, tool-calling framework is site-agnostic, allowing it to scale to any treatment site

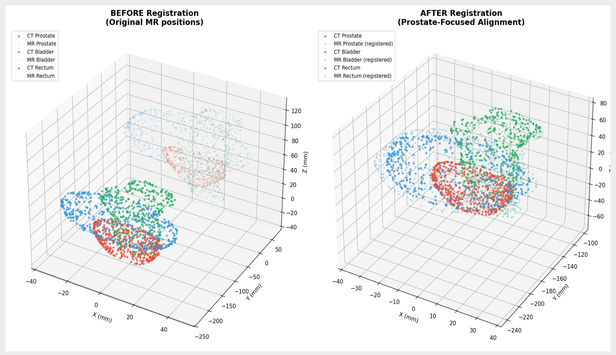

Figure 4. 3D visualization of all structure alignment from prostate-focused registration

CONCLUSIONS

- Successfully translated clinical intent into execution, shifting registration from manual parameter tuning to an automated, natural-language workflow
- Text commands predictably drove consistent anatomical alignments through MCP, creating a scalable framework for fully automated planning