CLARITY: Medical World Model for Guiding Treatment Decisions by Modeling Context-Aware Disease Trajectories in Latent Space

Tianxingjian Ding, Yuanhao Zou, Chen Chen, Mubarak Shah, Yu Tian
Institute of Artificial Intelligence, University of Central Florida

Personal Page | Project Page

ECCV 2026 | MALMÖ | SEPT 8–12

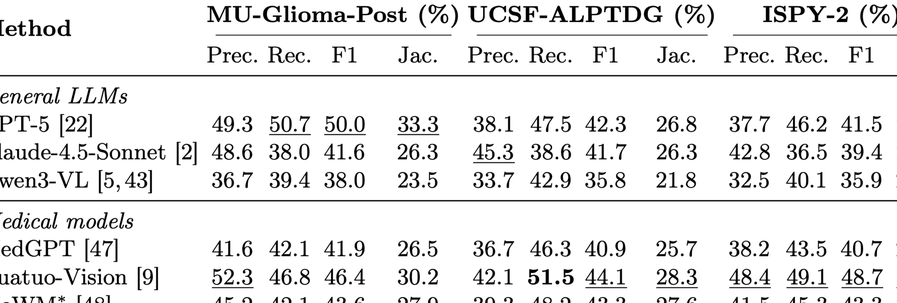
Purpose

Motivation:

■Why World Models?

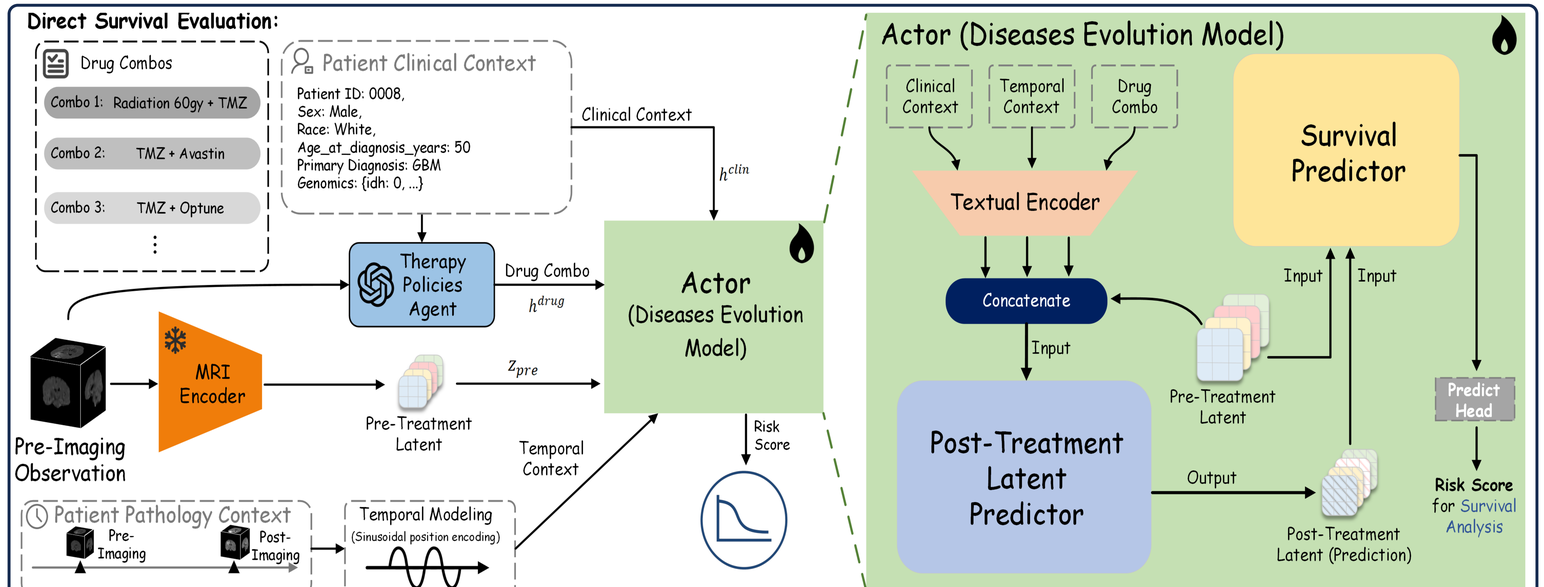
- Clinical treatment decisions are ***sequential*** rather than *static*.
- World models model disease trajectories under **different treatments.**
- Treatment planning requires **long-horizon** evaluation of candidate interventions.

Why predict a *Static* outcome When you can plan a *Dynamic* one?

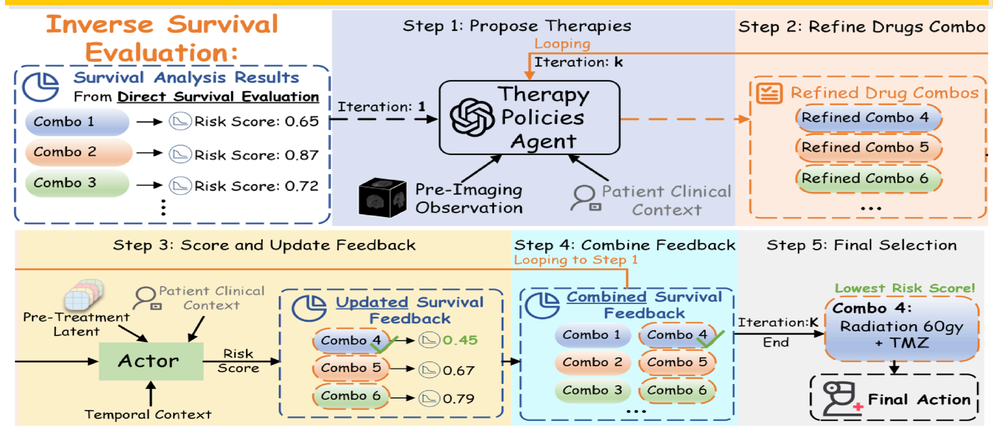


1.Therapy Policy Agent: ***Generates therapy proposals from multimodal patient inputs.***
2.Diseases Evolution Model : ***Simulates treatment-conditioned disease trajectories directly.***
3.Inverse Survival Evaluation: ***Iteratively optimizes therapies via survival-driven feedback.***

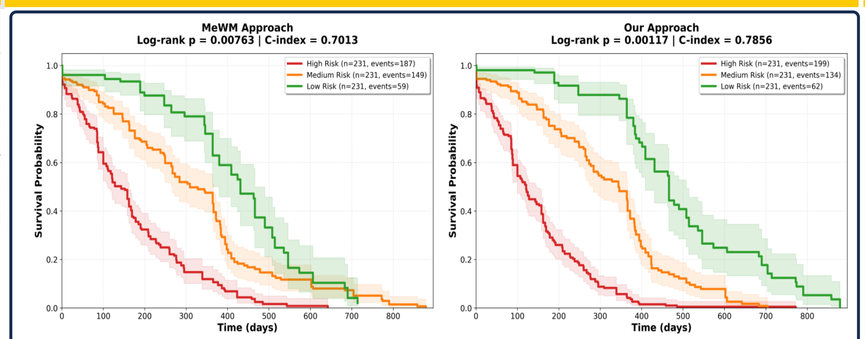
Methodology - Direct Survival Evaluation



Methodology - Inverse Survival Evaluation

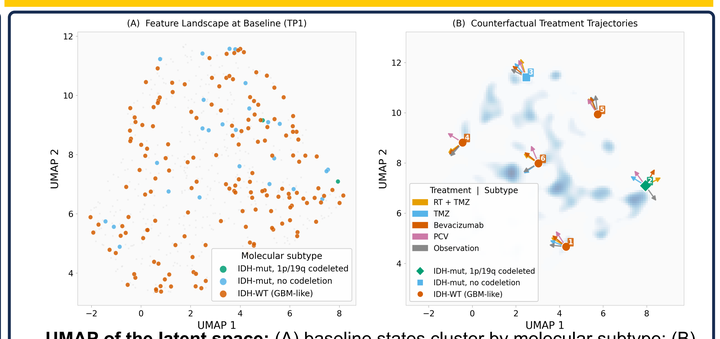


Survival Analysis



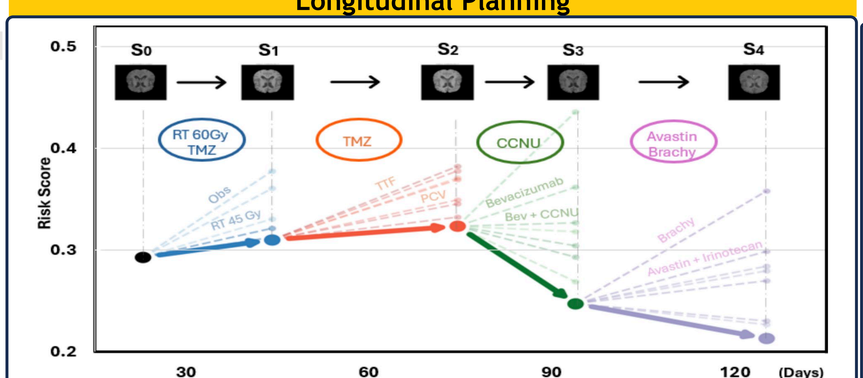
Kaplan-Meier curves: diffusion-based MeWM* vs CLARITY. C-index 0.794.

Latent trajectory visualization



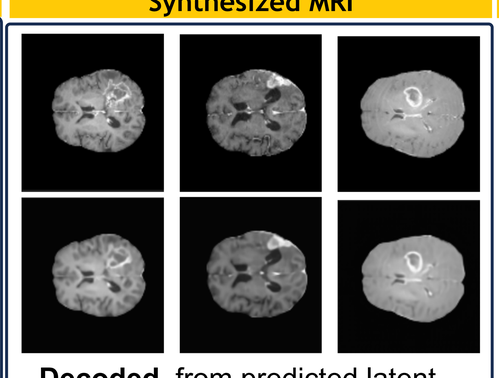
UMAP of the latent space: (A) baseline states cluster by molecular subtype; (B) counterfactual therapy rollouts form distinct, smooth trajectories.

Longitudinal Planning



- **At each MRI follow-up, multiple candidate therapies and their future risks are predicted.**
- **The lowest-risk branch is selected, and planning repeats with the next observation.**

Synthesized MRI



Decoded from predicted latent
PSNR 27.91 dB, SSIM 0.890

Takeaway

Disease Modeling
Treatment-conditioned latent trajectories

Clinical Context
Temporal + multimodal conditioning

Decision Making
Adaptive treatment recommendations

Results

Method	MU-Glioma-Post (%)				UCSF-ALPTDG (%)				ISPY-2 (%)			
	Prec.	Rec.	F1	Jac.	Prec.	Rec.	F1	Jac.	Prec.	Rec.	F1	Jac.
General LLMs												
GPT-5 [22]	49.3	50.7	50.0	33.3	38.1	47.5	42.3	26.8	37.7	46.2	41.5	26.2
Claude-4.5-Sonnet [2]	48.6	38.0	41.6	26.3	45.3	38.6	41.7	26.3	42.8	36.5	39.4	24.5
Qwen3-VL [5, 43]	36.7	39.4	38.0	23.5	33.7	42.9	35.8	21.8	32.5	40.1	35.9	21.9
Medical models												
MedGPT [47]	41.6	42.1	41.9	26.5	36.7	46.3	40.9	25.7	38.2	43.5	40.7	25.6
Huatuo-Vision [9]	52.3	46.8	46.4	30.2	42.1	**51.5**	44.1	28.3	48.4	49.1	48.7	32.1
MeWM* [48]	45.2	42.1	43.6	27.9	39.3	48.2	43.3	27.6	41.5	45.3	43.3	27.7
Our Approach	**61.3**	**53.5**	**57.1**	**39.9**	**52.6**	48.0	**50.2**	**33.5**	**56.2**	**51.8**	**53.9**	**36.8**