CLINADAPT: SELF-ADAPTIVE MULTI-AGENT LEARNING FOR LONGITUDINAL CLINICAL EVENT MODELING IN LUNG CANCER EARLY DETECTION

Anonymous authors
Paper under double-blind review

ABSTRACT

Longitudinal clinical event modeling from electronic health records (EHRs) requires reasoning over sparse, noisy, and long-context multimodal sequences. Existing LLM-based multi-agent systems alleviate context-length constraints but process each patient in isolation, failing to emulate how clinicians exploit accumulated experience from similar prior cases. We introduce ClinAdapt, a self-adaptive multi-agent system with two complementary adaptation mechanisms. First, an Adaptive Experience Memory (AEM) serves as non-parametric memory, indexing rejection-sampled reasoning traces to retrieve similar patients as few-shot contexts. Second, Multi-Agent Policy Adaptation (MAPA) with reward-ranked fine-tuning parametrically optimizes inter-agent and agent-memory collaboration. A leave-one-out cross-retrieval strategy unifies these mechanisms, aligning training- and inference-time behavior under retrieval augmentation. On lung cancer prediction using up to five years of multimodal EHRs, ClinAdapt outperforms nine strong baselines on both the overall population and a challenging never-smoker population. Analyses of the adaptation dynamics reveal three key findings. First, as the AEM expands, the optimal retrieval strategy shifts from diverse to specific samples. Second, under MAPA, the manager agent's prediction loss converges rapidly, while worker agents' temporal reasoning continues to benefit from more verified patients. Third, the two mechanisms are complementary in predicted risk, with AEM improving specificity and MAPA improving sensitivity.

1 INTRODUCTION

Lung cancer is the leading cause of cancer-related mortality worldwide Sung et al. (2021); Lancaster et al. (2022), and early detection substantially improves patient outcomes Lancaster et al. (2022). Longitudinal electronic health records (EHRs) offer a uniquely powerful opportunity for early detection, as they accumulate a rich, multimodal clinical history including diagnoses, procedures, laboratory values, vital signs, medications, and unstructured clinical notes, which collectively encode subtle disease trajectories preceding a cancer diagnosis Jensen et al. (2012); Kim et al. (2019). Within these trajectories lie early signals of risk and their trends, such as recurrent respiratory symptoms, chronic pulmonary conditions, or incidental radiographic findings documented years before diagnosis D'Arcy et al. (2025); Ganti et al. (2021).

Extracting and temporally reasoning over these signals from long and noisy patient trajectories, however, is challenging. Recent studies evaluated LLM-based approaches for generalizable modeling from heterogeneous EHR data Cui et al. (2025); Kruse et al. (2025); Zeng et al. (2025; 2026). Among these, Traj-CoA is a multi-agent framework that leverages chain-of-agents and a long-term memory to facilitate temporal reasoning over patient trajectories for cancer early detection, eliminating complex feature engineering while achieving zero-shot performance comparable to supervised machine learning and deep learning models Zeng et al. (2025; 2026).

Despite these advances, existing LLM-based longitudinal EHR modeling systems share a fundamental limitation: they are static. Every patient is processed in isolation, relying solely on the LLM's frozen parametric knowledge and a fixed prompt. This stands in sharp contrast to expert clinical practice, where diagnostic judgement is continually refined by accumulated experience with

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

similar patients. This process is central to how clinicians recognise atypical presentations, such as early lung cancer in a never-smoker with an otherwise unremarkable history Eva (2005); Patel et al. (2005). For lung cancer early detection, where cases are clinically heterogeneous and often subtly distinguished from controls by patterns distributed across years of records, the inability of a system to learn from past verified cases can limit both performance and robustness, particularly in minority subgroups such as never-smokers.

Emerging research on Self-Adaptive LLM agents has the potential to address this gap. Rather than treating the model as immutable, Self-Adaptive agents continually update their behavior through interaction and feedback, evolving their memory, prompts, tools, or parameters as new experience accumulates Gao et al. (2025a); Zhang et al. (2025c). For example, memory-based approaches save the problem-solving trajectories as experience into an external database, which could guide future decisions through retrieval-augmented generation (RAG) Shinn et al. (2023); Zhao et al. (2024); Wu et al. (2025); Zhou et al. (2025a); Tang et al. (2025). In parallel, reinforcement-learning (RL) based approaches such as reward-ranked fine-tuning (RAFT) Dong et al. (2023); Xiong et al. (2025); Zhang et al. (2025b) and multi-agent RL variants Ma et al. (2024); Liao et al. (2025); Zhang et al. (2025a) enable collaborative agent systems to internalise successful reasoning patterns directly into their parameters.

In healthcare, Self-Adaptive agents have been explored in synthetic or simulated patient interactions Li et al. (2024); Almansoori et al. (2025) and medical question answering Chen et al. (2025). Designing Self-Adaptive systems for patient trajectory modeling in early cancer detection poses a distinct challenge: it requires complex temporal reasoning over years of noisy, multimodal data and the ability to draw reusable insights from heterogeneous clinical cases. Existing techniques may not be readily applicable to this scenario, and performance remains unclear. To our knowledge, no prior work has designed Self-Adaptive agents to enhance longitudinal EHR modeling for real-world early cancer detection.

To bridge this gap, we present ClinAdapt, a Self-Adaptive multi-agent framework for patient trajectory modelling that extends the Traj-CoA architecture Zeng et al. (2025) with two complementary evolutionary mechanisms, an evolving Adaptive Experience Memory (AEM) and Multi-Agent Policy Adaptation (MAPA). Collectively, these mechanisms enable ClinAdapt to learn from its own experience as it processes more patients, continually refining its temporal reasoning and learning from "patients-like-me", eventually improving performance over time. These two mechanisms transform patient trajectory modeling from a static, isolated prediction task into a continuously improving clinical learning system.

We evaluated ClinAdapt on a large longitudinal cohort from a medical center, using five years of multimodal EHR history to predict incident lung cancer within the subsequent year among the overall population and the particularly challenging subgroup of never-smokers. We benchmarked against a comprehensive suite of baselines spanning clinical risk models, supervised machine learning, sequential deep learning, clinical BERT-based models, and LLM-based systems.

The main contributions of this work are:

- We introduce ClinAdapt, to our knowledge, the first Self-Adaptive multi-agent framework for longitudinal EHR modelling applied to a real-world clinical prediction task.
- We design two complementary evolutionary mechanisms: an evolving Adaptive Experience Memory (AEM) that provides non-parametric, few-shot "patients-like-me" retrieval, and a MAPA procedure that parametrically optimizes inter-agent and agent-memory collaboration using rejection-sampled high-reward trajectories.
- We demonstrate that ClinAdapt achieves state-of-the-art discrimination for one-year lung cancer prediction in the overall population and the challenging never-smoker population.
- We provide detailed analyses of the Self-Adaptive dynamics, supporting the vision of a continuously improving clinical decision-support system.

2 Related Works

LLM-based Patient Trajectory Modeling Recent work increasingly leverages strong LLMs for zero- or few-shot reasoning over heterogeneous clinical histories, including DT-GPT for clinical variable forecasting Makarov et al. (2025), EHR2Path for scalable patient pathway prediction Pelle-

grini et al. (2025), TIMER for temporal instruction tuning Cui et al. (2025), and Kruse et al. (2025) for long-context summarization. Yet single-LLM pipelines remain limited by the lost-in-the-middle phenomenon Liu et al. (2024) on very long EHRs and by the complexity in specific clinical prediction tasks, motivating multi-agent designs to decompose longitudinal EHR modeling into simpler subtasks. MoMA Gao et al. (2025b) coordinates modality-specialized agents for clinical prediction, CARE-AD Li et al. (2025b) and ClinNoteAgents Zhou et al. (2025b) decompose reasoning across specialist agents, and Traj-CoA Zeng et al. (2025; 2026) extend chain-of-agents Zhang et al. (2024) with long-term memory for cancer early detection. Complementary efforts such as CliCARE Li et al. (2025a) and TRACE Qu & Färber (2026) further explore temporal knowledge graph and dual-memory approaches. However, these systems are static: each patient is reasoned about in isolation, with no mechanism or evaluation for accumulating verified clinical experience over time.

Self-Adaptive Agents Gao et al. (2025a) organize Self-Adaptive agents along three axes: what, when, and how to evolve. Along what to evolve, prior work targets memory, prompts, tools, or model parameters; along when, adaptation can be intra- or inter-test-time; along how, it is driven by textual feedback or scalar rewards in single- or multi-agent settings. For example, memory-evolving methods such as Reflexion Shinn et al. (2023), ExpeL Zhao et al. (2024), Memento Zhou et al. (2025a), and Agent KB Tang et al. (2025) store and retrieve past trajectories as non-parametric experience. Parameter-evolving methods internalize successful experience via model training, including supervised fine-tuning and reinforcement learning Zelikman et al. (2022); Zuo et al. (2025); Dong et al. (2023); Wang et al. (2025) and multi-agent extensions Zhang et al. (2025a); Ma et al. (2024); Liao et al. (2025). These two families are typically pursued in isolation. In healthcare, self-evolution has so far been confined to simulated or interactive settings, including Agent Hospital Li et al. (2024), MedAgentSim Almansoori et al. (2025), MDTeamGPT Chen et al. (2025), and EvoClinician He et al. (2026).

3 METHODS

3.1 PROBLEM FORMULATION

Lung Cancer Early Detection Let $\mathcal{P} = \{p_i\}_{i=1}^{N}$ denote a cohort of patients. For each patient p_i, we observe a longitudinal multimodal EHR sequence

$$\mathcal{X}_i = \left\{(t_{i,j}, e_{i,j})\right\}_{j=1}^{T_i}, \quad t_{i,j} \le t_i^\star, \tag{1}$$

where $t_i^\star$ is the patient-specific index date (time of prediction), T_i is the number of dated entries within the available EHR, and each event $e_{i,j}$ at time $t_{i,j}$ consists of either a structured record (diagnosis, medication, lab, vital, or procedure code) or unstructured clinical text (notes and radiology reports). The binary target $y_i \in \{0, 1\}$ indicates whether p_i receives a first primary lung-cancer diagnosis within one year after $t_i^\star$.

The task of lung cancer early detection is to learn a function $f_\theta : \mathcal{X}_i \mapsto (s_i, r_i)$ that maps the longitudinal record to an integer risk score $s_i \in \{1, \ldots, 10\}$ and a natural-language rationale r_i.

Self-Adaptive System Beyond standard generalization, we additionally require f_θ to improve as more patients are seen and verified, yielding a Self-Adaptive system. We formalize this as follows. Patients arrive sequentially as a stream $p_1, p_2, \ldots$. At each step t, the system maintains an experience set $\mathcal{E}_t$ that summarises all patients that have been processed and verified by a known diagnostic status so far. The primary objective of a Self-Adaptive system is to continuously improve the performance of $f_{\theta(\mathcal{E}_t)}$ as the experience set $\mathcal{E}_t$ grows. This captures two coupled challenges: $f_{\theta(\mathcal{E}_t)}$ must perform complex temporal reasoning over long and noisy EHRs to produce (s_t, r_t), and it additionally needs to construct and leverage $\mathcal{E}_t$ so that performance improves with $|\mathcal{E}_t|$.

We randomly split the dataset into training $\mathcal{D}_{tr}$, validation $\mathcal{D}_{val}$, and test set $\mathcal{D}_{test}$. In practice, we simulated the stream of verified patients by drawing a random sample from the training set that grew incrementally over time.

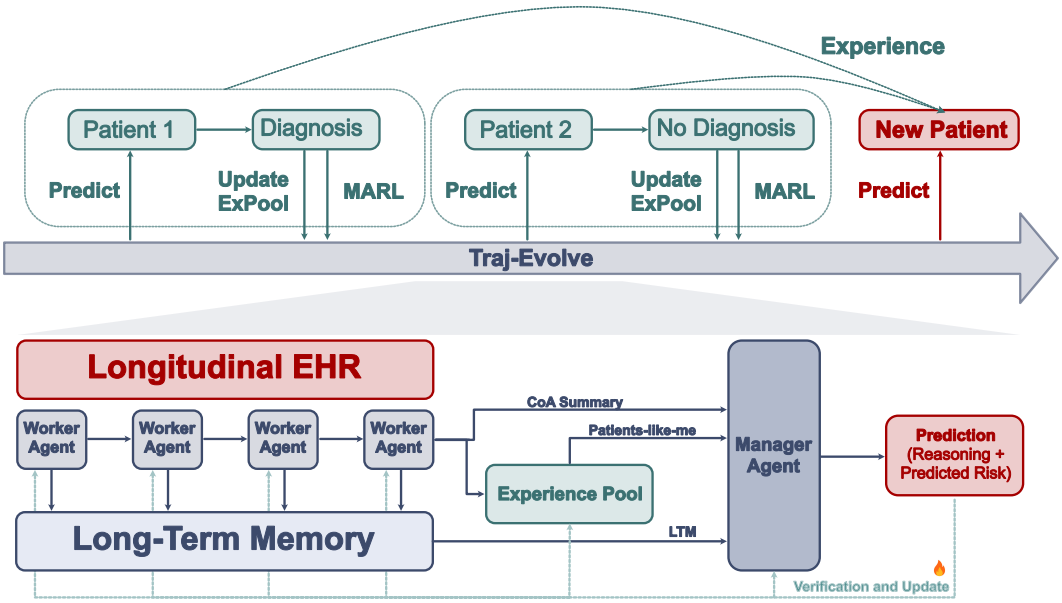

Figure 1: **Overview of the ClinAdapt architecture and Self-Adaptive workflow.** The top panel illustrates the Self-Adaptive process, wherein the system accumulates experience from prior verified patients to iteratively update ClinAdapt and facilitate prediction for a new patient. The bottom panel details the pipeline.

3.2 Background: Traj-CoA Base System

We build on Traj-CoA (Zeng et al., 2025; 2026), a static chain-of-agents (CoA) Zhang et al. (2024) backbone that handles long-context EHRs. Each $\mathcal{X}_i$ is serialized into an LLM-friendly XML representation and segmented into C_i chronologically ordered chunks $\{c_{i,1}, \ldots, c_{i,C_i}\}$ such that each chunk fits within the LLM context limit.

A sequence of C_i worker agents, all parameterized by θ_w, processes the chunks sequentially. Concretely, the ℓ-th worker maintains a running summary $u_{i,\ell}$ and a long-term episodic memory (LTM) $\mathcal{M}_{i,\ell}$. The final worker summary $u_i = u_{i,C_i}$ encapsulates the full EHR trajectory, while $\mathcal{M}_i = \mathcal{M}_{i,C_i}$ stores a condensed timeline of lung-cancer-related events extracted across all chunks. A manager agent π_{θ_m} then synthesizes both signals to produce the final output:

$$(s_i, r_i) = \pi_{\theta_m}(u_i, \mathcal{M}_i). \tag{2}$$

This zero-shot pipeline decomposes complex temporal reasoning into inter-agent and agent-memory collaboration tasks, improving patient trajectory modeling performance. However, it treats each patient in isolation, conditioning only on the LLM's parametric knowledge. This creates a gap with expert clinical practice, which relies on accumulated case experience. To bridge this gap, we design two complementary Self-Adaptive mechanisms atop Traj-CoA for our Self-Adaptive ClinAdapt system (Figure 1).

3.3 Evolving Adaptive Experience Memory (AEM)

Inspired by the procedural memory approach that saves successful reasoning traces for future problem solving of similar tasks Tang et al. (2025), we design an evolving experience pool (AEM, Figure 2A). AEM equips ClinAdapt with a non-parametric procedural memory that saves certain reasoning traces of verified patients. As ClinAdapt generates predictions that can be subsequently verified against ground-truth diagnostic status (e.g., confirmed cancer diagnosis or benign status), these verified reasoning traces can serve as experience for future cases. This design is analogous to expert clinical reasoning, in which a patient's presentation is rarely adjudicated in isolation but rather by analogy to remembered cases that share a similar longitudinal clinical pattern. Correct reasoning

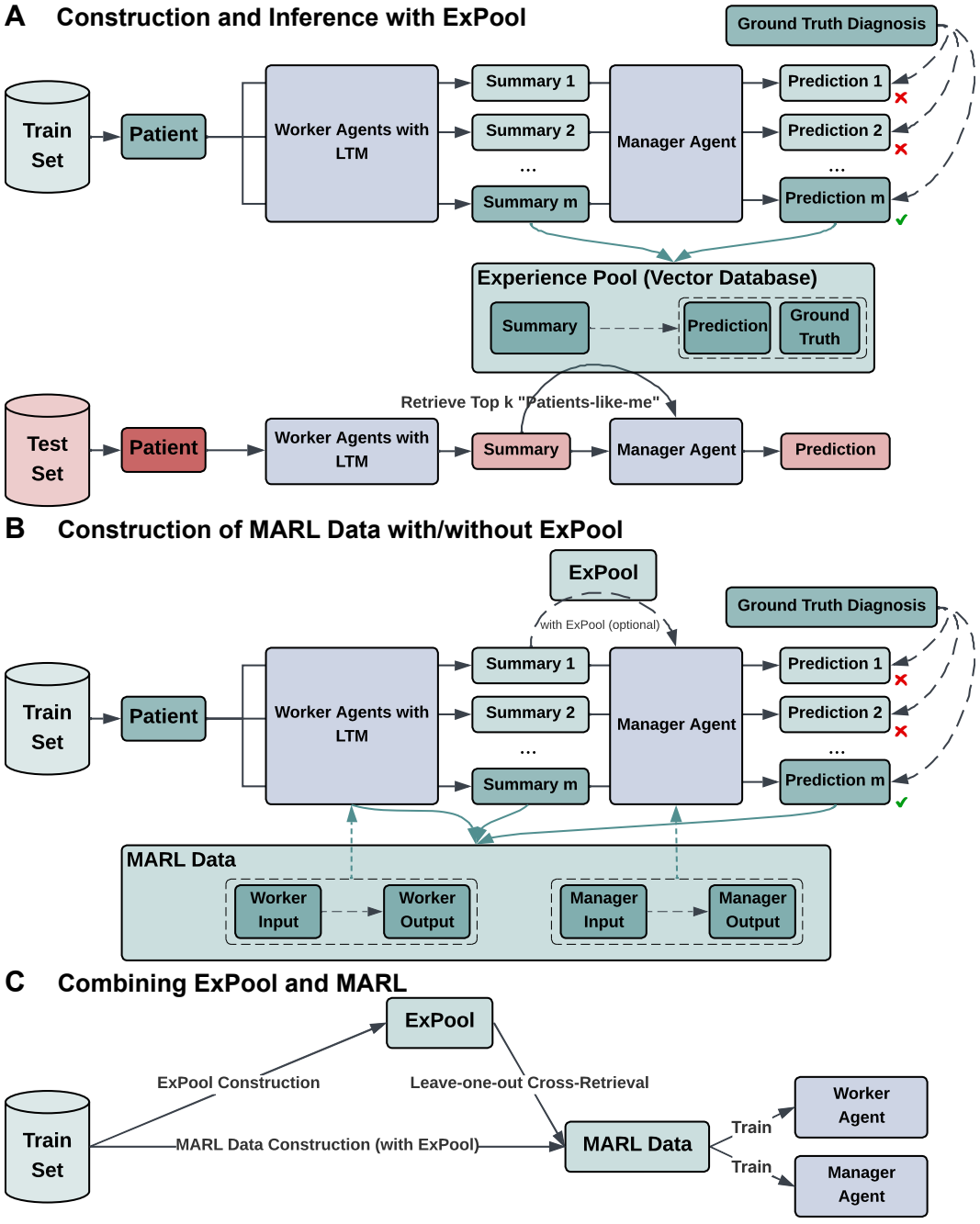

Figure 2: **Methodological design of the Self-Adaptive mechanisms.** (A) Construction and inference pipeline for the AEM. (B) Construction of MAPA training data. (C) Integration of AEM and MAPA.

traces may offer reusable diagnostic patterns, while incorrect predictions may expose model vulnerabilities and trigger self-reflection. These data-driven patterns evolve the system beyond its current boundary, forming the Self-Adaptive capability.

AEM Construction To capture the current system's boundary, i.e., the optimal diagnostic ability, we constructed AEM using a rejection sampling approach that selected best-of-N from a set of roll-out reasoning traces. Formally, for each patient $p_i \in \mathcal{D}_{tr}$, we draw m independent roll-outs from

the base system under an elevated sampling temperature $\tau > 1$ to diversify reasoning traces,

$$\left\{(u_i^{(j)}, \mathcal{M}_i^{(j)}, s_i^{(j)}, r_i^{(j)})\right\}_{j=1}^{m} \sim f_\theta(\mathcal{X}_i; \tau), \tag{3}$$

and retain a single optimal trace via label-conditioned rejection sampling

$$j_i^\star = \begin{cases} \arg\max_j s_i^{(j)}, & y_i = 1, \\ \arg\min_j s_i^{(j)}, & y_i = 0. \end{cases} \tag{4}$$

Notably, these optimal reasoning traces may mix correct and incorrect predictions, which provides different signals for AEM.

These selected traces then populate AEM, a vector database where the retrieval keys are the embeddings of the final worker agent summaries:

$$\mathbf{v}_i = \phi(u_i^\star), \tag{5}$$

in which $\phi(\cdot)$ denotes an embedding model and index each experience as a key-value pair

$$\mathcal{E} = \left\{\left(\mathbf{v}_i,\ (r_i^\star, s_i^\star, y_i)\right)\right\}_{i \in \mathcal{D}_{\text{tr}}}, \tag{6}$$

where the value stores the manager's rationale, predicted risk, and ground-truth label. This design makes it feasible to embed and index long patient trajectories in the latent space, providing an efficient approach for experience retrieval.

AEM Inference During inference, AEM functions using a retrieval-augmented generation (RAG) Xu et al. (2024) mechanism. For each new patient p_q, we adopt semantic retrieval by querying $\mathcal{E}$ with $\mathbf{v}_q = \phi(u_q)$ and returning the top-k nearest neighbors as "patients-like-me" using cosine similarity:

$$\mathcal{N}_k(q) = [\text{Top-}k]_{i \in \mathcal{E}} \ \frac{\mathbf{v}_q^\top \mathbf{v}_i}{\|\mathbf{v}_q\| \, \|\mathbf{v}_i\|}. \tag{7}$$

The semantic matching is chosen over exact matching because it yields a soft neighborhood that balances diversity and specificity. Retrieved patients may exhibit different clinical profiles, matching the patient p_q in diverse ways. Rather than collapsing this heterogeneity into a hard label via majority voting, we delegate the comparative reasoning to the manager agent, conditioning it on the full retrieved set:

$$(s_q, r_q) = \pi_{\theta_m}\big(u_q,\ \mathcal{M}_q,\ \{(r_i^\star, s_i^\star, y_i)\}_{i \in \mathcal{N}_k(q)}\big). \tag{8}$$

AEM therefore provides the manager agent with diverse but clinically relevant "patients-like-me" as in-context examples, and the manager agent comprehensively reasons over all information for prediction. As AEM continuously scales with newly verified patients, the retrieval of highly specific neighbors becomes increasingly precise, shifting the system from isolated predictions to a progressively Self-Adaptive framework.

3.4 Multi-Agent Policy Adaptation (MAPA)

AEM adapts the system at inference time, but it leaves the underlying agent parameters (θ_w, θ_m) unchanged and acts only on the final manager step. The intermediate worker reasoning, which decides what gets distilled into $\mathcal{M}_i$ (agent-memory communication) and how the chunk-by-chunk summary is built up (inter-agent communication), does not receive a learning signal from the verified outcomes. To close this gap, we introduce a second Self-Adaptive mechanism that internalizes successful reasoning traces directly into the model parameters, adapting reward-ranked fine-tuning (RAFT) (Dong et al., 2023), a reinforcement learning strategy that optimizes models using the reasoning trace with the highest reward among all self-generated roll-outs, to our multi-agent setting. (Figure 2B)

Reward and Accepted Set Parallel to the construction of AEM, we generated m roll-outs with an elevated temperature. We use the ground-truth label y_i as a binary reward signal. Unlike AEM,

which retains the optimal trace per patient regardless of its correctness, MAPA retains only clinically consistent traces in its accepted set $\mathcal{A}$:

$$\mathcal{A} = \big\{(i,j) : R(s_i^{(j)}, y_i) = 1\big\}, \tag{9}$$

where the reward function is

$$R(s,y) = \begin{cases} \mathbb{1}[s \geq 6], & y = 1, \\ \mathbb{1}[s \leq 4], & y = 0. \end{cases} \tag{10}$$

Eq. equation 10 acts as a hard rejection filter, ensuring the resulting fine-tuning data consists exclusively of logically sound and clinically accurate reasoning traces.

Decoupled Optimization For each accepted trace $(i,j) \in \mathcal{A}$, we decompose the roll-out into per-agent input-output pairs. Let $\mathbf{x}^w_{i,j,c}$ and $\mathbf{y}^w_{i,j,c}$ denote the input and output of the c-th worker, and let $\mathbf{x}^m_{i,j}, \mathbf{y}^m_{i,j}$ denote the manager pair. To preserve temporal balance across the chain, we subsample worker positions to retain the first, the last, and two randomly selected intermediate worker agents per reasoning trace.

$$\mathcal{C}_{i,j} = \{1,\, C_i\} \cup \mathrm{Sample}_2\big(\{2, \ldots, C_i - 1\}\big), \tag{11}$$

yielding the worker and manager training sets

$$\mathcal{D}_w = \big\{(\mathbf{x}^w_{i,j,c}, \mathbf{y}^w_{i,j,c}) : (i,j) \in \mathcal{A},\, c \in \mathcal{C}_{i,j}\big\}, \tag{12}$$

$$\mathcal{D}_m = \big\{(\mathbf{x}^m_{i,j}, \mathbf{y}^m_{i,j}) : (i,j) \in \mathcal{A}\big\}. \tag{13}$$

The worker and manager parameters are updated independently to preserve their specialized roles:

$$\theta^\star_w = \arg\min_{\theta_w} \sum_{(\mathbf{x},\mathbf{y}) \in \mathcal{D}_w} -\log \pi_{\theta_w}(\mathbf{y} \mid \mathbf{x}), \tag{14}$$

$$\theta^\star_m = \arg\min_{\theta_m} \sum_{(\mathbf{x},\mathbf{y}) \in \mathcal{D}_m} -\log \pi_{\theta_m}(\mathbf{y} \mid \mathbf{x}). \tag{15}$$

Optimizing θ_w refines sequential inter-agent reasoning and worker-memory collaboration, while optimizing θ_m refines how the final summary and the LTM are aggregated into a risk estimation. As verified patient cohort expands, this MAPA approach benefits from larger training data and drives continuous self-evolution through reinforcing successful temporal reasoning and collaborative memory distillation to progressively enhance ClinAdapt's intrinsic diagnostic capabilities.

3.5 Combining AEM and MAPA

AEM provides retrieved "patients-like-me" but does not update parameters; Vanilla MAPA updates parameters but lacks an explicit mechanism to incorporate similar patients during final risk estimation. We unify the two through a leave-one-out cross-retrieval procedure that prevents data leakage and matches training and inference input formats, while injecting retrieval signals into the optimization. (Figure 2C)

For each $p_i \in \mathcal{D}_{\text{tr}}$, we construct a patient-specific dynamic pool $\mathcal{E}_{-i} = \mathcal{E} \backslash \{i\}$ from which we retrieve $\mathcal{N}_k(i)$ for the manager agent during the MAPA roll-out phase. The manager input is augmented with the retrieved patients for prediction, after which the rejection filter in Eq. equation 10 is applied to obtain the accepted set $\mathcal{A}^{\text{loo}}$. Formally, the accepted set is defined as:

$$\mathcal{A}^{\text{loo}} = \Big\{(i,j) : R\big(s_i^{(j)}(\mathcal{E}_{-i}),\, y_i\big) = 1\Big\}, \tag{16}$$

where $s_i^{(j)}(\mathcal{E}_{-i})$ denotes the j-th roll-out conditioned on neighbours retrieved from $\mathcal{E}_{-i}$. This methodological synthesis yields optimal traces that benefit from the augmented context provided by "patients-like-me" from the dynamic AEM, thereby generating potentially better training data compared to MAPA in isolation. The worker and manager agents were subsequently trained using the standard MAPA protocol, while inference was conducted utilizing the standard AEM RAG methodology on the full AEM from all training patients.

Table 1: Baseline characteristics of cases and controls in the overall and never-smoker cohorts.

	Overall sample		Never-smoker sample	
Characteristic	Cases (n=90)	Controls (n=910)	Cases (n=27)	Controls (n=808)
Demographics				
Age at index date, years, median (IQR)	68 (60–76)	68 (59–78)	74 (64–78)	68 (58–80)
Sex, n (%)				
Female	49 (54.4)	423 (46.5)	24 (88.9)	463 (57.3)
Male	41 (45.6)	487 (53.5)	3 (11.1)	345 (42.7)
Race, n (%)				
White	74 (82.2)	657 (72.2)	21 (77.8)	572 (70.8)
Black or African American	8 (8.9)	102 (11.2)	0 (0.0)	80 (9.9)
Asian	5 (5.6)	77 (8.5)	6 (22.2)	99 (12.3)
Other	1 (1.1)	28 (3.1)	0 (0.0)	22 (2.7)
Unknown	2 (2.2)	46 (5.1)	0 (0.0)	35 (4.3)
Ethnicity, n (%)				
Hispanic or Latino	4 (4.4)	48 (5.3)	4 (14.8)	47 (5.8)
Not Hispanic or Latino	80 (88.9)	758 (83.3)	23 (85.2)	722 (89.4)
Unknown	6 (6.7)	104 (11.4)	0 (0.0)	39 (4.8)
Clinical index and EHR history				
Index year, median (IQR)	2016 (2014–2018)	2016 (2014–2018)	2018 (2016–2018)	2017 (2015–2018)
EHR look-back duration, years, median (IQR)	4.5 (2.0–4.9)	3.7 (1.1–4.9)	4.2 (2.4–4.8)	4.5 (2.8–4.9)
Time to event, years, median (IQR)	0.5 (0.4–0.7)	—	0.5 (0.3–0.8)	—
XML token count, median (IQR)	61 678 (26 535–151 282)	51 904 (18 429–125 871)	84 406 (46 292–131 768)	75 070 (35 060–153 634)
Distinct dated EHR entries, median (IQR)	44 (19–113)	39 (12–95)	74 (30–102)	58 (28–114)
Index exam modality, n (%)				
Computed radiography (CR)	45 (50.0)	516 (56.7)	15 (55.6)	493 (61.0)
Computed tomography (CT)	45 (50.0)	394 (43.3)	12 (44.4)	315 (39.0)
Index exam body site, n (%)				
Chest	59 (65.6)	548 (60.2)	18 (66.7)	550 (68.1)
Abdomen/pelvis	8 (8.9)	135 (14.8)	5 (18.5)	140 (17.3)
Extremity	5 (5.6)	95 (10.4)	2 (7.4)	49 (6.1)
Other	0 (0.0)	4 (0.4)	0 (0.0)	4 (0.5)
Missing	18 (20.0)	128 (14.1)	2 (7.4)	65 (8.0)
Low-dose CT lung cancer screening, n (%)	1 (1.1)	18 (2.0)	0 (0.0)	0 (0.0)
Clinical text and structured data volume, median (IQR)				
Epic notes	10 (0–50)	6 (0–28)	32 (7–58)	20 (8–45)
ORCA notes	2 (0–12)	4 (0–13)	0 (0–10)	2 (0–12)
Radiology reports	10 (5–19)	8 (3–14)	9 (6–18)	8 (3–15)
Diagnosis entries	144 (53–340)	118 (34–346)	283 (115–377)	202 (86–451)
Medication entries	52 (9–200)	59 (10–164)	77 (37–173)	72 (23–184)
Procedure entries	118 (37–220)	98 (34–221)	100 (72–170)	124 (52–272)
Laboratory entries	167 (54–538)	177 (49–484)	293 (160–627)	230 (93–577)
Vital-sign entries	28 (0–118)	18 (0–83)	81 (22–172)	61 (24–135)
Lifestyle factors				
Smoking status, n (%)				
Current	16 (17.8)	82 (9.0)	0 (0.0)	0 (0.0)
Former	29 (32.2)	180 (19.8)	0 (0.0)	0 (0.0)
Never	11 (12.2)	289 (31.8)	27 (100.0)	808 (100.0)
Passive exposure only	0 (0.0)	2 (0.2)	0 (0.0)	0 (0.0)
Unknown	34 (37.8)	357 (39.2)	0 (0.0)	0 (0.0)
Alcohol use, n (%)				
Current use	16 (17.8)	219 (24.1)	10 (37.0)	285 (35.3)
No current use	30 (33.3)	260 (28.6)	11 (40.7)	426 (52.7)
Unknown	44 (48.9)	431 (47.4)	6 (22.2)	97 (12.0)

4 EXPERIMENTS

Dataset We predict first primary lung cancer diagnoses within a one-year window using an in-house longitudinal EHR dataset. The input $\mathcal{X}_i$ comprises up to five years of EHR history before the index date $t_i^\star$ (the completion of a chest radiology exam). Cases and controls are 1:10 matched on index exam type and date. The training set $\mathcal{D}_{tr}$ contains $13,629$ patients. We evaluate on two disjoint, held-out test sets: an overall cohort ($n = 1,000$) and a clinically more challenging never-smoker cohort ($n = 835$). Notably, patients inputs are exceptionally long, with median XML token counts exceeding 60k overall and 80k for never-smokers (Table 1).

Implementation Details ClinAdapt uses GPT-OSS-20B OpenAI et al. (2025) as the base LLM and nomic-embed-text-v1.5 Nussbaum et al. (2024) for AEM embeddings. During roll-out, we sample $m = 4$ traces at temperature $\tau = 1.5$. Both MAPA agents are trained via QLoRA Dettmers et al. (2023) for one epoch.

Baselines We compare against five baseline categories: (1) clinical risk models (LCRAT (Katki et al., 2016)); (2) supervised ML (Logistic Regression, XGBoost (Chen & Guestrin, 2016)); (3)

sequential deep learning (RETAIN (Choi et al., 2016), PatientTM (Silva & Matos, 2022)); (4) clinical BERT (Clinical ModernBERT (Lee et al., 2025)); and (5) GPT-OSS-20B-based LLM pipelines (vanilla LLM, RAG, and Traj-CoA (Zeng et al., 2025)).

Evaluation For binary risk classification, we report AUROC, AUPRC, and F1 as primary metrics, alongside sensitivity, specificity, PPV, and NPV from the predicted risk s_i and binary target y_i. All results are reported as bootstrap means and standard errors over $1,000$ resamples of each test set.

5 Conclusion

We present ClinAdapt, a Self-Adaptive multi-agent framework for lung cancer early detection from longitudinal EHRs. By combining non-parametric few-shot retrieval (AEM) with parametric reasoning internalization (MAPA), ClinAdapt emulates the continuous learning of experienced clinicians. ClinAdapt establishes a new state-of-the-art for one-year lung cancer prediction and shows particular robustness in challenging subgroups like never-smokers. Furthermore, we demonstrate that AEM and MAPA are highly complementary, improving specificity and sensitivity, respectively. Future work will extend this framework to real-world streaming settings with richer reward signals and additional clinical tasks.

References

Mohammad Almansoori, Komal Kumar, and Hisham Cholakkal. Medagentsim: Self-evolving multi-agent simulations for realistic clinical interactions. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 362–372. Springer, 2025.

Kai Chen, Xinfeng Li, Tianpei Yang, Hewei Wang, Wei Dong, and Yang Gao. Mdteamgpt: A self-evolving llm-based multi-agent framework for multi-disciplinary team medical consultation. *arXiv preprint arXiv:2503.13856*, 2025.

Tianqi Chen and Carlos Guestrin. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, August 2016. doi: 10.1145/2939672.2939785. URL http://arxiv.org/abs/1603.02754. arXiv:1603.02754 [cs].

Edward Choi, Mohammad Taha Bahadori, Jimeng Sun, Joshua Kulas, Andy Schuetz, and Walter Stewart. Retain: An interpretable predictive model for healthcare using reverse time attention mechanism. *Advances in neural information processing systems*, 29, 2016.

Hejie Cui, Alyssa Unell, Bowen Chen, Jason Alan Fries, Emily Alsentzer, Sanmi Koyejo, and Nigam H Shah. Timer: Temporal instruction modeling and evaluation for longitudinal clinical records. *npj Digital Medicine*, 8(1):577, 2025.

Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, and Luke Zettlemoyer. QLoRA: Efficient Finetuning of Quantized LLMs, May 2023. URL http://arxiv.org/abs/2305.14314. arXiv:2305.14314 [cs].

Hanze Dong, Wei Xiong, Deepanshu Goyal, Yihan Zhang, Winnie Chow, Rui Pan, Shizhe Diao, Jipeng Zhang, Kashun Shum, and Tong Zhang. Raft: Reward ranked finetuning for generative foundation model alignment. *arXiv preprint arXiv:2304.06767*, 2023.

Monica E D'Arcy, Ruth M Pfeiffer, Marie C Bradley, Phuc H Hoang, Thi-Van-Trinh Tran, John P McElderry, Mengying Li, Michael Kebede, Curt T DellaValle, Sara Rivas, et al. Inflammatory diseases and risk of lung cancer among individuals who have never smoked. *Nature communications*, 16(1):5095, 2025.

Kevin W Eva. What every teacher needs to know about clinical reasoning. *Medical education*, 39 (1):98–106, 2005.

Apar Kishor Ganti, Alyssa B Klein, Ion Cotarla, Brian Seal, and Engels Chou. Update of incidence, prevalence, survival, and initial treatment in patients with non–small cell lung cancer in the us. *JAMA oncology*, 7(12):1824–1832, 2021.

Huan-ang Gao, Jiayi Geng, Wenyue Hua, Mengkang Hu, Xinzhe Juan, Hongzhang Liu, Shilong Liu, Jiahao Qiu, Xuan Qi, Yiran Wu, et al. A survey of self-evolving agents: What, when, how, and where to evolve on the path to artificial super intelligence. *arXiv preprint arXiv:2507.21046*, 2025a.

Jifan Gao, Mahmudur Rahman, John Caskey, Madeline Oguss, Ann O'Rourke, Randall Brown, Anne Stey, Anoop Mayampurath, Matthew M. Churpek, Guanhua Chen, and Majid Afshar. MoMA: a mixture-of-multimodal-agents architecture for enhancing clinical prediction modelling. *npj Digital Medicine*, 9(1):46, December 2025b. ISSN 2398-6352. doi: 10.1038/s41746-025-02219-4. URL https://www.nature.com/articles/s41746-025-02219-4.

Yufei He, Juncheng Liu, Zhiyuan Hu, Yulin Chen, Yue Liu, Yuan Sui, Yibo Li, Nuo Chen, Jun Hu, Bryan Hooi, Xinxing Xu, and Jiang Bian. EvoClinician: A Self-Evolving Agent for Multi-Turn Medical Diagnosis via Test-Time Evolutionary Learning, January 2026. URL http://arxiv.org/abs/2601.22964. arXiv:2601.22964 [cs.AI].

Peter B Jensen, Lars J Jensen, and Søren Brunak. Mining electronic health records: towards better research applications and clinical care. *Nature Reviews Genetics*, 13(6):395–405, 2012.

Hormuzd A Katki, Stephanie A Kovalchik, Christine D Berg, Li C Cheung, and Anil K Chaturvedi. Development and validation of risk models to select ever-smokers for ct lung cancer screening. *Jama*, 315(21):2300–2311, 2016.

Ellen Kim, Samuel M Rubinstein, Kevin T Nead, Andrzej P Wojcieszynski, Peter E Gabriel, and Jeremy L Warner. The evolving use of electronic health records (ehr) for research. In *Seminars in radiation oncology*, volume 29, pp. 354–361. Elsevier, 2019.

Maya Kruse, Shiyue Hu, Nicholas Derby, Yifu Wu, Samantha Stonbraker, Bingsheng Yao, Dakuo Wang, Elizabeth Goldberg, and Yanjun Gao. Large language models with temporal reasoning for longitudinal clinical summarization and prediction. In *Findings of ACL. EMNLP. Conference on Empirical Methods in Natural Language Processing*, volume 2025, pp. 20715–20735, 2025.

Harriet L Lancaster, Marjolein A Heuvelmans, and Matthijs Oudkerk. Low-dose computed tomography lung cancer screening: Clinical evidence and implementation research. *Journal of internal medicine*, 292(1):68–80, 2022.

Simon A. Lee, Anthony Wu, and Jeffrey N. Chiang. Clinical ModernBERT: An efficient and long context encoder for biomedical text, April 2025. URL http://arxiv.org/abs/2504.03964. arXiv:2504.03964 [cs].

Dongchen Li, Jitao Liang, Wei Li, Xiaoyu Wang, Longbing Cao, and Kun Yu. CliCARE: Grounding Large Language Models in Clinical Guidelines for Decision Support over Longitudinal Cancer Electronic Health Records, July 2025a. URL https://arxiv.org/abs/2507.22533v2.

Junkai Li, Yunghwei Lai, Weitao Li, Jingyi Ren, Meng Zhang, Xinhui Kang, Siyu Wang, Peng Li, Ya-Qin Zhang, Weizhi Ma, et al. Agent hospital: A simulacrum of hospital with evolvable medical agents. *arXiv preprint arXiv:2405.02957*, 2024.

Rumeng Li, Xun Wang, Dan Berlowitz, Jesse Mez, Honghuang Lin, and Hong Yu. CARE-AD: a multi-agent large language model framework for Alzheimer's disease prediction using longitudinal clinical notes. *npj Digital Medicine*, 8(1):541, August 2025b. ISSN 2398-6352. doi: 10.1038/s41746-025-01940-4. URL https://www.nature.com/articles/s41746-025-01940-4.

Junwei Liao, Muning Wen, Jun Wang, and Weinan Zhang. Marft: Multi-agent reinforcement fine-tuning. *arXiv preprint arXiv:2504.16129*, 2025.

Nelson F. Liu, Kevin Lin, John Hewitt, Ashwin Paranjape, Michele Bevilacqua, Fabio Petroni, and Percy Liang. Lost in the Middle: How Language Models Use Long Contexts. *Transactions of the Association for Computational Linguistics*, 12:157–173, 2024. doi: 10.1162/tacl_a_00638. URL https://aclanthology.org/2024.tacl-1.9/. Place: Cambridge, MA.

Hao Ma, Tianyi Hu, Zhiqiang Pu, Boyin Liu, Xiaolin Ai, Yanyan Liang, and Min Chen. Coevolving with the other you: Fine-tuning llm with sequential cooperative multi-agent reinforcement learning. *Advances in Neural Information Processing Systems*, 37:15497–15525, 2024.

Nikita Makarov, Maria Bordukova, Papichaya Quengdaeng, Daniel Garger, Raul Rodriguez-Esteban, Fabian Schmich, and Michael P Menden. Large language models forecast patient health trajectories enabling digital twins. *npj Digital Medicine*, 8(1):588, 2025.

Zach Nussbaum, John X. Morris, Brandon Duderstadt, and Andriy Mulyar. Nomic Embed: Training a Reproducible Long Context Text Embedder, February 2024. URL https://arxiv.org/abs/2402.01613v2.

OpenAI, Sandhini Agarwal, Lama Ahmad, and et al. gpt-oss-120b & gpt-oss-20b Model Card, August 2025. URL http://arxiv.org/abs/2508.10925. arXiv:2508.10925 [cs].

Vimla L Patel, José F Arocha, and Jiajie Zhang. Thinking and reasoning in medicine. *The Cambridge handbook of thinking and reasoning*, 14(727-750):1, 2005.

Chantal Pellegrini, Ege Özsoy, David Bani-Harouni, Matthias Keicher, and Nassir Navab. From ehrs to patient pathways: Scalable modeling of longitudinal health trajectories with llms. *arXiv preprint arXiv:2506.04831*, 2025.

Zhan Qu and Michael Färber. TRACE: Temporal Reasoning via Agentic Context Evolution for Streaming Electronic Health Records (EHRs), February 2026. URL https://arxiv.org/abs/2602.12833v1.

Noah Shinn, Federico Cassano, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. Reflexion: Language agents with verbal reinforcement learning. *Advances in neural information processing systems*, 36:8634–8652, 2023.

João Figueira Silva and Sérgio Matos. Modelling patient trajectories using multimodal information. *Journal of Biomedical Informatics*, 134:104195, October 2022. ISSN 1532-0464. doi: 10.1016/j.jbi.2022.104195. URL https://www.sciencedirect.com/science/article/pii/S1532046422002003.

Hyuna Sung, Jacques Ferlay, Rebecca L Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 71 (3):209–249, 2021.

Xiangru Tang, Tianrui Qin, Tianhao Peng, Ziyang Zhou, Daniel Shao, Tingting Du, Xinming Wei, Peng Xia, Fang Wu, He Zhu, et al. Agent kb: Leveraging cross-domain experience for agentic problem solving. *arXiv preprint arXiv:2507.06229*, 2025.

Zihan Wang, Kangrui Wang, Qineng Wang, Pingyue Zhang, Linjie Li, Zhengyuan Yang, Xing Jin, Kefan Yu, Minh Nhat Nguyen, Licheng Liu, Eli Gottlieb, Yiping Lu, Kyunghyun Cho, Jiajun Wu, Li Fei-Fei, Lijuan Wang, Yejin Choi, and Manling Li. RAGEN: Understanding Self-Evolution in LLM Agents via Multi-Turn Reinforcement Learning, May 2025. URL http://arxiv.org/abs/2504.20073. arXiv:2504.20073 [cs.LG].

Rong Wu, Xiaoman Wang, Jianbiao Mei, Pinlong Cai, Daocheng Fu, Cheng Yang, Licheng Wen, Xuemeng Yang, Yufan Shen, Yuxin Wang, et al. Evolver: Self-evolving llm agents through an experience-driven lifecycle. *arXiv preprint arXiv:2510.16079*, 2025.

Wei Xiong, Jiarui Yao, Yuhui Xu, Bo Pang, Lei Wang, Doyen Sahoo, Junnan Li, Nan Jiang, Tong Zhang, Caiming Xiong, et al. A minimalist approach to llm reasoning: from rejection sampling to reinforce. *arXiv preprint arXiv:2504.11343*, 2025.

Zhentao Xu, Mark Jerome Cruz, Matthew Guevara, Tie Wang, Manasi Deshpande, Xiaofeng Wang, and Zheng Li. Retrieval-Augmented Generation with Knowledge Graphs for Customer Service Question Answering. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 2905–2909, July 2024. doi: 10.1145/3626772.3661370. URL http://arxiv.org/abs/2404.17723. arXiv:2404.17723 [cs].

Eric Zelikman, Yuhuai Wu, Jesse Mu, and Noah D. Goodman. STaR: Bootstrapping Reasoning With Reasoning, May 2022. URL http://arxiv.org/abs/2203.14465. arXiv:2203.14465 [cs.LG].

Sihang Zeng, Yujuan Fu, Sitong Zhou, Zixuan Yu, Lucas Jing Liu, Jun Wen, Matthew Thompson, Ruth Etzioni, and Meliha Yetisgen. Traj-coa: Patient trajectory modeling via chain-of-agents for lung cancer risk prediction. *arXiv preprint arXiv:2510.10454*, 2025.

Sihang Zeng, Young Won Kim, Wilson Lau, Ehsan Alipour, Ruth Etzioni, Meliha Yetisgen, and Anand Oka. Trajonco: a multi-agent framework for temporal reasoning over longitudinal ehr for multi-cancer early detection. *arXiv preprint arXiv:2604.10386*, 2026.

Kaiyan Zhang, Kai Tian, Runze Liu, Sihang Zeng, Xuekai Zhu, Guoli Jia, Yuchen Fan, Xingtai Lv, Yuxin Zuo, Che Jiang, et al. Marti: A framework for multi-agent llm systems reinforced training and inference. In *The Fourteenth International Conference on Learning Representations*, 2025a.

Kaiyan Zhang, Yuxin Zuo, Bingxiang He, Youbang Sun, Runze Liu, Che Jiang, Yuchen Fan, Kai Tian, Guoli Jia, Pengfei Li, Yu Fu, Xingtai Lv, Yuchen Zhang, Sihang Zeng, Shang Qu, Haozhan Li, Shijie Wang, Yuru Wang, Xinwei Long, Fangfu Liu, Xiang Xu, Jiaze Ma, Xuekai Zhu, Ermo Hua, Yihao Liu, Zonglin Li, Huayu Chen, Xiaoye Qu, Yafu Li, Weize Chen, Zhenzhao Yuan, Junqi Gao, Dong Li, Zhiyuan Ma, Ganqu Cui, Zhiyuan Liu, Biqing Qi, Ning Ding, and Bowen Zhou. A Survey of Reinforcement Learning for Large Reasoning Models, October 2025b. URL http://arxiv.org/abs/2509.08827. arXiv:2509.08827 [cs].

Yusen Zhang, Ruoxi Sun, Yanfei Chen, Tomas Pfister, Rui Zhang, and Sercan Ö Arik. Chain of Agents: Large Language Models Collaborating on Long-Context Tasks, June 2024. URL http://arxiv.org/abs/2406.02818. arXiv:2406.02818 [cs].

Zeyu Zhang, Quanyu Dai, Xiaohe Bo, Chen Ma, Rui Li, Xu Chen, Jieming Zhu, Zhenhua Dong, and Ji-Rong Wen. A survey on the memory mechanism of large language model-based agents. *ACM Transactions on Information Systems*, 43(6):1–47, 2025c.

Andrew Zhao, Daniel Huang, Quentin Xu, Matthieu Lin, Yong-Jin Liu, and Gao Huang. Expel: Llm agents are experiential learners. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 19632–19642, 2024.

Huichi Zhou, Yihang Chen, Siyuan Guo, Xue Yan, Kin Hei Lee, Zihan Wang, Ka Yiu Lee, Guchun Zhang, Kun Shao, Linyi Yang, et al. Memento: Fine-tuning llm agents without fine-tuning llms. *arXiv preprint arXiv:2508.16153*, 2025a.

Rongjia Zhou, Chengzhuo Li, Carl Yang, and Jiaying Lu. ClinNoteAgents: An LLM Multi-Agent System for Predicting and Interpreting Heart Failure 30-Day Readmission from Clinical Notes, December 2025b. URL https://arxiv.org/abs/2512.07081v2.

Yuxin Zuo, Kaiyan Zhang, Shang Qu, Li Sheng, Xuekai Zhu, Biqing Qi, Youbang Sun, Ganqu Cui, Ning Ding, and Bowen Zhou. TTRL: Test-Time Reinforcement Learning, April 2025. URL http://arxiv.org/abs/2504.16084. arXiv:2504.16084 [cs].