Supplementary Information
Recursive Self-Improvement in Healthcare AI Agents

Dengzhe Hou, Zihao Wu, Yuwen Zeng, Lingyu Jiang, Fangzhou Lin, Kazunori D Yamada

4 November 2026

Supplementary Table S3 provides complete mechanism summaries for all 85 mapped studies. Extraction records (S1), screening records and exact queries (S2), and the 25-study imaging view (S4) are available with the evaluation templates at `https://github.com/dengzhe-hou/rsi-healthcare-agents/tree/evidence-2026-11-04-v0.3.2/resources`. References in this PDF are numbered independently of the main article.

Supplementary Table S3. Complete study-level mechanism evidence

Each study appears once across the fourteen continued table blocks. Level definitions and the aggregate profile are in the main text. Publication tags describe source status, not clinical validation; SD denotes standard deviation. Clinical setting follows reported evaluation design; Coach Iris is retrospective, with its operating patient service recorded separately. The highest evidence levels, clinical-setting counts and inclusion decisions are unchanged.

Table S3: Study-level evidence: memory and experience across encounters. J/C = published journal or conference paper; A = officially listed accepted paper; A* = acceptance stated by authors in the arXiv record; P = preprint examined. Status describes publication, not clinical validation. Full sources and extraction fields appear in Supplementary Table S1.

Study/status	Changes and fixed components	Feedback and timing	Evaluation	Evidence boundary
Agent Hospital [P] [1]	Case database and principles change; model and reflection design fixed.	Synthetic reference decisions; retained across simulated encounters.	Separate simulated tests, MedQA and memory ablations.	Experience reuse; no changing-improver control or prospective care.
Agent Mental Clinic [P] [2]	Tertiary diagnostic memory changes; psychiatrist backbone fixed.	Supervisor feedback after simulated psychiatric interviews.	Dialogue-derived train/test cases; memory ablations.	Simulated diagnosis, with unchanged memory-update procedure.
MedAgentSim [J/C] [3]	Successful cases and reflections retained; doctor model fixed.	Simulated patient/measurement interactions and answer feedback.	Diagnostic simulations; component ablations.	Full-system gain combines memory, reasoning and measurement; not patient-care evidence.
GSEM [J/C] [4]	Graph-structured diagnostic memory and weights change; update rules fixed.	Feedback-driven experience updates over case iterations.	Clinical-reasoning/QA benchmarks and component analysis.	Snapshot/evolution-case separation insufficiently specified; gains not monotonic.
Evo-MedAgent [P] [5]	Error patterns, successful strategies and preferences change; models fixed.	Reference answer revealed after each chest-X-ray case.	Sequential ChestAgentBench evaluation and memory ablations.	Retrospective adaptation; proposed tool governance lacks quantitative ablation.
DxEvolve [P] [6]	Diagnostic cognition profiles accumulate; construction prompts and retrieval fixed.	Diagnosed historical cases distilled before later evaluations.	Non-overlapping MIMIC-CDM cases; 293 external hospital cases; physician profile review.	External retrospective evidence; improved experience assets, not an isolated updater gain.
HealthClaw [P] [7]	Profile, procedures and episodes change; shared rules and induction design fixed.	Post-episode writeback; disabled during probes.	Hidden-ledger synthetic-year probes and nine biomedical tasks.	A synthetic year is not clinical follow-up; task comparisons also differ in models.
EMR [P] [8]	Principles, patterns and cases change; weights, retrieval and merge rule fixed.	Gold-answer trajectory feedback; offline mining iterations.	Separated mining/test subsets; hierarchy and transfer analyses.	Several custom dataset halves; library transfer does not isolate an evolving updater.
PrivMeSA [P] [9]	Lesson memory grows and LoRA policy is RL-trained; distiller, judge and specialists fixed.	Recorded labels and privacy penalties in training; label-free lesson distillation after episodes.	Patient-disjoint MIMIC-IV-ED replay; baselines, ablations and paired memory-on/off run.	Retrospective replay with synthetic identifiers; single run; no evolving-versus-fixed improver test.
AgentClinic [J/C] [10]	Doctor-written diagnostic notebook is overwritten after each case; models and meta-prompt fixed.	Correct diagnosis shown after each simulated case; 1,000-character cap.	AgentClinic-MedQA tool experiments across six backbones.	Notebook written and tested on one case stream; gains vary in sign across models.
MedExpMem [J/C] [11]	Self-written pairwise differential-diagnosis notes change; VLM, retriever and prompts fixed.	Ground-truth diagnosis and expert discussion, offline over 6,228 pre-2025 cases.	Temporal split of 227 Eurorad 2025 cases; seven VLMs; RAG comparison and ablations.	Retrospective teaching-file accuracy; no variance or tests; construction procedure fixed.

Table S3: Study-level evidence: memory and experience across encounters (continued). Abbreviations as in the first part of this table.

Study/status	Changes and fixed components	Feedback and timing	Evaluation	Evidence boundary
EHRAgent [J/C] [12]	Changes a memory of successful code solutions; GPT-4, tools and prompts stay fixed.	Ground-truth success labels gate memory admission, online across the evaluated question stream.	No separate held-out split; ablation without memory on MIMIC-III and eICU benchmarks.	Offline EHR QA benchmarks; labels used at test time; no clinician or prospective evaluation.
ReflecTool [J/C] [13]	Changes trajectory memory and tool-wise experience; Qwen2 backbone, toolbox and verifiers fixed.	Ground-truth answers on a small training subset, offline before frozen inference.	Separate CAB test samples; ablations on 200 samples per dimension, single runs.	Offline public benchmarks; trajectory correctness unverified; no clinician or prospective evaluation.
MDTeamGPT [J/C] [14]	Changes correct-case and error-reflection knowledge bases; backbone and consultation protocol stay fixed.	Ground-truth correctness routes each training consultation; accumulated offline before testing.	Official test sets, five runs, KB ablation and cross-dataset transfer.	Offline QA and VQA benchmarks; labels gate admission; no deployment or prospective evaluation.
MACD [P] [15]	Self-learned diagnostic knowledge text changes; LLMs, summarizer and refiner fixed.	Correct diagnoses on learning cases; offline accumulation before frozen held-out testing.	3,076 held-out MIMIC cases; guideline knowledge, CoT, LoRA baselines; refiner ablation; reader study.	Single-site retrospective text; non-random test residue; knowledge procedure fixed, untested.
KEA [J/C] [16]	Experience and reflection banks grow; backbone LLM, retriever and knowledge base fixed.	Gold department labels gate storage during offline training-set rounds before testing.	Held-out 30% PG-Bench split; component ablation; frequency-shift test; three backbones.	Synthetic triage dialogues; label-gated memory reuse, not clinical deployment or improver change.
SpAgents [J/C] [17]	Labelled case memory grows; LLM agents, imaging tool and retrieval fixed.	Physician labels written immediately via learning task; experimental update timing unreported.	Internal validation, 51-patient external test, memory ablation, seven-physician reader comparison.	Retrospective; memory contents during testing unreported, so static labelled-case retrieval possible.
FHIR-AgentEval [J/C] [18]	Reflection memory changes; evaluation agent, reflector, validators and tools fixed.	Validator outcomes and evaluator critique during a separate offline training phase.	Unseen variations and 22 held-out tasks; five configurations; three runs each.	FHIR sandbox administrative tasks; cross-task transfer smaller; no clinical outcome evidence.
ClinicalReTrial [J/C] [19]	Local and global redesign memory change; agents, predictor reward and validator fixed.	Predictor-derived rewards each iteration; global memory updated after each trial converges.	60 held-out failed trials, paired ablations, independent HINT re-scoring, retrospective case studies.	Predicted trial success only; global-memory contribution not isolated from local memory.
R4Agent [P] [20]	Exemplar memory grows from outputs; VLM backbones, router and agent prompts fixed.	Self-critique and pass@k selection per case; outputs written to memory during inference.	IU X-ray reports, VinBigData boxes; model-judge and overlap metrics; pass@k ablation.	Memory contribution not isolated; stored exemplars are unverified self-generated outputs.
AgentEHR [J/C] [21]	Experience memory bank grows; LLM weights, prompts, tools and retriever fixed.	Ground-truth comparison on 100 training cases per task, offline before testing.	Held-out MIMIC-IV cohorts, two shifted subsets, evolved-versus-plain ablations, memory baselines.	Single-center retrospective EHR; memory built once; train-test disjointness not stated.
TheraAgent [P] [22]	Case-outcome memory and subgroup patterns grow; GPT-4o agents and trial knowledge fixed.	Documented PSA and survival outcomes added after treatment; evaluation ordering not reported.	35 real patients, 400 synthetic cases, agent baselines, memory-removal ablation.	Single-center tiny cohort; possible self-retrieval unaddressed; no significance testing.

Table S3: Study-level evidence: memory and experience across encounters (continued). Abbreviations as in the first part of this table.

Study/status	Changes and fixed components	Feedback and timing	Evaluation	Evidence boundary
OpenHospital [P] [23]	Reflection-derived physician experience accumulates; LLM weights and arena fixed.	Ground-truth diagnosis, exams and treatment critique after each case, 22 batches.	Metric trajectories across training batches; no ablation, comparator, or repeated runs.	Synthetic text-only simulation; retained artifact and test-set use unspecified.
SkinGPT-X [P] [24]	Case memory and per-disease guidelines evolve; vision, diagnosis and review agents fixed.	Expert labels of stored training cases; threshold-triggered guideline rewrites, offline.	Held-out splits of six datasets, memory ablation, blinded dermatologist review.	Offline public data; test-time memory growth unreported; no prospective evaluation.
SEA [P] [25]	Case and rule memory change in stream; RL-trained weights, candidate sets fixed	Ground-truth label after each case; offline RL, then online memory writes	ER-Reason single-pass stream, untrained-writer memory controls; MedCaseReasoning test without carry-over	Closed-set 200 candidates, one stream order, no intervals, one-physician rule rating
Traj-Evolve [P] [26]	Experience pool and LoRA agent weights change; architecture, embedder, reward rule fixed	Ground-truth cancer labels; offline rejection sampling and one fine-tuning epoch	Held-out overall and never-smoker test sets versus static Traj-CoA and ML baselines	Single-institution retrospective case-control; simulated patient stream; single training round only
MedEvoEval [A*] [27]	Rule-card memory bank grows; doctor, simulator, scorer and case order fixed	Manager rubric scores; online write-back after every simulated episode	No-memory vs mature-memory on 200 held-out cases with Holm-corrected paired tests	Simulated episodes; repeated-stream gains confounded; held-out outcome gains not significant
TopoAgent [P] [28]	Within-dataset failure-correction memory grows; GPT-4o, tools, distilled skill set fixed	LLM self-check of feature quality; memory written online after failed reflections	Balanced accuracy on 26 benchmark datasets; memory ablations; three external datasets	Memory reset per dataset; skills distilled from evaluation benchmark; non-clinical endpoint
SAGEAgent [J/C] [29]	Episodic store and reflected rules change; frozen LLM, predictor and rewards fixed	Uncertainty and burden rewards; reflection every 10 training patients before testing	Nested 5x5 cross-validation with memory ablations, fusion, RL and Reflexion baselines	Small public glioma cohort; uncertainty-based reward; test-time memory updating unreported
MIRA [P] [30]	VLM weights and principle memory change; tools, judges, validation gate fixed	Judge-scored rewards; memory edits gated on validation during online RL	Nine medical VQA test splits; RL ablation isolating reflection memory contribution	Benchmark VQA only; single runs; memory role in reward not isolated
AMR [A*] [31]	Per-agent reflection memories change; GPT-4o, routing graph and RAG corpus fixed	Ground-truth label on wrong answers; test-time write-back disabled, filling phase unreported	MedQA test and MedMCQA subset; batches of 50, sizes unreported	Exam multiple-choice QA only; no intervals; memory provenance at test unclear
CDEG [P] [32]	Evidence graph built from trajectories; doctor backbones, tools and environment fixed	Reference-diagnosis correctness plus counterfactual verifier; one offline construction pass	200 patient-disjoint MIMIC-IDx cases and 107 AgentClinic OOD cases	Interactive retrospective records; LLM-judged accuracy; no repeated runs reported
LearnActCoder [P] [33]	Role-routed error memory changes; GPT-5, Coder-Judge workflow and lookup table fixed	Reference-code mismatches from 25 LEARN notes; built once offline, then frozen	Held-out matched MIMIC-III (150) and MIMIC-IV (250) notes; 1,000-note stability run	Retrospective coding; one model; CPT gain only; ICD F1 unchanged

Table S3: Study-level evidence: skills, tools and workflow adaptation. J/C = published journal or conference paper; A = officially listed accepted paper; A* = acceptance stated by authors in the arXiv record; P = preprint examined. Status describes publication, not clinical validation. Full sources and extraction fields appear in Supplementary Table S1.

Study/status	Changes and fixed components	Feedback and timing	Evaluation	Evidence boundary
EvoClinician [P] [34]	Actor prompts and action-conditioned memory change; Evolver and Grader fixed.	Process feedback and diagnosis score in a sequential case stream.	Case-stream evaluations and memory/strategy ablations.	Judge score feeds adaptation; not an independent frozen release test.
EvoMDT [J/C] [35]	Prompts and agent configuration adapt; framework/evaluator prescribed.	Expert development-case feedback, metrics and reported post-task outcomes.	QA datasets, retrospective records and blinded physician scoring.	Evolution gain not isolated; automatic release criteria not detailed.
PsychAgent [P] [36]	Skills and shared model parameters change; session-selection rule fixed.	Simulated counseling feedback; skill evolution and rejection fine-tuning.	Separate client profiles; human ratings of simulated dialogues.	Shared-model change can affect skill generation; its contribution is not isolated.
SkeMex [A] [37]	Skill repository and utility state change; online update operators fixed.	Task trajectories, novelty/quality and utility; offline and online phases.	Held-out in-domain cases, unseen benchmark families and cross-model skill transfer.	Transfer of skills is tested; improvement of update operators is not established.
HealthFlow [J/C] [38]	Plans, workflow code and experience change; writeback/retrieval rules fixed.	Execution outcomes; reflection and outcome-aware memory after tasks.	External model judge and blinded expert report review; online memory runs.	Independent assessment exists; benchmark-time memory and clinical efficacy remain distinct.
MedRSI [P] [39]	Tools and trainable tool models change; reasoner, execution and judge fixed.	Failure-driven capability construction; candidate validation before registration.	Public imaging tests and frozen-checkpoint retrospective cohorts.	Generated capabilities may aid later construction; improver benefit not independently isolated.
MedForge-RSI [P] [40]	Experience, tools and optional directives change; detector and acceptance design fixed.	Verified errors; replay/validation checks over 20 rounds.	Disjoint final 4,000-image test; mechanism/seed analyses.	Bounded self-refinement; one earlier resample set contaminated; no clinical trial.
TCM-SEA [P] [41]	Clinical skill document edited; LLM, edit loop, gate and simulator fixed.	Expert monograph analysis of failed held-in cases; nine offline edit rounds.	Gate on held-in and validation splits; sealed test split; outcome-only ablation.	Cases from published monographs with a lookup simulator; one seed; no physician review.
Asclepius [J/C] [42]	Operating manual rewritten; backbone, skills library, subagent prompts and proposer fixed.	Judge scores and traces between offline search-batch iterations.	Four held-out simulated emergency-department batches; add-one ablations and manual-only contrast.	Simulated shifts; evolved manual alone adds +0.02 on held-out batches, not significant.
Learning to Be A Doctor [J/C] [43]	Changes workflow graph and node prompts; backbone LLM, tools and evolution procedure fixed.	Ground-truth diagnostic errors on 50 training images, offline batches with validation checks.	Held-out test images; static baselines and operation ablations; single seed, three backbones.	Small dermatology splits, one synthetic dataset; offline only, no clinician evaluation.
MedDCR [J/C] [44]	Coding workflows and design archive change; Designer, Reflector and selection rule fixed.	Validation F1 score plus Reflector critique across 100 offline search iterations.	MDACE and ACI-BENCH micro-F1; archive and feedback ablations; no separate test split described.	Benchmark workflow search; selection and reporting data separation unclear; no clinical deployment.

Table S3: Study-level evidence: skills, tools and workflow adaptation (continued). Abbreviations as in the first part of this table.

Study/status	Changes and fixed components	Feedback and timing	Evaluation	Evidence boundary
CodeClinic [P] [45]	Verified function library grows; inference weights, verifier, threshold, guidelines fixed	Accuracy on labeled verification cases with error traces; offline, then frozen	56,700 held-out queries and separate ICU surveillance task; zero-shot, GRPO controls	Single-center MIMIC-IV, one builder model, small verification sets, LLM-written guidelines
MediSkill-Evo [P] [46]	Four typed knowledge banks change; backbone, harness, simulator and evaluator fixed	Post-encounter label-based reflection; offline over 700 training encounters, frozen at test	300 held-out MIMIC-IV encounters, 180 stress conditions, 100 NEJM cases	Simulated encounters; LLM judge; one rollout; no clinician validation

Table S3: Study-level evidence: parameter updates and changing training generators. J/C = published journal or conference paper; A = officially listed accepted paper; A* = acceptance stated by authors in the arXiv record; P = preprint examined. Status describes publication, not clinical validation. Full sources and extraction fields appear in Supplementary Table S1.

Study/status	Changes and fixed components	Feedback and timing	Evaluation	Evidence boundary
TissueLab [P] [47]	Tool classifiers and model preferences change; foundations/framework fixed.	Clinician annotations, acceptance feedback and active learning.	Retrospective imaging/omics tasks and patient-level splits.	Human-guided adaptation; autonomous experience-to-skill synthesis unclear.
MACRO [P] [48]	Memory, composite tools and policy change; discovery/training design fixed.	Training episodes; correctness gates retention, composite use earns reward.	Public imaging benchmarks and static/component comparisons.	Persistent artifacts arise during training; deployment-time learning not established.
VIBEMed [J/C] [49]	Memory/model/code changes described; coordinated update workflow prescribed.	Case feedback, model judges and sandbox tests; user confirms code release.	Base-versus-tuned MedBench comparison and oncology demonstration.	Evolution components not isolated; two evaluated task scores decline after tuning.
EvoPatient [J/C] [50]	Doctor trajectory memory changes patient-training interactions; validation rules fixed.	Simulated doctor/patient feedback across coevolution rounds.	Doctor-question and patient-ability analyses; disease transfer.	Interaction-generator change tested; patient gains use bundled doctor ablations.
MedZERO [A*] [51]	Examiner and Reasoner parameters plus knowledge graph change; reward/curation rules fixed.	Alternating frontier-focused question generation and RL training.	Medical benchmarks and curriculum-source/component analyses.	Learned generator changes; its isolated benefit versus a frozen Examiner not established.
ECG-RSI [P] [52]	Waveform operators, prompts, skills and learner change; acceptance criteria fixed.	Construction feedback repairs generator; separate feedback and selection cohorts.	Patient-separated tests; frozen-generator-component controls; physician sample review.	Positive bounded generator evidence; accepted-data budgets match, candidate counts differ; total-cost matching unestablished.
ReVision3D [P] [53]	Perception adapters, training recipe and inference program change; LLM designer, verifier and objective fixed.	Mask-grounded attribution and development objective; revisions kept above seed noise.	Test CT patients, partly reused from earlier comparisons; frontier, recipe-control and designer comparisons.	Matched-compute recipe control on retrospective CT; recipe origin inconsistently reported; no improver-revision test.
UltrasoundAgents [P] [54]	Sub-agent and main-agent VLM weights change; tools, taxonomy, rewards and recipe fixed.	Ground-truth labels drive GRPO stages, then self-distilled trajectory SFT, all pre-test.	Three in-domain test splits, external BrEaST OOD set, stage ablations, oracle upper bounds.	One-off training on public data; frozen at test; low BI-RADS agreement.
Healthcare AI GYM [P] [55]	Qwen3.5-9B weights change by GRPO/TT-OPD; gym tasks, tools, reward fixed	Episode 5D reward plus EMA-teacher KL; online RL during 60 training steps	18 benchmarks with same tools; Base+AR and GRPO controls; validation 307 tasks	Train tasks from same benchmark families, single run, no intervals, inconsistent counts
SepsisAgent [P] [56]	Qwen3-4B weights change across SFT, BC, GRPO; world model and rewards fixed	Simulated survival, SOFA, lactate and guideline reward; offline staged training only	725 held-out MIMIC-IV episodes by OPE; untrained world-model agents as controls	Off-policy estimates, same-source world model, no seeds, no clinician evaluation
Cura 1T [P] [57]	Training mixture and 1T weights change; harness skills, benchmarks, human review fixed	Graded benchmark trajectories; offline keep-or-revert rounds per capability loop	Same healthcare benchmarks used for selection; out-of-domain checks not used for selection	No independent healthcare test; no ablations; human decides each round

Table S3: Study-level evidence: parameter updates and changing training generators (continued). Abbreviations as in the first part of this table.

Study/status	Changes and fixed components	Feedback and timing	Evaluation	Evidence boundary
BaT [P] [58]	Agent policy and stage curriculum change; execution environment, verifier, controller fixed	Rubric-verified rewards and held-out stage scores; offline multi-round GRPO	AutoMedBench-Lite also steers training; ABRA, MedXpertQA, external benchmarks unselected	Benchmark guides training; no matched ablation; research workflows rather than care
CASE [P] [59]	8B VLM policy weights change; tools, judge, evidence replay and prompts fixed	Answer, rubric and self-distillation rewards; offline RSFT then one GRPO epoch	4,850 questions from 455 held-out UK Biobank participants, three roles	Single cohort; MCQ labels from ICD-10; LLM judges; no external site

Table S3: Study-level evidence: diagnostic training and informative adaptation comparators. J/C = published journal or conference paper; A = officially listed accepted paper; A* = acceptance stated by authors in the arXiv record; P = preprint examined. Status describes publication, not clinical validation. Full sources and extraction fields appear in Supplementary Table S1.

Study/status	Changes and fixed components	Feedback and timing	Evaluation	Evidence boundary
ClinicalAgents [J/C] [60]	Within-case working state changes; experience store remains static.	Current-case evidence, confidence and search feedback.	Experience/test separation and clinician-checked model judge.	Within-case orchestration comparator; no new cross-case experience learned.
DiagAgent [P] [61]	Diagnostic policy parameters change; DiagGym and reward procedure fixed.	RL feedback on virtual EHR-derived consultations.	Physician-validated diagnostic benchmark cases.	Learner improvement in a simulated environment; no changing-improver test.
DoctorAgent-RL [J/C] [62]	Doctor policy updates; interaction and reward design fixed.	Multi-turn patient-simulator feedback during training.	Dialogue-derived simulated consultation evaluation.	Conference-aligned version; later extended-manuscript clinical results not imported.
OAIMS [J/C] [63]	Segmentation weights adapt; online update procedure fixed.	Label-derived simulated correction clicks, within/across cases.	Retrospective streams with imaging modality/pathology shifts.	Non-LLM comparator; simulated clicks do not establish clinician-in-the-loop outcomes.
ReclAIm [J/C] [64]	Image-classifier weights adapt; agent workflow and alarm rule fixed.	Labeled monitoring data trigger classifier fine-tuning.	Separate fine-tuning cohort; same monitoring cohort used for reassessment.	Automated maintenance, not updater learning; no untouched final external test.
TEMPLAR [P] [65]	Radiology reporting template and knowledge graphs built once; LLMs, prompts and rules fixed.	Literature grounding and one structuring audit on training reports.	Held-out PET/CT and CT reports; ablations, cross-dataset transfer and expert ratings.	One-off offline construction, not post-deployment learning; mostly LLM-judged metrics.
VDSS [J/C] [66]	Clinician preference bandit state updates; LLM agents, safety checks, backbones fixed	Clinician accept/reject and accepted settings; online update at each cycle close	Multi-center retrospective replay and expert ratings of 100 cycles; NoPref ablation	Learning shown only as one-clinician regret trend; bandit and patient span unstated
ClinVerse [P] [67]	Memories, skills, playbooks, tool registry change; Qwen3-32B, simulator, rule-based judge fixed	Rule-based EHR-grounded trajectory scores; offline 5,000 or online 1,000 trajectories	100 held-out EHR cases, five runs; zero-shot transfer to AIHospital	Simulated EHR environment; borrowed methods; gains often within one SD; preprint

Table S3: Study-level evidence: Memory and experience across encounters (continued, OpenAlex additions). Status tags follow Table S3.

Study/status	What changes; what stays fixed	Feedback and timing	Evaluation	Evidence boundary
EMAI [J/C] [68]	Report-derived experience rules evolve; backbone and rule induction/selection procedure remain fixed.	Impression scores filter and refine cluster rules offline for later unseen reports.	MIMIC-CXR and Open-I report summarisation with component ablations and clinician ratings.	Task rules evolve under fixed induction; no changing-improver control or prospective patient evaluation.
SeM-Agents [J/C] [69]	Correct-case and error-reflection banks grow; backbone, consultation protocol and reviewer prompts stay fixed.	Reference answers route training consultations into separate banks before frozen test evaluation.	MedQA and PubMedQA test sets, memory variants and two-backbone evolution curves.	Training-only memories; no prospective care or changing-improver control; MedQA option counts differ.
Shennong-Agent [J/C] [70]	Clinician corrections and prompts enter vector memory; Qwen-7B and module design remain fixed.	Human edits are encoded for reuse in later similar pharmaceutical-care situations.	461 retrospective outpatient records; pharmacist scoring and with/without human-intervention comparison.	Persistent-memory contribution not isolated; no adaptation/test split or prospective deployment evaluation.
U-EHR [J/C] [71]	Positive/negative memory and retrieval estimates change; backbone, scorers and UCB rules fixed.	Prior action scores and task success update the bank after each clinical task.	MIMIC EHR tasks and private CN-EHR, four backbones; sequential adaptation and ablations.	No separate memory-learning holdout or prospective patient outcomes; fixed update procedure.
SR-MAPR [J/C] [72]	Physician-assessment heuristics change; backbone, reflection routine and retrieval design remain fixed.	Expert-rating discrepancies trigger error reflections and rules retained for later profile evaluations.	7362 physician profiles, stated80/20 split, backbone comparisons and rules/case/RAG ablations.	Professional-profile benchmark; rule-update/test scheduling unspecified; no patient outcomes or improver control.

Table S3: Study-level evidence: Memory and experience across encounters (continued, OpenAlex additions). Status tags follow Table S3.

Study/status	What changes; what stays fixed	Feedback and timing	Evaluation	Evidence boundary
MDTace [J/C] [73]	Structured consultation bullets change; backbone, lead/reflector prompts and memory procedure remain fixed.	Training-case deliberation and tool-use feedback are distilled into a global reusable Playbook.	Four clinical QA/VQA test sets; training-only bullets, Playbook and context-layer ablations.	Offline benchmarks; no changing-improver control or prospective care; first online2026, imprint2027.

Table S3: Study-level evidence: Skills, configurations and agent policies (continued, OpenAlex additions). Status tags follow Table S3.

Study/status	What changes; what stays fixed	Feedback and timing	Evaluation	Evidence boundary
GuideSkill [P] [74]	Executable diagnostic scoring skills evolve; backbone, compiler and skill generator remain fixed.	Labelled case rationales distilled offline before evaluation on different clinical benchmark cases.	Four diagnostic benchmarks; evolved skills assessed across four fixed task backbones.	No prospective evaluation; learned task skills do not revise the improvement procedure.
Role-based norm agent [J/C] [75]	Norm memory and role prompts evolve; Qwen-plus weights and feedback routines remain fixed.	Fixed AI feedback drives prompt edits and periodic identity changes across episodes.	Three simulated healthcare communication tasks compare full framework and component ablations.	No prospective evaluation; judge sampling and identity-consistency measurement are reported inconsistently.
MedGuard [J/C] [76]	Agent policy weights change; prompts, retrieval knowledge base and training generator stay fixed.	Full risk-alerting trajectories receive format and outcome rewards during offline GRPO.	Separate consultation benchmarks, retrospective records and expert ratings assess risk-alerting outputs.	Patient memory and planner refinement are case-local; no prospective patient outcome evaluation.
SEMPA [J/C] [77]	Prescribing and Critic policies evolve; backbone, rewrite routine and safety rules stay fixed.	Cross-encounter utility pools guide periodic rewrites before frozen held-out encounter evaluation.	One hundred retrospective MIMIC-IV encounters, six runs and component ablations.	Critic scores inform later updates; its evolving evaluator role is not isolated.
Skill-CDPO [J/C] [78]	Experience-derived tool skills and agent weights change; expert and selection procedure stay fixed.	Historical invocation failures form skills; rollout outcome gaps guide offline preference optimization.	Three medical retrieval and QA benchmarks compare skills, SFT and DPO variants.	One-round parameter optimization; skill reuse supports eligibility, without changing-improver causal control.

Table S3: Study-level evidence: Skills, configurations and agent policies (continued, OpenAlex additions). Status tags follow Table S3.

Study/status	What changes; what stays fixed	Feedback and timing	Evaluation	Evidence boundary
Gated adapter / OpenClaw [J/C] [79]	Persistent guidance and final adapter gate change; encoders and update routine remain fixed.	Report-simulated physician feedback updates special-case memory and guidance every hundred iterations.	Private fundus images; conflict counts assessed across repeated adaptation rounds.	Agent backbone unspecified; reused test cases and conflict counts do not establish independent accuracy.
BridgeGap [J/C] [80]	Diagnostic dialogue policy learns; patient simulator, evidence verifier and reward routine stay fixed.	Environment-verified belief progress and terminal diagnosis guide PPO across simulated patient episodes.	HealthBench, AgentClinic and static medical QA assess the resulting frozen policy.	Belief state is case-local; learned weights persist, without prospective patient outcome evidence.
Evolved decision rules [P] [81]	Medical VQA prompt-program configurations evolve; proposer, VLM and selection pipeline remain fixed.	Evolution-case accuracy guides offline search, followed by separate holdout selection and testing.	MedFrameQA internal video-grouped split; five main-method repeats on 855 final-test examples.	No prospective evaluation or revised improver; strategy feedback subset sizes differ.
Living-OEP [J/C] [82]	Annotation instructions and examples change; clinical LLM backbones and human update procedure fixed.	Human annotation corrections guide repeated instruction refinement before locked later-publication validation.	1997 independent breast-cancer publications and eight clinical evidence-retrieval scenarios; human quality review.	No patient outcomes or isolated updater control; trial-level evidence interpretation remains human.

Table S3: Study-level evidence: Skills, configurations and agent policies (continued, OpenAlex additions). Status tags follow Table S3.

Study/status	What changes; what stays fixed	Feedback and timing	Evaluation	Evidence boundary
Coach Iris / FAST [J/C] [83]	Shared prompts, retrieval data and responses change; backbone and optimization details unreported.	Human reviewers flag dialogue failures; clinical experts and the ML team refine behavior.	Framework face validity on recorded dialogues; five reviewers initially assess twelve dialogues.	Deployed service monitoring; no prospective adaptation trial or quantified post-update clinical benefit.
RHIS [P] [84]	Action-governance rules change; Kimi backbone, benchmark source and change-admission protocol remain fixed.	Execution failures trigger versioned controller revisions between frozen benchmark runs and stages.	300 virtual-EHR tasks; named same-stage retests separated from cross-version first-pass totals.	No held-out update-transfer or frozen-improver control; candidate governance labels await independent review.

Table S3: Study-level evidence: Training generators and evaluators (continued, OpenAlex additions). Status tags follow Table S3.

Study/status	What changes; what stays fixed	Feedback and timing	Evaluation	Evidence boundary
MedEnv / Diagnosis-R1 [J/C] [85]	Doctor policy changes and supplies a specified reward likelihood; GRPO and simulators fixed.	Interactive rollouts receive diagnosis outcomes, confidence-gain rewards and unsupported-action penalties.	Two thousand held-out cases per author response; matched reward/training ablations.	Simulation; scoring-weight synchronization and patient-level split unspecified; evolving scorer not isolated.

Supplementary references

[1] J. Li, Y. Lai, W. Li, J. Ren, M. Zhang, X. Kang, S. Wang, P. Li, Y.-Q. Zhang, W. Ma, Y. Liu, Agent hospital: A simulacrum of hospital with evolvable medical agents (2024). arXiv:2405.02957.

[2] K. Lan, B. Jin, Z. Zhu, S. Chen, S. Zhang, K. Q. Zhu, M. Wu, Depression Diagnosis Dialogue Simulation: Self-improving Psychiatrist with Tertiary Memory, arXiv preprint arXiv:2409.15084Version 2, 2024-10-09 (2024). doi:10.48550/arXiv.2409.15084.
URL https://arxiv.org/abs/2409.15084

[3] M. Almansoori, K. Kumar, H. Cholakkal, MedAgentSim: Self-evolving multi-agent simulations for realistic clinical interactions, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2025, Vol. 15968 of Lecture Notes in Computer Science, Springer Nature Switzerland, 2025, pp. 362–372. doi:10.1007/978-3-032-05114-1_35.
URL https://papers.miccai.org/miccai-2025/0537-Paper2575.html

[4] X. Han, Y. Fan, S. Zhao, H. Wang, B. Qin, GSEM: Graph-based self-evolving memory for experience augmented clinical reasoning, in: Proceedings of the 2026 Conference on Empirical Methods in Natural Language Processing, 2026, pp. 11637–11660. doi:10.18653/2026.emnlp-main.507.
URL https://aclanthology.org/2026.emnlp-main.507/

[5] W. Shen, B. Jian, J. Li, C. Liu, J. Moll, X. Hu, D. Rueckert, H. B. Li, J. Pan, Evo-MedAgent: Beyond One-Shot Diagnosis with Agents That Remember, Reflect, and Improve, arXiv preprint arXiv:2604.14475Version 1, 2026-04-15 (2026). doi:10.48550/arXiv.2604.14475.
URL https://arxiv.org/abs/2604.14475

[6] R. Ren, Y. Wang, Y. Liang, L. Luo, J. Liu, H. Wang, C. Feng, Y. Zhang, C. Miao, J.-R. Wen, W. X. Zhao, Emulating Clinician Cognition via Self-Evolving Deep Clinical Research, arXiv preprint arXiv:2603.10677Version 1, 2026-03-11 (2026). doi:10.48550/arXiv.2603.10677.
URL https://arxiv.org/abs/2603.10677

[7] H. Li, J. Deng, T. Jin, J. Han, Y. Wang, Z. Wang, Q. Si, W. Gong, X. Zhuang, J. You, W. Cheng, J. Feng, H. Guo, A Self-Evolving Agent for Longitudinal Personal Health Management, arXiv preprint arXiv:2607.13940 (2026). arXiv:2607.13940.
URL https://arxiv.org/abs/2607.13940v1

[8] D. Shi, Y. Li, X. Yi, L. Wang, EMR: Self-Evolving Medical Multi-Agent System via Experience Mining and Reuse, arXiv preprint arXiv:2609.15161 (2026). arXiv:2609.15161.
URL https://arxiv.org/abs/2609.15161v1

[9] D. Wang, Y. Zhang, B. Sun, A. Stinard, Y. Shang, S. Wang, Y. Tian, PrivMeSA: Privacy-Aware Self-Evolving Multi-Agent System for Medicine via Local-Remote LLM Collaboration (2026). arXiv:2609.38458, doi:10.48550/arXiv.2609.38458.
URL https://arxiv.org/abs/2609.38458v1

[10] S. Schmidgall, R. Ziaei, C. Harris, J. W. Kim, E. P. Reis, J. Jopling, M. Moor, AgentClinic: a multimodal benchmark for tool-using clinical AI agents, npj Digital Medicine 9 (2026) 499. doi:10.1038/s41746-026-02674-7.

[11] Q. Feng, Z. Huang, Y. Zhu, Y. Gu, W. C. W. Chu, X. Zhang, Q. Dou, MedExpMem: Adapting experience memory for differential diagnosis, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2026, Lecture Notes in Computer Science, Springer, 2026, pp. 430–440. doi:10.1007/978-3-032-38098-2_41.

[12] W. Shi, R. Xu, Y. Zhuang, Y. Yu, J. Zhang, H. Wu, Y. Zhu, J. C. Ho, C. Yang, M. D. Wang, EHRAgent: Code Empowers Large Language Models for Few-shot Complex Tabular Reasoning on Electronic Health Records, in: Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, 2024, pp. 22315–22339. doi:10.18653/v1/2024.emnlp-main.1245.

[13] Y. Liao, S. Jiang, Y. Wang, Y. Wang, ReflecTool: Towards Reflection-Aware Tool-Augmented Clinical Agents, in: Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2025, pp. 13507–13531. doi:10.18653/v1/2025.acl-long.663.

[14] K. Chen, X. Li, T. Yang, H. Wang, G. Yang, J. Huo, Y. Gao, MDTeamGPT: Mitigating context collapse and enabling self-evolution in medical multi-agent reasoning, in: M. Liakata, V. P. Moreira, J. Zhang, D. Jurgens (Eds.), Findings of the Association for Computational Linguistics: ACL 2026, Association for Computational Linguistics, San Diego, California, United States, 2026, pp. 28578–28606. doi:10.18653/v1/2026.findings-acl.1427.
URL https://aclanthology.org/2026.findings-acl.1427/

[15] W. Li, R. Yan, X. Zhang, L. Chen, H. Zhu, J. Zhao, J. Li, M. Li, W. Cao, Z. Jiang, W. Wei, K. Zhang, S. K. Zhou, MACD: Multi-Agent Clinical Diagnosis with Self-Learned Knowledge for LLM (2025). arXiv:2509.20067, doi:10.48550/arXiv.2509.20067.
URL https://arxiv.org/abs/2509.20067v5

[16] W. Lu, K. Liu, J. Wang, X. Peng, T. Shen, F. Zhu, W. Zhang, J. Zhu, T. Xin, A. V. Vasilakos, Advancing Chinese Conversation-based Patient Guidance with a Benchmark and Knowledge-Evolvable Assistant, IEEE Journal of Biomedical and Health Informatics (2025) 1–12doi:10.1109/jbhi.2025.3639805.

[17] X. Ji, Z. Li, L. Zeng, L. Hu, Y. Wang, K. Zhang, L. Shi, M. Wei, L. Chen, L. Guo, J. Dong, A. Wang, L. Sun, Y. Song, H. Wang, J. Wang, Y. Lei, W. Yue, Z. Zhao, J. Zhu, F. Huang, J. Zhang, T. Li, K. Li, Early diagnosis of axial spondyloarthritis in primary care using multi-agent systems, npj Digital Medicine 9 (1) (2026) 185. doi:10.1038/s41746-026-02372-4.

[18] Y. Mokssit, K. Ravi, M. Nie, J. Kim, C. Liu, FHIR-AgentEval: a modular sandbox for benchmarking clinical LLM agents with memory-augmented configurations, npj Digital Medicine (2026). doi:10.1038/s41746-026-03348-0.

[19] S. Xing, K. Wu, X. Xia, H. He, M. Jiang, J. Chen, T. Fu, ClinicalReTrial: Clinical trial redesign with self-evolving agents, in: H.-Y. Lee, S. Dev, J. Li, D. Elliott (Eds.), Findings of the Association for Computational Linguistics: EMNLP 2026, Association for Computational Linguistics, Budapest, Hungary, 2026, pp. 16296–16330. doi:10.18653/2026.findings-emnlp.772.
URL https://aclanthology.org/2026.findings-emnlp.772/

[20] M. F. A. Sayeedi, R. Rahman, S. T. Bhuiyan, S. Wasi, A. Islam, S. B. Alam, A. M. Rahman, Route, Retrieve, Reflect, Repair: Self-Improving Agentic Framework for Visual Detection and Linguistic Reasoning in Medical Imaging (2026). arXiv:2601.08192, doi:10.48550/arXiv.2601.08192.
URL https://arxiv.org/abs/2601.08192v2

[21] Y. Liao, C. Xuan, Y. Cai, L. Yang, Z. Chen, Y. Wang, Y. Wang, AgentEHR: Advancing autonomous clinical decision-making via retrospective summarization, in: H.-Y. Lee, S. Dev, J. J. Li, D. Elliott (Eds.), Proceedings of the 2026 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Budapest, Hungary, 2026, pp. 34870–34917. doi:10.18653/2026.emnlp-main.1529.
URL https://aclanthology.org/2026.emnlp-main.1529/

[22] Z. Chen, J. Wang, Y. Chen, X. Ji, X. Hu, J. Hong, W. A. Bosbach, A. Rominger, A. Afshar-Oromieh, H. Shan, K. Shi, TheraAgent: Multi-Agent Framework with Self-Evolving Memory and Evidence-Calibrated Reasoning for PET Theranostics (2026). arXiv:2603.13676, doi:10.48550/arXiv.2603.13676.
URL https://arxiv.org/abs/2603.13676v1

[23] P. Liu, R. Ding, Y. Mao, Z. Jiang, Y. Ye, Y. Gao, Y. Zhang, R. Sun, L. Lai, Z. Qian, OpenHospital: A Thing-in-itself Arena for Evolving and Benchmarking LLM-based Collective Intelligence (2026). arXiv:2603.14771, doi:10.48550/arXiv.2603.14771.
URL https://arxiv.org/abs/2603.14771v3

[24] Z. Chen, Y. Shen, F. Widjaja, Y. Xu, L. Sun, Z. Wang, H. Chen, W. Dai, J. Zhou, SkinGPT-X: A Self-Evolving Collaborative Multi-Agent System for Transparent and Trustworthy Dermatological Diagnosis (2026). arXiv:2603.26122, doi:10.48550/arXiv.2603.26122.
URL https://arxiv.org/abs/2603.26122v1

[25] B. Li, S. Du, Y. Guo, Joint Optimization of Reasoning and Dual-Memory for Self-Learning Diagnostic Agent (2026). arXiv:2604.07269, doi:10.48550/arXiv.2604.07269.
URL https://arxiv.org/abs/2604.07269v2

[26] S. Zeng, M. Thompson, R. Etzioni, M. Yetisgen, Traj-Evolve: A Self-Evolving Multi-Agent System for Patient Trajectory Modeling in Lung Cancer Early Detection (2026). arXiv:2606.02812, doi:10.48550/arXiv.2606.02812.
URL https://arxiv.org/abs/2606.02812v1

[27] H. Zhang, MedEvoEval: Evaluating Continual Evolution of Doctor Agents through Simulated Clinical Episodes (2026). arXiv:2606.28900, doi:10.48550/arXiv.2606.28900.
URL https://arxiv.org/abs/2606.28900v3

[28] G. Meng, P. Gu, X. Li, Y. Shi, E. W. Chambers, D. Z. Chen, TopoAgent: An Agentic Framework for Automated Topology Learning in Medical Imaging (2026). arXiv:2606.29763, doi:10.48550/arXiv.2606.29763.
URL https://arxiv.org/abs/2606.29763v1

[29] C. Qu, C. Cui, Z. Lu, J. Zhu, T. Yao, J. Guo, J. Xiong, Y. Zhu, Y. Yang, B. A. Landman, Y. Huo, SAGEAgent: A Self-evolving Agent for Cost-Aware Modality Acquisition in Multimodal Survival Prediction, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2026, 2026, pp. 536–546. doi:10.1007/978-3-032-38059-3_51.

[30] S. Wang, J. Yang, K. Wu, X. Ji, Y. Ye, Z. Chen, M. Xiong, W. Fang, M. Liu, M. Xu, M. Shan, C. Liu, B. He, Q. Liu, MIRA: Medical Image Reflection for Agentic Diagnosis (2026). arXiv:2608.10827, doi:10.48550/arXiv.2608.10827.
URL https://arxiv.org/abs/2608.10827v1

[31] P. Murugesan, L. Yang, X. Chen, X. Fan, Adaptive Memory and Reflection Multi-Agent System for Medical Question Answering (2026). arXiv:2608.19029, doi:10.48550/arXiv.2608.19029.
URL https://arxiv.org/abs/2608.19029v1

[32] X. Dai, Z. Meng, Z. Fan, Y. Tang, G. Jiang, Z. Niu, Z. Liu, CDEG: Learning Decision-Critical Evidence for Long-Horizon Diagnostic Agents (2026). arXiv:2608.22899, doi:10.48550/arXiv.2608.22899.
URL https://arxiv.org/abs/2608.22899v3

[33] M. Ghaffari, B. Sen, N. Sabetpour, N. Fatehi, A. Agarwal, C. Morato, LearnActCoder: Role-Aware Error Memory for Adaptive Clinical Coding Agents (2026). arXiv:2609.19721, doi:10.48550/arXiv.2609.19721.
URL https://arxiv.org/abs/2609.19721v1

[34] Y. He, J. Liu, Z. Hu, Y. Chen, Y. Liu, Y. Sui, Y. Li, N. Chen, J. Hu, B. Hooi, X. Xu, J. Bian, EvoClinician: A Self-Evolving Agent for Multi-Turn Medical Diagnosis via Test-Time Evolutionary Learning, arXiv v1, 2026-01-30 (2026). arXiv:2601.22964, doi:10.48550/arXiv.2601.22964.
URL https://arxiv.org/html/2601.22964v1

[35] Q. Liu, Z. Hu, T. Huang, Y. Niu, X. Zhang, S. Ma, C. Lin, G. K. Huat, H. E. Kwon, F. Gao, X. Sun, Z. Ying, G. Qiang, EvoMDT: a self-evolving multi-agent system for structured clinical decision-making in multi-cancer, npj Digital Medicine 9 (2026) 124. doi:10.1038/s41746-025-02304-8.
URL https://www.nature.com/articles/s41746-025-02304-8

[36] Y. Yang, J. Li, Q. Pan, J. Zhou, K. Chen, Q. Chen, J. Zhao, N. Zhou, X. Li, L. He, PsychAgent: An Experience-Driven Lifelong Learning Agent for Self-Evolving Psychological Counselor, arXiv v3, 2026-04-28; v1 2026-04-01 (2026). arXiv:2604.00931, doi:10.48550/arXiv.2604.00931.
URL https://arxiv.org/html/2604.00931v3

[37] H. Sun, W. Li, Y. Zhang, Z. Lin, F. Zhang, K. Chen, X. He, Y. Li, M. Liu, L. Liu, Y. Jiang, Experience Makes Skillful: Enabling Generalizable Medical Agent Reasoning via Self-Evolving Skill Memory, Accepted at NeurIPS 2026, accepted at NeurIPS 2026 (2026). doi:10.48550/arXiv.2606.09365.
URL https://openreview.net/forum?id=i09nlt9B6r

[38] Y. Zhu, Z. Wang, Y. Qi, L. Gu, D. Sui, H. Hu, X. Zhang, Z. He, Y. Wang, J. He, L. Ma, L. Yu, HealthFlow: automating electronic health record analysis via a strategically self-evolving multi-agent framework, npj Digital Medicine 9 (2026) 660. doi:10.1038/s41746-026-03097-0.
URL https://www.nature.com/articles/s41746-026-03097-0

[39] J. Wu, J. Zhu, M. Hu, F. Liu, J. Pan, MedRSI: Recursive self-improvement for medical agents via clinically aligned self-evolution (2026). arXiv:2609.24838, doi:10.48550/arXiv.2609.24838.
URL https://arxiv.org/abs/2609.24838v1

[40] Z. Chen, M. Feng, MedForge-RSI: Medical Deepfake Detection via Recursive Self-Improvement, arXiv preprint arXiv:2609.36549 (2026). arXiv:2609.36549.
URL https://arxiv.org/abs/2609.36549v1

[41] Y. Wang, Z. Lin, J. Liu, T. Zhang, N. Wang, Learning from Expert-Corrected Misdiagnoses: A Self-Evolving Consultation Agent for Traditional Chinese Medicine after Failed Treatment (2026). arXiv:2610.40663, doi:10.48550/arXiv.2610.40663.
URL https://arxiv.org/abs/2610.40663v1

[42] G. C. Yuan, X. Zhang, S. E. Kim, L. Luo, P. Rajpurkar, Asclepius: An adaptive harness for long-horizon clinical agents, in: Findings of the Association for Computational Linguistics: EMNLP 2026, Association for Computational Linguistics, 2026, pp. 44611–44629. doi:10.18653/2026.findings-emnlp.2094.
URL https://aclanthology.org/2026.findings-emnlp.2094/

[43] Y. Zhuang, W. Jiang, J.-Y. Zhang, Z. Yang, J. T. Zhou, C. Zhang, Learning to Be a Doctor: Searching for Effective Medical Agent Architectures, in: Proceedings of the 33rd ACM International Conference on Multimedia, 2025, pp. 6996–7005. doi:10.1145/3746027.3755559.

[44] J. Zheng, I. Nassar, T. Vu, X. Zhong, Y. Lin, T. Liu, L. Duong, Y.-F. Li, MedDCR: Learning to design agentic workflows for medical coding, in: M. Liakata, V. P. Moreira, J. Zhang, D. Jurgens (Eds.), Findings of the Association for Computational Linguistics: ACL 2026, Association for Computational Linguistics, San Diego, California, United States, 2026, pp. 12878–12893. doi:10.18653/v1/2026.findings-acl.627.
URL https://aclanthology.org/2026.findings-acl.627/

[45] T. Ossowski, X. Liu, D. Maqbool, V. Dhanuka, S. Zhang, H. Poon, M. Afshar, T. Bradshaw, J. Hu, CodeClinic: Evaluating Automation of Coding Skills for Clinical Reasoning Agents (2026). arXiv:2605.09675, doi:10.48550/arXiv.2605.09675.
URL https://arxiv.org/abs/2605.09675v1

[46] R. Wu, S. Xie, Y. Sun, H. Tong, F. Zhao, MediSkill-Evo: Process-Constrained Self-Evolution for Evidence-Grounded Clinical Interaction (2026). arXiv:2608.23397, doi:10.48550/arXiv.2608.23397.
URL https://arxiv.org/abs/2608.23397v2

[47] S. Li, J. Xu, T. Bao, Y. Liu, Y. Liu, Y. Liu, L. Wang, W. Lei, S. Wang, Y. Xu, Y. Cui, J. Yao, S. Koga, Z. Huang, A co-evolving agentic AI system for medical imaging analysis (2025). arXiv:2509.20279.

[48] L. Fan, P. Dai, Z. Deng, H. Wang, X. Gong, Y. Zheng, Y. Ou, Evolving medical imaging agents via experience-driven self-skill discovery (2026). arXiv:2603.05860.

[49] Q. Zhang, Y. Ren, S. Qin, X. Zhang, L. Zhang, J. Huang, Z. Liu, C. Liu, H. Feng, J. Chen, Y. Ding, W. You, H. Jiang, Y. Pan, Y. Zhou, J. Chen, L. Chen, W. Liu, T. Liu, Z. Zhao, L. Zhang, Toward Vibe Medicine: A Self-Evolving Multi-Agent Framework for Clinical Decision Support, Meta-Radiology 4 (2) (2026) 100223. doi:10.1016/j.metrad.2026.100223.

[50] Z. Du, L. Zheng, R. Hu, Y. Xu, X. Li, Y. Sun, W. Chen, J. Wu, H. Cai, H. Ying, LLMs Can Simulate Standardized Patients via Agent Coevolution, in: Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, 2025, pp. 17278–17306. doi:10.18653/v1/2025.acl-long.846.
URL https://aclanthology.org/2025.acl-long.846/

[51] X. Dang, W. Ruan, X. Yang, J. Wang, X. Hu, J. Li, P.-A. Heng, MedZERO: Self-evolving agents for open-ended medical reasoning through controlled knowledge accumulation, author-reported acceptance at NeurIPS 2026 (2026). arXiv:2610.08327, doi:10.48550/arXiv.2610.08327.
URL https://arxiv.org/abs/2610.08327v1

[52] X. Wang, Z. Chen, C. Zheng, J. Lao, Q. Wang, H. Lin, J. Liu, Z. Lu, H. Jin, J. Wang, J. Gu, L. Wu, ECG-RSI: Recursive self-improvement through targeted electrocardiogram waveform evolution (2026). arXiv:2610.17718, doi:10.48550/arXiv.2610.17718.
URL https://arxiv.org/abs/2610.17718v1

[53] H. H. Lee, Y. Zhou, Y. Yu, S. Gu, Y. Wu, ReVision3D: Attribution-Guided Recursive Self-Improvement for 3D Medical Perception (2026). arXiv:2609.32984, doi:10.48550/arXiv.2609.32984.
URL https://arxiv.org/abs/2609.32984v2

[54] Y. Zhu, K. Zhou, D. Wu, G. Meng, UltrasoundAgents: Hierarchical Multi-Agent Evidence-Chain Reasoning for Breast Ultrasound Diagnosis (2026). arXiv:2603.10852, doi:10.48550/arXiv.2603.10852.
URL https://arxiv.org/abs/2603.10852v1

[55] M. Jeong, Healthcare AI GYM for Medical Agents (2026). arXiv:2605.02943, doi:10.48550/arXiv.2605.02943.
URL https://arxiv.org/abs/2605.02943v1

[56] M. Wu, Y. Yan, Z. Cai, K. Ji, C. Fang, Z. Sheng, X. Wang, R. Wang, H. Zhang, S. Li, B. Wang, H. Zha, Agentifying Patient Dynamics within LLMs through Interacting with Clinical World Model (2026). arXiv:2605.14723, doi:10.48550/arXiv.2605.14723.
URL https://arxiv.org/abs/2605.14723v1

[57] H. Chen, L. Qi, S. Brown, D. Metelski, T. Xia, J. Lee, Q. Wang, K. Riley, F. Wang, W. Yao, Cura 1T: Healthcare Foundation Model via Recursive Self-Improvement (2026). arXiv:2607.15314, doi:10.48550/arXiv.2607.15314.
URL https://arxiv.org/abs/2607.15314v3

[58] J. Liu, Y. He, Y. He, P. Guo, D. Yang, A. Myronenko, C. Zhao, H. Ye, T. Qi, Y. Zhou, D. Xu, Y. Tang, BaT: Towards Self-Evolving Medical Research Agent with Stage Rubrics (2026). arXiv:2608.16211, doi:10.48550/arXiv.2608.16211.
URL https://arxiv.org/abs/2608.16211v1

[59] M. Shao, C. Yu, Y. Wu, F. Wang, L. Shao, Y. Long, J. Wang, H. Yuan, S. Liu, From Given to Gathered Evidence: Agentic Learning for Longitudinal Medical Reasoning (2026). arXiv:2609.39566, doi:10.48550/arXiv.2609.39566.
URL https://arxiv.org/abs/2609.39566v3

[60] Z. Ge, H. Li, Y. Wang, N. Hu, C. J. Zhang, Q. Li, ClinicalAgents: Multi-agent orchestration for clinical decision making with dual-memory, in: Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2, 2026. arXiv:2603.26182, doi:10.1145/3770855.3818931.

[61] P. Qiu, C. Wu, J. Liu, Q. Zheng, Y. Liao, H. Wang, Y. Yue, Q. Fan, S. Zhen, J. Wang, J. Gu, Y. Wang, Y. Zhang, W. Xie, Evolving interactive diagnostic agents in a virtual clinical environment (2025). arXiv:2510.24654.

[62] Y. Feng, J. Wang, L. Zhou, Z. Lei, Y. Li, DoctorAgent-RL: A multi-agent collaborative reinforcement learning system for multi-turn clinical dialogue, in: ICASSP 2026 - 2026 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2026, pp. 16952–16956. arXiv:2505.19630, doi:10.1109/ICASSP55912.2026.11460976.

[63] W. Xu, Z. Liang, H. Anthony, Y. Ibrahim, F. Cohen, G. Yang, K. Kamnitsas, You point, I learn: Online adaptation of interactive segmentation models for handling distribution shifts in medical imaging, in: International Conference on Learning Representations (ICLR), 2026.
URL https://arxiv.org/abs/2503.06717

[64] E. Tzanis, M. E. Klontzas, ReclAIm: A multiagent framework for monitoring and correcting performance decline in medical imaging AI, Radiology: Artificial Intelligence 8 (4) (2026) e250923. doi:10.1148/ryai.250923.
URL https://pubs.rsna.org/doi/10.1148/ryai.250923

[65] X. Hu, M. Liu, Z. Wang, X. Zhang, Y. Huang, Z. Wang, K. Anmahaepong, Y. Li, Y. Chen, H. Yang, Z. Li, Q. Tian, TEMPLAR: agentic induction and evolution of standardized radiology reporting templates from large-scale clinical corpora (2026). arXiv:2610.05247, doi:10.48550/arXiv.2610.05247.
URL https://arxiv.org/abs/2610.05247v1

[66] S. Li, X. Tan, Q. Wang, W. Zhao, C. Zhan, T. Hao, X. Wang, L. Gu, R. Eils, X. Qiu, Human-In-The-Loop Multi-agent Ventilator Decision Support with Contextual Bandit Preference Learning, in: Medical Image Computing and Computer Assisted Intervention – MICCAI 2026, 2026, pp. 290–300. doi:10.1007/978-3-032-38260-3_28.

[67] S. Chen, F. Chen, X. Ding, W. Chen, Z. Wei, ClinVerse: A Realistic Multi-Agent Diagnostic Environment for Training and Evaluating Medical Agents (2026). arXiv:2610.34878, doi:10.48550/arXiv.2610.34878.
URL https://arxiv.org/abs/2610.34878v1

[68] J. Li, T. Zhou, Z. Zhou, X. Wei, H. Song, Z. Wang, Y. Chen, H. Lv, Experience-guided multi-agent interpretable framework for radiology report summarization, Computer Methods and Programs in Biomedicine 273 (2026) 109078. doi:10.1016/j.cmpb.2025.109078.
URL https://www.sciencedirect.com/science/article/abs/pii/S016926072500495X

[69] K. Chen, J. Qi, J. Huo, P. Tian, F. Meng, X. Yang, Y. Gao, A Self-Evolving Framework for Multi-Agent Medical Consultation Based on Large Language Models, in: ICASSP 2025 - IEEE International Conference on Acoustics, Speech and Signal Processing, 2025. doi:10.1109/ICASSP49660.2025.10889517.
URL https://ieeexplore.ieee.org/document/10889517

[70] Y. Dou, Z. Deng, T. Xing, J. Xiao, S. Peng, Autonomous Pharmaceutical Care with Large Language Models, in: 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2024. doi:10.1109/BIBM62325.2024.10822302.
URL https://ieeexplore.ieee.org/document/10822302

[71] H. Wu, Z. Wang, Z. Rao, H. Lin, J. Zhu, Q. Cai, Y. Zhou, A. Lin, H. Wang, H. Xiong, U-EHR: A Self-Evolving EHR Agent with Step-Level Credit Assignment and UCB-Guided Memory Retrieval, in: Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V.2, 2026, pp. 12420–12431. doi:10.1145/3770855.3819038.
URL https://dl.acm.org/doi/pdf/10.1145/3770855.3819038

[72] S. Mao, Y. Yang, A Self-Reflective Multi-Agent Physician Recommendation Framework, in: CNML 2026, 2026. doi:10.1109/CNML68938.2026.11452303.
URL https://ieeexplore.ieee.org/document/11452303

[73] K. Chen, Y. Gao, MDTace: Agentic Context Engineering for Multi-disciplinary Team Medical Consultation, in: Intelligent Computing Theories and Application (ICIC 2026), Vol. 16649, 2027, pp. 81–92, first online 19 July 2026; publisher citation year 2027. doi:10.1007/978-981-92-3403-5_6.
URL https://link.springer.com/chapter/10.1007/978-981-92-3403-5_6

[74] L. Cao, Y. Shen, T. Luo, S. Du, H. Peng, Y. Guo, GuideSkill: Evolving Executable LLM Agent Skills for Guideline-Grounded Clinical Reasoning (2026). arXiv:2607.26160, doi:10.48550/arXiv.2607.26160.
URL https://arxiv.org/html/2607.26160v1

[75] H. Ruan, X. Jiang, Z. Li, W. Jia, X. Xu, X.-F. Shan, S. Chen, X. Ye, A Self-Evolving LLM Agent Framework for Role-Based Norm Compliance in Healthcare, in: Findings of the Association for Computational Linguistics: ACL 2026, 2026, pp. 22562–22577. doi:10.18653/v1/2026.findings-acl.1133.
URL https://aclanthology.org/2026.findings-acl.1133.pdf

[76] Y. Shi, Q. Wang, S. Gao, J. Zhou, J. Lin, Z. Ying, X. Zeng, C. Lin, MedGuard: an LLM-based gatekeeper for detecting clinical risks in Chinese telemedicine consultations, npj Digital Medicine (2026). doi:10.1038/s41746-026-03116-0.
URL https://www.nature.com/articles/s41746-026-03116-0_reference.pdf

[77] J. Jiang, Y. Lei, Y. Xu, Z. Wang, Z. Wang, L. Zong, H. Huang, Z. Li, Z. Huo, Y. Chen, J. Chen, X. Fan, Y. Li, Self-Evolving Dual-Role Deliberation for Medication Recommendation from Clinical Records, in: Findings of the Association for Computational Linguistics: EMNLP 2026, 2026, pp. 20985–21004. doi:10.18653/2026.findings-emnlp.994.
URL https://aclanthology.org/2026.findings-emnlp.994.pdf

[78] Y. Li, J. Lin, Y. Wang, Skill-CDPO: Evolving Agent Tool-Use via Critical Step Preference Optimization, in: Findings of the Association for Computational Linguistics: EMNLP 2026, 2026, pp. 42309–42326. doi:10.18653/2026.findings-emnlp.1985.
URL https://aclanthology.org/2026.findings-emnlp.1985.pdf

[79] S. Feng, P. Su, Self-Evolving Diagnostic Framework based Gated Residual Adapters and OpenClaw-Based Medical Agents, International Journal of Computer Science and Information Technology 8 (4) (2026) 1–11. doi:10.62051/ijcsit.v8n4.01.
URL https://ijocsit.org/pdf/2026/10.62051-ijcsit.v8n4.01.pdf

[80] W. Li, T. Yang, J. Yang, F. Chao, BridgeGap: Belief-Grounded Training for Medical Dialogue Agents, in: Proceedings of the 2026 Conference on Empirical Methods in Natural Language Processing, 2026, pp. 44932–44947. doi:10.18653/2026.emnlp-main.1974.
URL https://aclanthology.org/2026.emnlp-main.1974.pdf

[81] S. Li, J. Hao, X. Liu, Inference-Time Agentic Decision Rules Beat Longer Evolving Search for Multi-Image Medical Reasoning (2026). arXiv:2607.27564, doi:10.48550/arXiv.2607.27564.
URL https://arxiv.org/abs/2607.27564v1

[82] H. Rugo, A. Forsythe, D. Flora, S. Glück, S. Grieve, R. Campden, R. J. Liu, M. Musat, J. M. Rege, P. A. Kaufman, R. L. Coleman, P. Tarantino, J. O'Shaughnessy, L. Schwartzberg, Development and validation of a human-supervised AI-augmented living oncology evidence platform: a breast cancer pilot study, ESMO Real World Data and Digital Oncology 13 (2026) 100770. doi:10.1016/j.esmorw.2026.100770.
URL https://pmc.ncbi.nlm.nih.gov/articles/PMC13563553/

[83] M. Neary, E. Fulton, V. Rogers, J. Wilson, Z. Griffiths, R. Chuttani, P. M. Sacher, Think FAST: a novel framework to evaluate fidelity, accuracy, safety, and tone in conversational AI health coach dialogues, Frontiers in Digital Health 7 (2025) 1460236. doi:10.3389/fdgth.2025.1460236.
URL https://www.frontiersin.org/journals/digital-health/articles/10.3389/fdgth.2025.1460236/full

[84] W. Chen, From API Acknowledgment to Clinical State: Runtime Governance, Controlled Self-Evolution, and Cost in MedAgentBench, SSRN working paper (2026). doi:10.2139/ssrn.7482519.
URL https://papers.ssrn.com/sol3/papers.cfm?abstract_id=7482519

[85] Z. Fan, X. Ren, Y. Wang, R. Chen, Z. Meng, K. Hu, J. Liu, J. Wu, W. Zhai, H. Xu, Z. Liu, MedEnv: Scaling Multimodal Virtual Medical Environments for Long-Horizon Diagnosis, in: Medical Image Computing and Computer Assisted Intervention - MICCAI 2026, Vol. 16886, 2026, pp. 420–429. doi:10.1007/978-3-032-38098-2_40.
URL https://papers.miccai.org/miccai-2026/paper/5084_paper.pdf