# RSI-HA evidence coding guide

Operational rules for healthcare agent studies | Version 0.3.0 | Search cutoff 4 November 2026

**LLM-assisted extraction.** Verification status is documented in REVIEW_STATUS.md. Apply this guide to the examined source version. It operationalizes existing RSI concepts for this review and does not certify clinical readiness.

### 1. Identify the changing object and its role

Record what changes, what stays fixed, the feedback and retained artifact, and when updating, selection, freezing and testing occur. Ask whether the changed object executes the current task or produces later updates. Keep original section/table/page locators.

### 2. Assign the highest level supported by the examined report

| Level | What the evidence shows |
|---|---|
| 0 | Only current-case or session state changes. Static experience retrieval alone does not show newly learned cross-case memory. |
| 1 | Memory, skills, tools, configuration or learner weights persist across different cases, under a fixed update procedure. |
| 2 explicit | A changed component explicitly constructs, selects or evaluates later training data, interactions, trajectories or recipes. The procedure admitting its changes stays fixed. |
| 2 implicit | The changing component has a documented update-producing or update-assessing role as well as a task role, but its update-related contribution is not isolated. |
| 3 | The update-proposing, selection or evaluation procedure itself is revised, and its effect on subsequent learning is tested. |

Merely executing tasks whose outputs are later mined is insufficient for Level 2. A Level 2 assignment does not itself establish causal benefit. A Level 3 claim requires testing later learning against the fixed procedure, with independent evaluation and data/compute matching reported.

### 3. Keep mechanism, benefit and clinical evaluation separate

Record publication status; controls for the changed component; and evaluation in simulation, offline benchmarks, retrospective records, external retrospective cohorts or prospective care. Code the study evaluation design separately from whether a patient service is operating. "Not reported" is a reporting gap, not proof of absence.

### 4. Check the boundary examples as coded in this review

**ClinicalAgents | L0** Static experience memory and changing within-case search do not establish new cross-case learning. Source.

**Agent Hospital | L1** Cases and principles accumulate while the reflection and validation procedure stays fixed. Source.

**ECG-RSI | L2 explicit** Changed synthesis operators and skills create later training data. Frozen-component controls exist, with unequal generation-search costs. Source.

**SEMPA | L2 implicit** The changing Critic affects later update selection; its evolving evaluator contribution is not isolated. Source.

No mapped healthcare study reaches Level 3 or reports prospective clinical evaluation. These are the review's source-version-specific judgments; consult the full extraction and its limitations.

### Use the source-linked extraction and preserve attribution

Start from the blank extraction template and field codebook. The four examples and 85-study map retain original source locators. Levels, causal benefit, clinical setting and publication status are separate fields; no aggregate quality score is assigned.

Conceptual sources: Liu et al. (2026), component versus improver; Han et al. (2026), L0-L3 roles; Pu et al. (2026), functional distinction. Cite those sources for their concepts and the RSI-HA manuscript for this healthcare map and operationalization.

Use the companion evaluation-design templates for controls, phase separation and experiment records.

dengzhe-hou.github.io/RSI-Healthcare-Agents