AI in Clinical Skills and Assessment

Active Learning

Macy Foundation / Millennium Conference • Preclinical M1/M2 • November 2026

Format. Structured reflective writing in pairs. 25 minutes total: 3 minutes orientation, 8 minutes individual writing, 6 minutes paired exchange, 8 minutes facilitated debrief. Reflective writing is the recommended format for this session because the objectives include identity-formation work — structured self-reflection on how AI will and will not be used — that benefits from committed, first-person writing.

Tier A and Tier B. This activity requires no technology and runs in any classroom; the two implementation tiers converge. If the session's PBL cases were run first, students may draw on them, but the activity is written to stand alone.

Student-facing instructions

Before you write: an honest premise

You are being asked to set rules for an environment you have not yet worked in. As a preclinical student, you have probably never watched an ambient scribe draft a note under your name, or had an EHR risk score disagree with your assessment. That is not a flaw in this exercise — it is the reason for it. You are writing this boundary statement before the habits form, not after — the way clinicians decide how they will handle a code before the code, not during it. Expect your statement to be provisional; you will be invited to revise it against experience before clerkships begin. What it must not be is vague.

Step 1 — Orient (3 minutes, as a class)

The AI tools you will meet in clinical training fall into families that deserve different rules. Two matter most for this exercise:

	Clinical decision support (CDS)	Documentation aids (incl. ambient scribes)
What it does	Suggests differentials, risk scores, workups from clinical data — it overlaps with your reasoning.	Drafts notes, summaries, and messages from the encounter — it overlaps with your record of what happened.
What's at stake for a learner	The development of your own diagnostic reasoning (never-skilling risk lives here).	The accuracy of what you sign, and your summarization-as-understanding skill.
The core question your boundary must answer	At what point in my reasoning do I consult it — and what must I have done first?	What do I verify before anything drafted for me goes out under my name?

Definitions of every term used here (CDS, ambient scribe, documentation aid, skill-tracking platform, deskilling, mis-skilling, never-skilling) are in the glossary at the end of this handout.

Step 2 — Write your boundary statement (8 minutes, individually)

Write a personal AI boundary statement: one to two paragraphs, in your own voice, describing how you intend to use AI tools during clinical training. It must be built from concrete, observable behaviors — things a classmate shadowing you could watch you do or not do — rather than intentions (“I will be careful”) or values (“I believe in patient-centered care”). Address all five prompts, in any order:

- **A — Decision support.** What will I always do myself, before consulting a CDS tool or asking an AI for a differential? Name the specific acts (e.g., “write my own three-bin differential and one do-not-miss diagnosis”).
- **B — Documentation.** When an AI drafts a note or summary under my name, what specific checks will I perform before I sign or present it? Name the steps.
- C — Same or different? Compare your answers to A and B. Is your approach to reasoning tools and documentation tools the same or different — and why? (Hint: one guards how you think; the other guards what you attest. Many strong statements treat them differently.)
- **D — My risk.** Which risk — deskilling, mis-skilling, or never-skilling — most applies to someone at my stage, in which specific skill? What is the concrete habit that protects it, and the situation where I will be most tempted to drop that habit?
- **E — The check.** How will I know, three months into clerkships, whether I am keeping this boundary? Name an observable sign — something that would show up in my behavior, not my intentions.

Step 3 — Exchange (6 minutes, in pairs)

Read your statement to a partner. Each partner returns exactly two pieces of feedback:

- The single most actionable commitment in the statement — the one a shadowing classmate could verify tomorrow.
- The single vaguest sentence — and a suggested rewrite that turns it into a behavior.

Step 4 — Debrief (8 minutes, as a class)

The facilitator will lead the closing discussion. Keep your statement; this curriculum invites you to revisit and revise it at the end of the M2 year, before clerkships begin.

FACILITATOR GUIDE — REFLECTIVE WRITING

Educational purpose and cognitive level. This is a Bloom's create-level activity: students are not judging an existing policy but authoring an original, personal artifact — a set of concrete behavioral commitments synthesized from the session's concepts and their own self-knowledge. Facilitate accordingly: the assessment target is the specificity and internal coherence of what they build, not whether their rules match an official answer.

Why the prospective framing matters. These are preclinical learners writing about tools most have never used. Left unacknowledged, that gap produces either generic aspirations or borrowed opinions. The activity therefore (a) names the gap explicitly in the student instructions, (b) anchors writing in the CDS-versus-documentation comparison so students have a concrete framework rather than a blank page, and (c) frames the statement as a draft to be revised against experience. If a student objects “how would I know?”, affirm it — then point out that the commitment to a verification behavior (prompt B) requires no clinical experience at all.

Setup. Paper or a blank document; no other technology. Run the orientation table briskly — it is scaffolding, not a second lecture. Time each phase visibly.

Anticipated student responses. Most first drafts are too vague; that is what Step 3 exists to fix. Expect honest anxieties: fear of being slower than AI-using peers, uncertainty about what is allowed at their site, doubt the boundary will survive fatigue. Treat these as material for prompt D, not problems to reassure away. On prompt C, expect a productive split — some students will hold one uniform rule, others will reason their way to stricter behavior around what they sign than around what they consult. Both are defensible if argued from the think/attest distinction.

Watch for. (1) Policy language (“students should…”) — redirect to first person, present tense. (2) Abstinence statements (“I will not use AI until residency”) — probe feasibility and cost: these tools will be in their environment regardless, and refusing to develop verification skills is its own risk. The curriculum's

through-line is stewardship, not abstinence. (3) Statements listing only reasoning boundaries and skipping documentation (or vice versa) — send them back to the weaker column. (4) Signs of genuine distress about AI and their future — note the offer of confidential follow-up.

Closing debrief prompts (choose two or three). “Read us one behavior from your statement that a shadowing classmate could verify.” • “Who treated decision support and documentation differently — and what's the argument?” • “What would have to be true about your clerkship environment for this commitment to be easy rather than heroic?” • “What will you look for, three months in, as your early-warning sign?”

Longitudinal follow-up (recommended). Collect statements (or have students self-archive them) and return them at the end of M2 for revision against actual clinical exposure. This converts a 25-minute activity into a longitudinal reflective practice and supplies a natural artifact for the session's evaluation plan.

Glossary — terms used in this activity

CLINICAL DECISION SUPPORT (CDS)

Software that takes clinical information and returns suggestions — differential diagnoses, risk scores, alerts, recommended tests. Some is rule-based; newer tools use machine learning or large language models. Its output is an input to your judgment, never a verdict.

AMBIENT SCRIBE (AMBIENT LISTENING SYSTEM)

A microphone-plus-AI system that records a clinical visit, transcribes it, and drafts the note for the clinician to review and sign. It reduces documentation burden, and it can also omit, alter, or invent content — which is why the review is the clinician's job.

AI-POWERED DOCUMENTATION AIDS

Drafting tools inside the EHR: autocomplete, AI-drafted replies to patient messages, chart summarizers. They work from text already in the record, so they can propagate existing errors as well as introduce new ones.

AUTOMATED SKILL-TRACKING PLATFORMS

Systems that log learner performance data (procedure counts, checklist scores, sometimes video) and flag patterns. Worth asking of any such tool: what is captured, who sees it, what decisions does it feed?

DESKILLING • MIS-SKILLING • NEVER-SKILLING

Deskilling: losing a skill you once had, through disuse, after a tool takes over the task.

Mis-skilling: learning a skill incorrectly because a tool shaped how you practiced it.

Never-skilling: never developing a foundational skill at all, because a tool absorbed the cognitive work before you ever did it yourself. This is the version specific to learners at your stage — and the hardest to notice, because you cannot miss what you never had. It is why prompt A asks what you will always do first.