T e a c h i n g

A

ACT 1 — THE SETUP

W e '

Our learners are getting an AI curriculum. We just aren't the ones teaching it.

M e d i c a l

8

of US medical students, per a 2026 JMIR national survey

W i t h

3

report any formal AI integration in their training — a gap we cannot ignore

OpenEvidence reports **over 40% of US physicians log in daily** — roughly 20 million clinical consultations per month by early 2026. This is already the dominant reality of training. We are not getting ahead of it.

THE CENTRAL ASYMMETRY

S a m e

> "It is one thing for me to use it to ask a question. I know the questions to ask, and I can interpret an answer and know if it is a good answer or a dumb answer. A first- or second-year medical student is not an expert. They are a novice."
> — Bob Wachter, MD

T h e

Uses AI like a **curbside consult** — calls in a question, weighs the answer against deep prior knowledge, accepts or rejects it..

T h e

Uses AI as a **crutch** — cannot interrogate the output because they haven't yet built the schema to evaluate it.

NAMING THE PROBLEM

Three

Before we solve the problem, we need to name it correctly. These come from Abdulnour et al., *NEJM* August 2025 — we'll return to this paper in a few minutes.

Deskilling

Losing a skill you once had. The experienced endoscopist who stops reading films carefully because AI flags the findings first.

Never-Skilling

Never developing the skill in the first place. The resident who never struggled through a differential — because the struggle was always bypassed.

Misskilling / Bias

Learning the AI's errors as truth. The resident who internalizes a hallucinated mechanism because no one taught them to verify.

Education

Academic

24 hospitalists. Epic integrated with ambient scribes. Chief resident tracking AI tool use. Structured feedback cycles. GME oversight infrastructure.

Critical Access

You, a single APP, a free ChatGPT tab, and overnight call. No IT governance. No curriculum committee. Just the clinical moment in front of you.

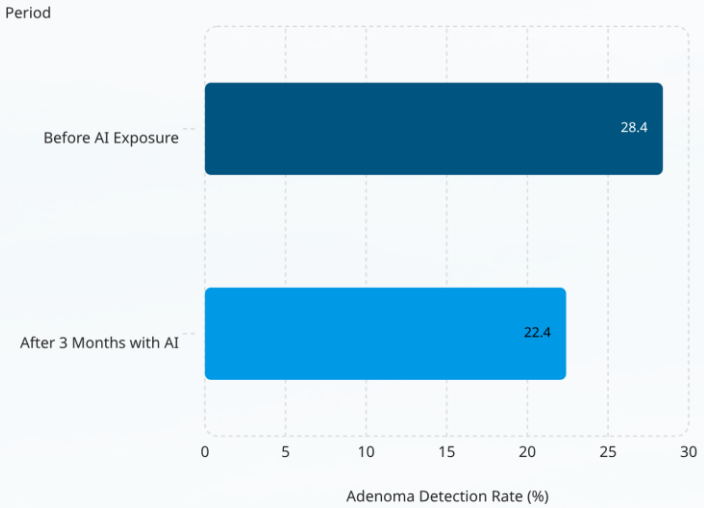
Both contexts are real. Both are yours. The **frameworks apply in both** — the institutional tools may not. The norm-and-question moves we'll land on cost nothing and require no IT ticket.

ACT 2 — THE EVIDENCE

T h e

Budzyń et al., *Lancet Gastroenterology & Hepatology*, August 2025. 1,443 colonoscopies. Four Polish centers. Endoscopists averaging 30 years of experience and over 2,000 prior procedures.

W h a t -A I



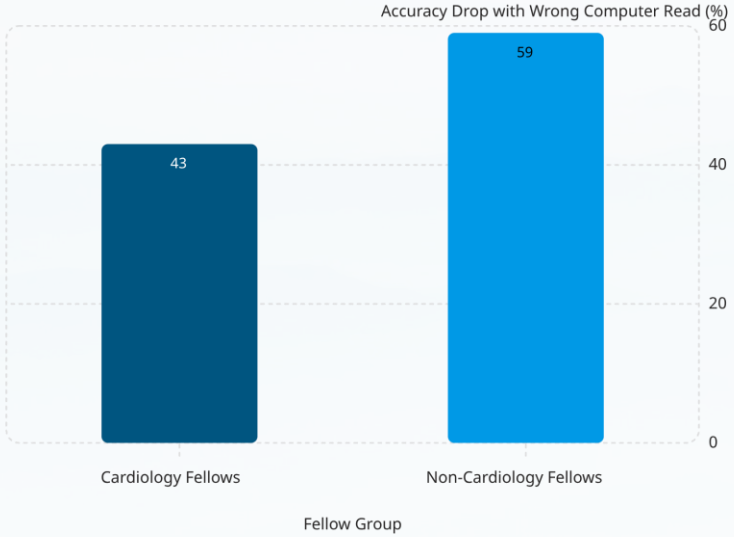
W h y

A **20% relative drop** in detection — in experts — after just three months of AI-assisted work.
As Rodman put it: *"If this is the level of deskilling that happens when somebody trained in the old way uses it for three months — what happens when somebody trains with this from the very beginning? Do they ever develop those skills?"*

THE EVIDENCE — TRAINEES

T h e

Bond et al., *Journal of Electrocardiology*, 2018. 30 cardiology and non-cardiology fellows. 9,000 ECG interpretations. One variable: whether the computer read was correct or incorrect.



A u t o m a t i o n

When the computer was wrong, fellows followed it — and the less experienced they were, the steeper the fall.

> "Automation bias could bring about the automation paradox — where humans lose their skills, vigilance, and a reliable sense of certainty in decision-making."
> **— Bond et al.**

W h e n

> "Let me start with a few confessions: I'm no longer very good at reading a map, doing long division, or listening to heart sounds for the telltale murmurs of aortic stenosis. And I periodically forget my wife's phone number — if I ever lose my iPhone, I fear I might never see her again. In short, I've been deskilled."
> **— Bob Wachter, MD, *Pattern Recognition* Substack, August 2025**

WHAT'S ACTUALLY SHIFTING

T h e — I t

The pivot in one line: **from "knowing" → "knowing how to verify, reason, and act."**

R e c a l l

GPT-4 passes the USMLE. It cannot sit in a room with a frightened family at 2am and make a goals-of-care decision.

D o c u m e n t a t i o n f u n c t i o n

Ambient scribes have removed the cognitive pressure that writing a SOAP note once created. But how does that replace structured thinking?

T h e

From teaching facts → supervising judgment. The expert-as-curbside model applies to precepting as much as it does to AI use.

T h e

When you don't model AI use explicitly, residents do it badly — and privately. Your silence is a curriculum.

DEFT-AI: -Minute

Abdulnour et al., *NEJM*, August 2025. A *supervisory* framework — not a curriculum redesign. You can use this tomorrow on rounds with zero IT involvement.

D — Diagnosis

Probe the learner's own reasoning *and* the prompts/questions they used. Separate their thinking from the machine's output.

E — Evidence

Ask for supporting *and* opposing data. Did they read the sources the AI cited? Did they do any independent research?

F — Feedback

Prompt self-reflection. Where did their reasoning diverge from the AI? Which divergence was right?

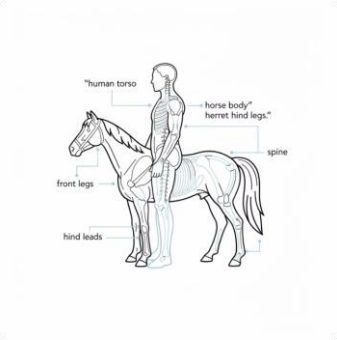
T — Teaching

Reinforce clinical reasoning *and* AI literacy simultaneously. Name the skill explicitly.

AI — AI

Guide safe, adaptive AI use. Every output is an untraceable judgment until the learner can defend it independently.

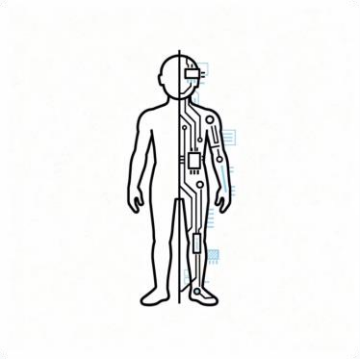
Centaur — Two -AI



Centaur

Distinct collaborators. Clear handoffs.

- Human reasons first; AI assists with a specific sub-task
- Boundary between human and AI work stays visible
- Example: resident builds differential → consults OpenEvidence → integrates selectively



Cyborg

Fused in the same task. Real-time.

- AI is woven into the cognitive flow as the work happens
- Boundary between human and AI work disappears
- Example: ambient scribe drafting the note while the interview unfolds

For trainees, centaur protects the struggle that builds judgment. Cyborg is faster — and carries the higher deskilling risk.

DEFT-AI IN PRACTICE

W h a t

"

t o

Surfaces whether the learner is thinking, then asking — or asking, then presenting. The expert uses AI like a curbside. The novice lets it do the thinking.

"

t o l d

Probes whether they have an independent basis for the claim. If they can't answer, that's a teaching moment, not a failure.

"

t h i n k–ianngd

The delta between their reasoning and the AI output is exactly where learning happens. Make it explicit. Teach the framework.

ASSESSMENT

Assessment

If your assessments don't change, your teaching won't either.

What'

- Written SOAP notes and case summaries — AI-undetectable at scale
- Multiple-choice shelf exams for knowledge recall alone
- Take-home written assignments without oral defense

What' —and

- **Oral exams** and bedside presentations without notes
- **Observed reasoning** — entrustment decisions with explicit AI-aware language
- **Observed clinical encounters** — direct observation of history-taking and physical exam, scored independently of any written output

ACT 4 — THE CLOSE

W h a t

The literature points to two domains with the strongest claim on being irreducibly human.

B e d s i d e

Physical exam. History-taking under genuine listening — not transcription. The Verghese / Stanford 25 argument: the ritual of the exam is itself therapeutic, and it cannot be outsourced.

M a n a g e m e n t

Rodman & Parsons, *JGIM* 2026: *"Diagnosis answers 'what's wrong?'; management answers 'what should we do?' Medicine has long privileged diagnostic virtuosity. This imbalance leaves us vulnerable as AI masters pattern recognition."*

For hospitalists specifically: **discharge planning, goals-of-care conversations, transitions of care** — the irreducibly human work that AI cannot do and families will not accept from a machine.

Your

Three things. No more. These replace existing habits — they don't pile on top of them.

State

Commit-then-consult is the rule on your service. Say it out loud, publicly, on day one of the rotation. It costs nothing. It changes everything.

Add Rounds

The DEFT-AI bedside trio: **What prompt? How verified? What changed?** Three minutes. Every presentation. Return to these until they are reflexive.

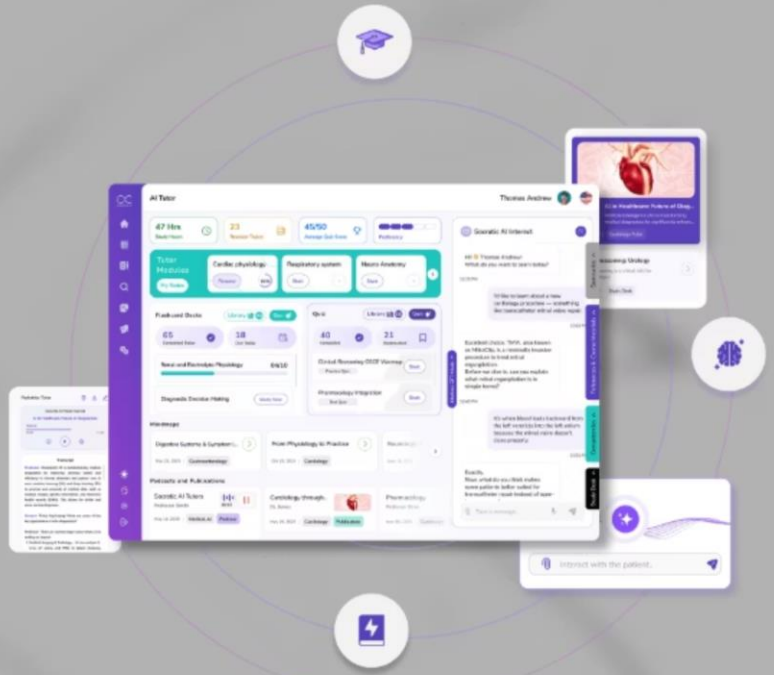
Swap

Replace one written case write-up per block with an **oral case presentation** — no notes, no AI assist. Observed reasoning is the new gold standard.

What you do **not** have to do: rewrite your curriculum, get IT involved, learn how transformers work, or ban anything.

Reimagine

Your Clinical Journey



Socratic

THE SOCRATIC METHOD, RE-DISCOVERED

The
the

Medical education has used Socratic questioning for 2,400 years — the chief on rounds asking 'why?' until the resident either earns the answer or earns the gap. Most AI tools removed the 'why?' They just give the answer. A Socratic AI puts it back.

Socrates

Ask, don't answer. Force the student to reason aloud.

Socratic

The same design pattern — scaled, available at 2am, infinitely patient.

Chief

Why? What else? What would change your mind?

Generic

Cyborg by default

1. Learner types question
2. AI returns answer
3. Struggle gone
4. Reasoning: optional

Socratic

Centaur by design

1. Learner commits to a thought
2. Tool probes with a question
3. Learner refines
4. Repeat — reasoning is the product

A

The tool probes. The learner does the work.

Learner

Patient has chest pain. I think it's an MI.

Socratic

What features of this presentation made MI your leading diagnosis? What two findings would lower your suspicion?

Learner

(types out reasoning, names features)

Socratic

Good. Your differential has three diagnoses. Which one would kill the patient fastest if you missed it?

"What does it mean to be a doctor in an era when in your pocket is a tool that's smarter than you are?"