SKILL-PRESERVING UX: DESIGNING AGAINST AI DESKILLING IN CLINICAL PRACTICE

***Hrozian Y.**, Founding Designer, Lumos AI, San Francisco, California, USA*

Abstract. *As AI assistance scales in clinical practice, evidence is mounting that continuous reliance erodes the very expertise it augments. Recent clinical trials show measurable decline in unaided performance among clinicians who work continuously with AI assistance. This paper introduces Skill-Preserving UX, a four-pattern framework (Stretch Tasks, Calibration Drills, Solo Mode, Shared-Authority Mode) for designing AI interfaces that protect rather than erode expert capability over time.*

Keywords: *Skill-Preserving UX, deskilling, clinical AI, healthcare AI, human-AI interaction, expertise preservation.*

The dominant story about AI in clinical practice is augmentation: AI enhances clinician performance, reduces workload, and improves outcomes. Recent evidence complicates that story [1]. A 2025 study in *Lancet Gastroenterology & Hepatology* found a 6 percentage-point drop in unaided adenoma detection among experienced endoscopists working continuously with AI assistance: from 28.4% before AI exposure to 22.4% after, observed when AI was unavailable [2]. This phenomenon, deskilling, has been documented for decades in aviation and industrial automation [3], but is only now arriving in clinical AI at scale. Most healthcare AI design has focused on automation efficiency. This paper argues deskilling is a design problem, and proposes Skill-Preserving UX, a four-pattern framework (Table 1).

Table 1 – Skill-Preserving UX Patterns

Pattern	What it does	Example implementation
Stretch Tasks	Periodically suppresses AI assistance on sampled cases to maintain unaided performance.	Random 5-10% of cases run AI-blind end-to-end; AI signal absent during the case.
Calibration Drills	Active practice cases with known ground truth, designed to recalibrate clinician judgment.	Quarterly drill modules with feedback on agreement with expert consensus.
Solo Mode	An interface mode where AI is fully suppressed, supporting unaided practice.	One-click toggle into solo mode; AI signal returns at session end as optional retrospective.
Shared-Authority Mode	AI presents reasoning only after the clinician commits to their own judgment.	Clinician answers first; AI verdict appears next; surfaced disagreement becomes learning signal.

Stretch Tasks suppress AI signal end-to-end on a random 5-10% of cases. The clinician handles those cases without AI suggestions, confidence scores, or nudges, as if AI did not exist for the case. Aviation training has used this approach for decades through “raw data” flying and unannounced manual segments [4]. Optional retrospective comparison can surface the AI’s verdict after the session ends, but never during the decision. The fraction matters: too low fails to maintain unaided skill; too high erodes AI’s productivity benefit.

Calibration Drills extend Stretch Tasks into structured practice with feedback. The system presents curated cases with known ground truth; the clinician records a judgment and receives feedback on agreement with expert consensus. Where Stretch Tasks sample randomly, drills target specific calibration questions, catching drift that random sampling alone might miss. Building and maintaining the case set requires curriculum investment, but drills can be amortized across many clinicians.

Solo Mode is a fully user-controlled deskilling defense. The clinician opts out of AI for a defined session and the system respects that choice end-to-end. Few production clinical AI products offer this, and the reason is product priority, not technical complexity. Where "off" toggles do exist, they typically suppress explicit AI output but leave its signal woven through pre-populated fields and ambient cues.

Shared-Authority Mode reframes the interaction order. The clinician commits to a judgment before seeing the AI's output; only after the human verdict is recorded does the AI's verdict appear. This forces unaided reasoning on every case while still benefiting from AI as a structured second opinion. Every case takes slightly longer, but the clinician keeps unaided judgment sharp, and every disagreement becomes a learning signal. Shared-Authority Mode is the most aggressive deskilling defense because it operates on every case, but it carries the highest workflow tax.

These patterns are not free. Stretch Tasks lose AI's accuracy boost on the cases run blind; Calibration Drills require curriculum and time; Solo Mode requires real product investment; Shared-Authority Mode adds cognitive overhead. These costs pay off over time: a clinician who can perform unaided can detect when the AI is wrong, override appropriately [5], and adapt when AI is unavailable. The FDA's 2025 guidance [6] requires ongoing performance monitoring but not yet skill preservation; designers can move ahead of regulation. For UX teams, the takeaway is that "make the user faster" is not the only goal. Sometimes, slow and unassisted is the right design choice.

References

1. Kaluhin, N., Vovk, O., & Chebotarova, I. (2024). The impact of artificial intelligence on future of humanity. Jóvenes en la ciencia, (26). https://www.jovenesenlaciencia.ugto.mx/index.php/jovenesenlaciencia/article/view/4235/3716.

2. Budzyń, K., Romańczyk, M., Kitala, D., et al. (2025). Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: A multicentre, observational study. The Lancet Gastroenterology & Hepatology, 10(10), 896–903. doi.org/10.1016/S2468-1253(25)00133-5.

3. Bainbridge, L. (1983). Ironies of automation. Automatica, 19(6), 775-779. doi.org/10.1016/0005-1098(83)90046-8.

4. Casner, S. M., Geven, R. W., Recker, M. P., & Schooler, J. W. (2014). The retention of manual flying skills in the automated cockpit. Human Factors, 56(8), 1506–1516. doi.org/10.1177/0018720814535628.

5. Amershi, S., Weld, D., Vorvoreanu, M., et al. (2019). Guidelines for Human-AI Interaction. 2019 CHI Conference on Human Factors in Computing Systems/ (p. 1-13). ACM. doi.org/10.1145/3290605.3300233.

6. U.S. Food and Drug Administration. (2025). Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations (Draft Guidance). fda.gov/regulatory-information/search-fda-guidance-documents.