Digital medicine

Preserving clinical skills in the age of AI assistance

How are clinicians to preserve core clinical skills in an era of algorithmic assistance? As artificial intelligence (AI) assumes a growing role in clinical practice, concern is mounting that off-loading clinical tasks and reasoning will lead to loss of skills (deskilling), adopting errors or bias from AI (mis-skilling), or failure to achieve competence (never-skilling; figure). Evidence for such skill attrition has been seen with automated interpretation of electrocardiograms or radiological images. An observational study published earlier this year, however, sharpens this concern, suggesting that experienced colonoscopists lost some proficiency in colon polyp detection when routine AI support was switched off.

At four Polish endoscopy centres, 19 gastroenterologists and surgeons who used an AI polyp detection system for 3 months showed a decline in adenoma detection rates when subsequently performing colonoscopies without AI. As a group, their post-AI polyp detection performance fell below their own baseline pre-AI performance. Some endoscopists experienced dramatic drops, whereas others maintained their performance. This finding suggests deskilling is not inevitable, and that we must look for approaches to mitigate physician deskilling as AI becomes woven into routine care.

We can learn from other industries that have long grappled with balancing automation and skill retention. In aviation, autopilot systems have brought safety gains, yet also raised concerns about erosion of manual flying skills. Regulators responded by requiring routine hands-on flight time and simulator sessions that mimic system failure. Nuclear plant operators rehearse critical scenarios at regular intervals to ensure that if automation fails, core skills remain intact.

The potential for AI-related deskilling extends well beyond gastrointestinal endoscopy. Surgical and interventional specialties require constant vigilance and readiness to respond to unexpected events. Dermatology, pathology, and radiology demand meticulous visual analysis and pattern recognition. Beyond procedures or image interpretation, clinical reasoning for diagnosis, decision making, and patient management requires critical thinking that could be eroded by AI interactions. Patient encounters are often hurried, leaving little space for deeper analysis and reasoning. Small wonder, then, that physicians would welcome the promise of AI to share the load. But when AI becomes entrenched in clinical practice, new risks emerge. Vigilance may diminish, skills could decay, and the more physicians rely on AI, the less confident we may become in our own, unaided judgement.

These concerns are especially relevant during medical education, where doctors-in-training may depend on AI before fully developing core competencies (never-skilling), leaving little foundation to draw on if technology fails, and curtailing the formative experiences that allow clinical intuition and judgement to fully develop. Protecting physician skills and critical thinking in an AI-enabled environment will not have a single, simple solution. The trajectory of deskilling, over-reliance, or loss of vigilance will differ across algorithms, specialties, and contexts, as will potential solutions. One safeguard might be to build short, deliberate "AI-off" or "AI-delay" intervals into clinical workflows, to recalibrate vigilance and capture objective measures of unaided performance. During training or skill-maintenance sessions, AI interpretations could be set to appear only after the reader logs their own findings. Another approach is to draw firmer boundaries: AI takes on rule-based, high-volume, low-ambiguity work, freeing physicians to focus more on contextual, ambiguous, and high-stakes decisions. For instance, in radiology, an AI algorithm might safely exclude normal chest x-rays so that radiologists can concentrate on subtle or complex findings. Such boundaries would help reduce automation bias (over-trusting incorrect AI outputs) and automation neglect (ignoring correct ones), which can arise when tasks are supposedly shared. Whatever the strategy, the foundation must be clinicians' ongoing ability to evaluate the trustworthiness of any AI tool.

The colonoscopy AI deskilling study is not an indictment of AI. Evidence from over 40 randomised trials for AI polyp detection shows it helps physicians detect more precancerous colon polyps. We should welcome such progress, but remain vigilant to how physicians' behaviour and clinical care will change as AI becomes embedded in practice. AI adoption is accelerating, yet workflows and habits are open to shaping. The choices we make now about how we design, integrate, and train around AI will determine whether these systems elevate our profession or quietly erode the skills that define it.

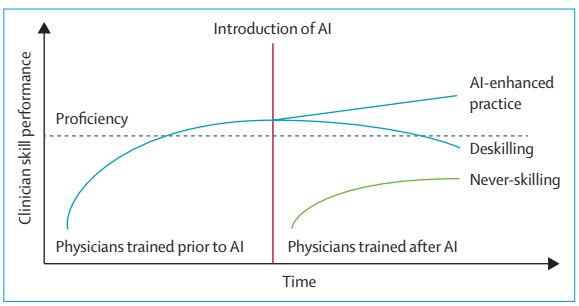

Figure: **Potential effects of AI on clinical skills over time**

Tyler M Berzin, Eric J Topol
Center for Advanced Endoscopy, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, MA 02215, USA (TMB); Scripps Research Translational Institute, La Jolla, CA, USA (EJT)
tberzin@bidmc.harvard.edu

Further reading

Abdulnour E, Gin B, Boscardin C, et al. Educational strategies for clinical supervision of artificial intelligence use. *N Engl J Med* 2025; **393:** 786–97

Budzyń K, Romańczyk M, Kitala D, et al. Endoscopist deskilling risk after exposure to artificial intelligence in colonoscopy: a multicentre, observational study. *Lancet Gastroenterol Hepatol* 2025; **10:** 896–903

Gaube S, Suresh H, Raue M, et al. Do as AI say: susceptibility in deployment of clinical decision-aids. *NPJ Digit Med* 2021; **4:** 31

Kostick-Quenet KM, Gerke S. AI in the hands of imperfect users. *NPJ Digit Med* 2022; **5:** 197

TMB reports consulting fees from Boston Scientific, Medtronic, Sanofi, and Beckman Coulter. EJT is supported by the NIH/National Center for Advancing Translational Sciences grant UL1TR001114 and is an adviser to Tempus Labs, Pheno AI, and Abridge.