

Viewpoint

# Enhancing Physician Resilience to Generative AI: Multilevel Framework for Shared Authority, Verification, and Skill Preservation

Hongxia Pan<sup>1</sup>, MMed; Jialin Liu<sup>2,3</sup>, MD; Siru Liu<sup>4</sup>, PhD

<sup>1</sup>Rehabilitation Medicine Center and Institute of Rehabilitation Medicine, West China Hospital of Sichuan University, Chengdu, Sichuan, China

<sup>2</sup>Department of Medical Informatics, West China Hospital of Sichuan University, Chengdu, Sichuan, China

<sup>3</sup>Department of Otolaryngology-Head and Neck Surgery, West China Hospital of Sichuan University, Chengdu, Sichuan, China

<sup>4</sup>Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, United States

**Corresponding Author:**

Jialin Liu, MD

Department of Medical Informatics

West China Hospital of Sichuan University

Guoxue Xiang Street

Chengdu, Sichuan 610041

China

Phone: 86 28-85422416

Email: [djlj8@163.com](mailto:djlj8@163.com)

## Abstract

As generative artificial intelligence (AI), particularly large language model-based tools, is increasingly integrated into diagnosis, triage, decision support, and treatment planning, it offers potential gains in efficiency and information access. However, real-world deployment also introduces important risks, including hallucinations, miscalibrated confidence, automation bias, and increased verification burden on physicians. This burden may divert attention from independent clinical reasoning, contribute to deskilling, and increase vulnerability when models fail silently or perform poorly in unfamiliar clinical contexts. Existing AI governance frameworks emphasize data quality, transparency, accountability, and ethical deployment, but pay less attention to physician-facing resilience, defined in this paper as the capacity to sustain independent and safe clinical judgment when collaborating with generative AI. In this viewpoint, we propose a multilevel governance framework organized around 3 coordinated domains: cognitive workload shaping, clinical authority governance and allocation, and organizational safety governance and accountability. Together, these domains aim to reduce verification burden, preserve physician decisional authority, and align institutional oversight with safe and context-sensitive AI use. The framework includes mechanisms such as risk-sensitive verification triggers, bounded delegation, structured interprofessional review, and organizational monitoring to support safe clinical integration while minimizing avoidable workflow disruption. At the same time, implementation may be limited by workflow friction, alert fatigue, variable institutional resources, and the need for ongoing monitoring and recalibration to ensure that safeguards remain clinically useful rather than burdensome. Accordingly, this paper outlines a structured governance framework to guide safer integration of generative AI into clinical care and inform future evaluation across specialties, workflows, and institutional settings.

*J Med Internet Res* 2026;28:e88058; doi: [10.2196/88058](https://doi.org/10.2196/88058)

**Keywords:** cognitive load; authority; accountability; artificial intelligence; AI; generative artificial intelligence; generative AI; clinical safety; physician

## Introduction

Large language model-based generative artificial intelligence (AI) is rapidly extending from documentation support to core clinical activities, including diagnosis, clinical decision support (CDS), triage, treatment recommendation,

and preventive care planning [1-6]. Early evaluations suggest potential gains in timeliness and efficiency; however, deployment in real-world clinical settings introduces sociotechnical risks with direct implications for safety and professional competence [7,8]. Generative AI can produce fluent but incorrect outputs (hallucinations); express

confidence that is not supported by the underlying evidence; and perform inconsistently under distribution shift, prompt artifacts, or algorithmic bias, potentially leading to clinically unjustified differences across demographic or socially marginalized groups [9]. In this paper, “distribution shift” refers to reduced performance when the AI encounters patients, settings, or data patterns that differ from its training data, and “prompt artifacts” refers to misleading changes in output caused by superficial wording or formatting differences rather than clinically meaningful distinctions. These risks impose a substantial verification burden: physicians must divert time and attention from intrinsic clinical reasoning to source checking, uncertainty appraisal, and contextual fit assessment, thereby increasing extraneous cognitive load [10-12]. In addition, assigning complex integrative tasks to AI may contribute to deskilling as independent diagnostic synthesis and risk communication decline, leaving physicians more vulnerable during model failures or out-of-distribution conditions [13,14]. Coupled with automation bias and overtrust, these pressures create structural vulnerabilities that cannot be mitigated solely through traditional safety training or single-discipline governance [15].

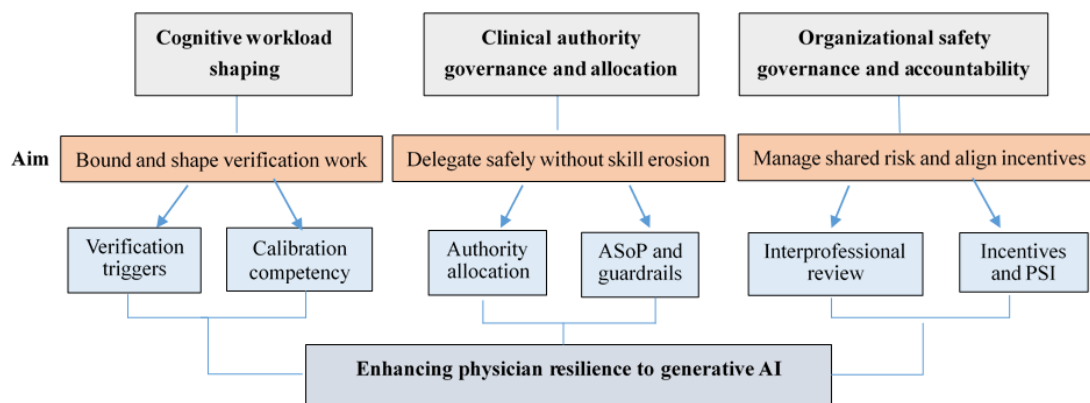
We use “resilience” in the sense established in resilience engineering: the capacity of a system or its actors to sustain safe and effective functioning under variable, degraded, or unanticipated conditions [16,17]. This usage differs from narrower psychological accounts that emphasize individual emotional coping or recovery from adversity. Applied to the physician-AI interface, physician resilience refers to the sustained capacity for independent clinical judgment when collaborating with systems whose outputs may be uncertain, miscalibrated, or context inappropriate. Although AI governance research has extensively examined data quality, transparency, and ethical principles, less attention has been paid to the combined cognitive, professional, and organizational conditions needed to preserve this resilience in physician-AI collaboration.

Sustaining this resilience depends on preserving physician autonomy. Autonomy in this case does not mean isolated clinical discretion or resistance to AI assistance. Rather, it

refers to the physician’s professionally accountable capacity to interrogate, contextualize, accept, modify, or override AI-generated suggestions based on clinical evidence, patient values, uncertainty, and situational constraints [18,19]. The relationship between resilience and autonomy is reciprocal: resilience requires preserved autonomy so that physicians can adapt safely when AI outputs are incomplete, biased, or miscalibrated; autonomy, in turn, requires resilient cognitive, professional, and organizational conditions that prevent clinical judgment from being displaced by automation bias, excessive verification burden, or institutional pressure to defer to algorithmic outputs [7,10,11,15]. Therefore, the governance challenge is not only whether generative AI outputs are accurate but also whether clinical environments preserve physicians’ capacity to remain accountable decision-makers within a model of shared human-AI authority.

To address this gap, we propose a multilevel governance framework organized around 3 coordinated domains (Figure 1): cognitive workload shaping, clinical authority governance and allocation, and organizational safety governance and accountability. This resilience-centered framing explains why governance must address not only model performance but also the cognitive, professional, and institutional conditions under which physicians can sustain safe judgment and accountable authority in AI-supported care. Together, these domains aim to bound and rebalance verification work, preserve physician decisional authority while enabling safe delegation, and strengthen institutional accountability for shared risk. Figure 1 depicts the relationships among these domains and their associated governance mechanisms. To clarify the 2 implementation levels of the framework, “institutional recommendations” refers to governance design, resource allocation, workflow integration, monitoring, and oversight structures. In contrast, “physician-facing recommendations” refers to verification practices, uncertainty appraisal, patient communication, sign-off, and preservation of clinical judgment. These levels cut across all 3 domains rather than mapping onto separate domains because safe physician action often depends on institutional design, resources, and oversight.

**Figure 1.** Three-pillar governance model for enhancing physician resilience to generative artificial intelligence (AI). ASoP: automation scope of practice; PSI: protected safety investment.



This viewpoint presents a normative, concept-driven framework developed through interdisciplinary,

literature-informed synthesis rather than formal empirical data collection or expert consensus procedures. We drew on

5 research areas: resilience engineering, clinical reasoning and metacognition, CDS stewardship, AI governance and regulation, and emerging evidence on automation bias and AI-induced deskilling. From these, we identified recurrent risks at the physician-AI interface and organized corresponding governance responses across the 3 domains. The framework is intended to guide implementation and evaluation of physician-facing safeguards for generative AI in clinical care and extend rather than replicate governance approaches developed for traditional rule-based CDS given the distinct properties of generative AI outputs.

## Cognitive Workload Shaping

### Overview

In a busy outpatient clinic, a physician uses an electronic health record-integrated generative AI tool to summarize a patient's history and suggest possible explanations for subacute dyspnea. The summary is fluent but omits a recent medication change and presents several low-probability diagnoses as a coherent leading narrative. The resulting risks are 2-fold: factual incompleteness or distortion and cognitive bias induced by a plausible but misleading narrative frame. The physician may use the AI-framed explanation as an anchor, divert attention to verifying low-probability claims, or overlook clinically important information omitted from the summary. Therefore, the central governance challenge is to reduce verification burden while preserving independent clinical reasoning when AI outputs are coherent but epistemically unreliable.

### Verification Triggers and Proportionate Safeguards

Institutions should classify verification triggers using a structured matrix based on 4 dimensions: (1) clinical harm severity if the AI-supported recommendation is wrong; (2) irreversibility and time criticality of the downstream action; (3) uncertainty or conflict intensity, including model uncertainty, weak provenance, or disagreement between AI output and clinician assessment; and (4) equity sensitivity, referring to decisions in which algorithmic bias may plausibly distort recommendations [20,21]. Because algorithmic bias is difficult for individual physicians to detect case by case, equity sensitivity should be operationalized primarily as a system-level trigger supplemented by clinician judgment [20, 22]. Institutions should predefine equity-sensitive contexts using local validation results, known subgroup performance gaps, and high-impact domains that affect access, triage, diagnosis, treatment intensity, or follow-up [23,24]. The trigger may activate automatically when an AI-supported recommendation involves these contexts, when patient characteristics fall outside the model's validated population, or when local audits identify subgroup-specific performance concerns [20,25]. Demographic variables used for this purpose should support bias surveillance rather than justify differential clinical treatment. Clinicians should also be able to invoke the trigger when they identify contextual vulnerabil-

ity or discordance between the AI output and the patient's clinical or social context.

Lower-tier triggers involve reversible, low-uncertainty tasks and require brief verification, whereas higher-tier triggers should prompt independent reasoning and explicit attestation. Escalation or second review should be reserved for cases involving high severity, irreversibility, time criticality, substantial uncertainty, conflict with clinician judgment, or equity-sensitive contexts. Thresholds may vary by service line, but classification logic should remain explicit, auditable, and periodically recalibrated. Attestation interfaces are designed to discourage verbatim or formulaic rationales, with free-text fields structured to elicit clinical specificity rather than generic confirmation. Mandatory pauses and other high-interruption mechanisms are reserved for high-risk decisions, whereas lower-risk situations rely on passive or minimally disruptive cues. Routine safety governance includes monitoring of override rates, bypass patterns, and clinician-reported burden to ensure that safeguards remain salient rather than becoming background noise [25].

### Uncertainty and Trust Calibration Competency

Institutions should ensure that physician training shifts from basic AI literacy to verification-oriented competencies and calibrated trust, defined as the ability to adjust reliance on AI according to clinical context, risk level, and uncertainty signals [26,27]. Core competencies include interpreting model uncertainty and calibration indicators as reliability signals rather than proxies for ground truth; calibration indicators refer to signals about whether the model's expressed confidence or uncertainty corresponds to observed correctness in the relevant clinical context [28]. Protocolized plausibility checks against established physiology, guideline constraints, and patient-specific factors are needed, with documented attestation when a check is triggered. Training should also cover prompt hygiene, defined as structuring prompts to reduce ambiguity and prevent misleading context from shaping the model's response, and contextual integrity, defined as maintaining appropriate boundaries around what clinical information should or should not be included in an AI interaction [29]. These practices require institutional templates and interface guardrails rather than individual technical improvisation. Physicians should also appraise the provenance and recency of model-cited evidence to determine whether AI-synthesized rationales should be accepted, discounted, or escalated [30]. Together, these competencies support calibrated trust and help contain verification burden in AI-supported care.

## Clinical Authority Governance and Allocation

### Overview

In an oncology service, a generative AI system drafts a chemotherapy dosing recommendation for a patient with declining renal function and evolving laboratory

abnormalities. The output is plausible but may not reflect local protocols, recent dose adjustments, or temporal trends in laboratory values. Therefore, the central governance challenge is not only whether the recommendation is clinically correct but also who is authorized to decide, verify, and sign off on a high-risk intervention.

### Decisional Responsibility and Role Allocation

In AI-supported clinical care, safe practice depends on clearly distinguishing the AI system’s informational support role from the physician’s decisional responsibility for clinical interpretation, patient communication, and sign-off [22,24]. AI outputs do not carry independent clinical authority and cannot, by themselves, commit a patient to action. Physicians retain responsibility for interpreting outputs; accepting, modifying, or rejecting recommendations; completing required disclosure and consent processes; and signing off on clinically consequential decisions [24,31]. Accountability should be distributed in proportion to operational control: physicians remain responsible for case-specific clinical judgment; institutions remain responsible for local validation, workflow integration, and performance monitoring; and developers remain responsible for model design, governance update, and management of known limitations [22,32]. This allocation preserves physician judgment while avoiding the inappropriate concentration of liability on physicians for system-level AI failures. In value-sensitive or preference-dependent decisions, decisional authority ultimately rests with the patient, whereas physicians remain responsible for transparent communication and informed, preference-sensitive decision-making.

At the institutional level, a designated AI governance committee or equivalent multidisciplinary body should approve AI deployment within specific workflows, define local validation requirements, and oversee monitoring and escalation pathways. Such a body should include representation from clinical leadership, informatics, quality and safety, legal or compliance, and relevant specialty expertise. Patient-facing transparency, including disclosure of AI involvement and relevant uncertainty, is a core component of safe human oversight and should be tailored to risk level and use case [22,33]. To make disclosure feasible in time-limited encounters, institutions can provide standardized plain-language phrasing, electronic health record-integrated documentation prompts, and risk-stratified consent workflows [31]. For low-risk uses such as AI-assisted summarization or documentation support, brief disclosure that AI was used and reviewed by the clinician may be sufficient. Clinically consequential recommendations, including AI-supported diagnostic interpretation, medication

adjustment, triage prioritization, or treatment planning, warrant communication of 4 core points: (1) AI supported the analysis, (2) the physician reviewed and contextualized the output, (3) uncertainty or alternative interpretations remain, and (4) the patient may ask questions or decline AI involvement where legally and ethically applicable. For example, phrasing such as “An AI-supported tool helped check part of the information, but I reviewed it in light of your clinical situation” can be integrated into consent, shared decision-making, or postvisit documentation workflows. When AI outputs conflict with clinical guidelines, documented patient goals, or physician judgment, a verification trigger should activate predefined escalation pathways, with the resulting review, attestation, and rationale recorded in the audit log [20, 26,30].

### Automation Scope of Practice and Safe Delegation

To operationalize safe delegation, institutions should apply automation scope of practice (ASoP) through a stepwise decision process that links task classification, verification triggers, and authority allocation [23,24,32]. First, the workflow is divided into discrete task units, such as summarization, differential diagnosis generation, medication review, triage prioritization, or treatment recommendation. Second, each task is assigned an initial ASoP tier based on clinical risk, verifiability, time criticality, and dependence on contextual physician judgment. Third, verification triggers, including harm severity, irreversibility and time criticality, uncertainty or conflict intensity, and equity sensitivity, are applied to determine whether the task requires a higher level of review. Fourth, authority allocation rules determine who may act on the output, who must supervise or attest, and when escalation or second review is required.

Institutions can apply a 4-tier ASoP rubric (Table 1). Tier 0 includes nondelegable, physician-led tasks involving high-stakes diagnostic or therapeutic judgment, weak verifiability, time-critical consequences, or potentially irreversible harm. Assistive tasks such as draft documentation, summarization, or structured information retrieval fall under tier 1, where AI output is low risk, readily verifiable, and not directly action enabling. Tier 2 applies to supervised analytic support, including preliminary differentials, medication conflict flags, or care plan drafting, for which physician review, contextualization, and explicit acceptance are required before clinical action. At tier 3, automation may proceed within predefined institutional rules, audit trails, and escalation criteria for conditionally delegated workflow actions such as low-risk protocol-based triage support, follow-up scheduling, or routine monitoring reminders [23, 32].

**Table 1.** Illustrative automation scope of practice (ASoP) tier assignments for representative clinical tasks<sup>a</sup>.

Example task	ASoP tier	Permitted AI <sup>b</sup> role	Verification and authority requirements
Drafting or formatting a clinical note	Tier 1: assistive	Drafting or structuring text	Clinician reviews for accuracy and omissions before signing; escalates if hallucinated or incorrect content is detected.

Example task	ASoP tier	Permitted AI <sup>b</sup> role	Verification and authority requirements
Summarizing prior laboratory results, imaging, or medication history	Tier 1: assistive	Extracting and summarizing defined source information	Clinician verifies source accuracy, temporal relevance, and omissions; escalates if data are incomplete, outdated, or outside validated sources.
Generating preliminary differential diagnoses	Tier 2: supervised analytic	Suggesting candidates but not determining diagnosis	Independent clinician reasoning and explicit acceptance, modification, or rejection; escalate for high uncertainty, conflict, or high-risk presentations.
Flagging medication interactions or renal dose concerns	Tier 2: supervised analytic	Flagging concerns or drafting options	Clinician, with pharmacist input where relevant, verifies history, laboratory trends, and contraindications; escalates high-harm or protocol-conflicting cases for second review.
Protocol-based follow-up scheduling or monitoring reminders	Tier 3: conditionally delegated	Recommending or initiating predefined workflow steps	Allowed only within institutional rules, audit trails, and retrospective monitoring; escalate for status change, missing data, or out-of-protocol cases.
Low-risk protocol-based triage support such as stable outpatient referrals	Tier 3: conditionally delegated	Supporting prioritization within bounded protocols	Bounded by predefined rules; higher-risk presentations escalate to clinician confirmation; monitor for red flags or atypical features.
Final diagnosis, chemotherapy dosing, treatment selection, or withholding urgent care	Tier 0: nondelegable	Providing information or options only	Physician-led judgment with explicit attestation; second review for high-stakes or contested cases; no autonomous execution pathway permitted.

<sup>a</sup>The categories are examples for local adaptation rather than fixed universal rules.

<sup>b</sup>AI: artificial intelligence.

Tier assignment should be dynamic rather than fixed. Tasks move toward more restrictive tiers when the patient is clinically unstable, input data are incomplete or unreliable, the output falls outside the model's validated use context, the recommendation affects vulnerable or underserved groups, or the consequences of error are difficult to reverse. These assessments determine what may be automated, what must remain under physician control, and what safeguards are required [22,23].

## Organizational Safety Governance and Accountability

### Overview

In a large health system, a generative AI tool drafts responses to patient portal messages about new symptoms, possible medication adverse effects, and requests for treatment advice. Some drafts are helpful, whereas others are overly reassuring, fail to identify escalation cues, or conflict with local triage protocols. Although individual physicians may detect some errors during review, safe use depends on organizational monitoring thresholds, audit mechanisms, update oversight, and accountability structures. Therefore, the central governance challenge is not only error detection at the point of care but also ensuring that institutional oversight sustains safe physician judgment beyond individual vigilance alone.

### Targeted Interprofessional Safety Review

For prespecified high-risk scenarios such as chemotherapy dose adjustment, physicians should retain final decisional authority while convening focused interprofessional review with relevant colleagues, including pharmacists, nurses, and specialists [24,34,35]. The purpose is not to duplicate individual verification but to support team-based assessment of feasibility, safety, and cross-disciplinary workflow fit [34-36]. Activation should follow verification triggers with team-level implications, and review intensity and timing

should be adapted to clinical urgency [24,32]. In time-critical situations, review may involve rapid, role-explicit confirmation. When immediate action is required before review can occur, documented retrospective review should follow. Workflow interfaces should capture role-specific input and physician attestation in the audit log to support traceability and accountability [32,36].

### Physician-Facing Incentives and Protected Safety Investment

Organizational policy should align incentives with safety rather than throughput alone [23,37,38]. Institutions should establish protected safety investment to support nonautomatable functions that strengthen physician decision-making, including time for complex case consultation, targeted interprofessional review training, and selected skill retention activities for high-risk scenarios [38]. AI-related gains can be documented through auditable operational measures such as clinician time returned, reduced after-hours documentation, or avoided downstream workflow failures rather than assumed from adoption alone. Where direct monetization is constrained by billing context, these measures can inform local business cases, and safety investment may need justification as prospective quality infrastructure. Performance monitoring should emphasize auditable safety process measures such as verification trigger adherence, appropriate override or decline rates, attested sign-off, and time to decision in triggered cases [37].

### Skill Retention and Learning-Oriented Error Review

Skill retention activities should be risk stratified and targeted to high-risk or high-delegation workflows. Where feasible, these activities can be integrated into protected simulation time, case-based continuing medical education, credentialing-linked refreshers, or quality improvement and morbidity and mortality review processes rather than assuming that clinicians have unscheduled downtime [13,38]. Institutions

should maintain a local learning-oriented, nonpunitive error taxonomy. This taxonomy would classify AI-related safety events by source and mechanism, including interface design problems, human factor contributors, data or workflow mismatches, and model output failures. The resulting signals can guide focused review, targeted training, and interface

improvement rather than individual blame [23,39]. To clarify how the 3 framework domains translate into institution-level and physician-facing responsibilities, Table 2 summarizes the major recommendations, primary responsible parties, and practical implications for implementation.

**Table 2.** Framework recommendations by domain and responsible party.

Framework domain and major recommendation	Primary responsible party	Practical implication
<b>Cognitive workload shaping</b>		
Use risk-sensitive verification triggers and proportionate safeguards	Institution led; physician applied	Institutions define thresholds, interface cues, escalation rules, and monitoring processes; physicians apply triggered checks, document rationales, and escalate when needed.
Develop uncertainty and trust calibration competency	Shared institutional and physician responsibility	Institutions provide training, templates, and interface guardrails; physicians interpret uncertainty, assess provenance, perform plausibility checks, and calibrate reliance on AI <sup>a</sup> outputs.
<b>Clinical authority governance and allocation</b>		
Clarify decisional responsibility and role allocation	Shared institutional and physician responsibility	Institutions define disclosure, consent, accountability, and escalation pathways; physicians interpret AI outputs, communicate AI-supported reasoning and uncertainty, and retain sign-off responsibility for consequential decisions.
Define an automation scope of practice for safe delegation	Institution led; physician applied	Institutions classify tasks by risk, verifiability, competence, and supervision level; physicians use AI only within approved delegation boundaries and apply required verification.
<b>Organizational safety governance and accountability</b>		
Activate targeted interprofessional safety review for high-risk scenarios	Shared institutional and physician responsibility	Institutions define team review triggers and documentation requirements; physicians retain final authority while seeking pharmacist, nursing, specialty, or other input when triggered.
Align incentives with protected safety investment and auditable performance monitoring	Institution led	Institutions protect time and resources for complex case consultation, targeted training, and safety monitoring while avoiding throughput-only measures of AI value.
Support skill retention and learning-oriented error review	Shared institutional and physician responsibility	Institutions use risk-stratified simulation, continuing education, credentialing-linked refreshers, quality improvement processes, and a nonpunitive error taxonomy; physicians participate in targeted reassessment and remediation when needed.

<sup>a</sup>AI: artificial intelligence.

## Discussion

### *Relation to Existing Knowledge*

The proposed framework is informed by and extends several established research areas. The verification trigger mechanism draws on cognitive load theory and meta-cognitive training in medical education [12,40]. The authority allocation construct builds on resilience and error management frameworks, whereas the organizational safety pillar is informed by World Health Organization guidance, the European Union AI Act, and emerging institutional initiatives such as the Trustworthy and Responsible AI Network [26,32,41,42]. These frameworks establish essential system-level expectations for transparency, accountability, human oversight, and institutional readiness but do not specify the cognitive and workflow conditions under which physicians can sustain independent judgment while collaborating with generative AI [26,32]. Physician autonomy literature further informs the authority governance component by clarifying why physicians must retain accountable authority to interpret, modify, or override AI-supported recommendations

[18,19]. Emerging evidence on automation bias and AI-related deskilling informs the verification, delegation, and skill retention components by identifying mechanisms that may erode independent clinical judgment [7,13-15]. This framework complements these foundations by focusing more explicitly on physician-facing safeguards, including verification triggers, bounded delegation, calibrated interpretation of uncertainty, and skill retention monitoring.

### *Distinct Governance Requirements for Generative AI*

This framework extends rather than replicates governance approaches used for traditional rule-based CDS. Conventional alert-based CDS is typically narrower in scope and anchored to prespecified logic; accordingly, governance has often focused on alert thresholds, override appropriateness, and workflow burden [24,25]. In contrast, generative AI systems produce context-sensitive natural language outputs that may appear coherent and persuasive while still being false, inaccurate, biased, or incomplete [28]. Their behavior may also vary with prompt design, the quantity and order of information provided, and model updates [29]. Therefore,

governance for physician resilience in AI-supported clinical care should extend beyond traditional alert management to include bounded verification triggers, calibrated interpretation of uncertainty, protection of unaided reasoning, and explicit boundaries for AI use within clinician-led decisions [32].

### **Limitations and Implementation Considerations**

As a normative, conceptual viewpoint, this framework has several limitations. First, it does not yet provide prospective evidence regarding workflow efficiency, clinician acceptance, patient outcomes, or unintended consequences; feasibility and net benefit require pilot-testing and context-specific evaluation [43,44]. Second, safeguards such as verification triggers, diagnostic time-outs, attestation requirements, and targeted interprofessional review may introduce workflow friction, increased administrative burden, alert fatigue, or delays in time-sensitive settings if applied too broadly [25]. Emergency override pathways may also be overused, underused, or applied inconsistently in practice, requiring explicit criteria and review processes to support timely care without undermining accountability. Third, implementation capacity is likely to vary substantially across institutions, particularly in resource-constrained settings lacking sufficient staffing, digital infrastructure, monitoring capacity, or protected implementation time [23,44]. Fourth, sustaining a living ASoP table and longitudinal skill retention activities requires version control, periodic review, and coordination across multiple functions; institutions will need proportionate update cycles, shared templates, and risk-prioritized review processes to keep these mechanisms sustainable.

Accordingly, the framework should be understood as risk stratified and locally adaptable rather than uniformly

applicable. In high-acuity contexts, immediate override with deferred review may be necessary to preserve timely care. Thresholds for verification, delegation, and authority sharing will require ongoing recalibration as models, interfaces, and clinical workflows evolve. Future work should identify which components are most feasible, which provide the greatest safety benefit, and how implementation burden can be minimized while preserving physician judgment. For institutions considering adoption, implementation should proceed incrementally, beginning with governance structures and bounded pilot workflows followed by prospective monitoring, recalibration, and staged expansion, with skill retention monitoring and incentive alignment incorporated from the outset rather than after scale-up [13,43,44].

### **Conclusions**

Physician resilience in AI-supported clinical care depends on more than improved prompting or isolated human oversight requirements. This viewpoint advances a governance approach to generative AI integration that emphasizes bounded verification, preserved physician judgment, and organizational accountability as conditions for safe clinical use. Future work should prospectively evaluate the feasibility, clinical effects, and implementation burden of these mechanisms across specialties, workflows, and institutional settings. Such evaluation should include bounded pilot studies, mixed methods assessments of physician workload and autonomy, safety event monitoring, equity-focused subgroup analyses, and longitudinal assessments of skill retention.

### **Acknowledgments**

After completing the manuscript, the authors used ChatGPT (GPT-5.5; OpenAI) to assist with language refinement. The authors reviewed and edited the output and take full responsibility for the final content of the manuscript.

### **Funding**

The authors declared no financial support was received for this work.

### **Data Availability**

Data sharing is not applicable to this paper as no datasets were generated or analyzed during this study.

### **Authors' Contributions**

JL and SL contributed to conceptualization. JL, HP, and SL contributed to literature review. HP, JL, and SL contributed to writing—original draft. All authors contributed to writing—review and editing.

### **Conflicts of Interest**

None declared.

### **References**

1. Goh E, Gallo R, Hom J, et al. Large language model influence on diagnostic reasoning: a randomized clinical trial. *JAMA Netw Open*. Oct 1, 2024;7(10):e2440969. [doi: [10.1001/jamanetworkopen.2024.40969](https://doi.org/10.1001/jamanetworkopen.2024.40969)] [Medline: [39466245](https://pubmed.ncbi.nlm.nih.gov/39466245/)]
2. Liu J, Wang C, Liu S. Utility of ChatGPT in clinical practice. *J Med Internet Res*. Jun 28, 2023;25:e48568. [doi: [10.2196/48568](https://doi.org/10.2196/48568)] [Medline: [37379067](https://pubmed.ncbi.nlm.nih.gov/37379067/)]
3. Gaber F, Shaik M, Allegra F, et al. Evaluating large language model workflows in clinical decision support for triage and referral and diagnosis. *NPJ Digit Med*. May 9, 2025;8(1):263. [doi: [10.1038/s41746-025-01684-1](https://doi.org/10.1038/s41746-025-01684-1)] [Medline: [40346344](https://pubmed.ncbi.nlm.nih.gov/40346344/)]

4. Liu S, Huang SS, McCoy AB, Wright AP, Horst S, Wright A. Optimizing order sets with a large language model-powered multiagent system. *JAMA Netw Open*. Sep 2, 2025;8(9):e2533277. [doi: [10.1001/jamanetworkopen.2025.33277](https://doi.org/10.1001/jamanetworkopen.2025.33277)] [Medline: [40986301](https://pubmed.ncbi.nlm.nih.gov/40986301/)]
5. Zeng J, Zou X, Li S, et al. Assessing the role of the Generative Pretrained Transformer (GPT) in Alzheimer's disease management: comparative study of neurologist- and artificial intelligence-generated responses. *J Med Internet Res*. Oct 31, 2024;26:e51095. [doi: [10.2196/51095](https://doi.org/10.2196/51095)] [Medline: [39481104](https://pubmed.ncbi.nlm.nih.gov/39481104/)]
6. Teo ZL, Thirunavukarasu AJ, Elangovan K, et al. Generative artificial intelligence in medicine. *Nat Med*. Oct 2025;31(10):3270-3282. [doi: [10.1038/s41591-025-03983-2](https://doi.org/10.1038/s41591-025-03983-2)] [Medline: [41053447](https://pubmed.ncbi.nlm.nih.gov/41053447/)]
7. Khera R, Simon MA, Ross JS. Automation bias and assistive AI: risk of harm from AI-driven clinical decision support. *JAMA*. Dec 19, 2023;330(23):2255-2257. [doi: [10.1001/jama.2023.22557](https://doi.org/10.1001/jama.2023.22557)] [Medline: [38112824](https://pubmed.ncbi.nlm.nih.gov/38112824/)]
8. Topaz M, Peltonen LM, Zhang Z. Beyond human ears: navigating the uncharted risks of AI scribes in clinical practice. *NPJ Digit Med*. Sep 24, 2025;8(1):569. [doi: [10.1038/s41746-025-01895-6](https://doi.org/10.1038/s41746-025-01895-6)] [Medline: [40993221](https://pubmed.ncbi.nlm.nih.gov/40993221/)]
9. Kim J, Podlasek A, Shidara K, Liu F, Alaa A, Bernardo D. Limitations of large language models in clinical problem-solving arising from inflexible reasoning. *Sci Rep*. Nov 11, 2025;15(1):39426. [doi: [10.1038/s41598-025-22940-0](https://doi.org/10.1038/s41598-025-22940-0)]
10. Kosior K, Wall T, Ferrero S. The role of metacognition in teaching clinical reasoning: theory to practice. *Educ Health Prof*. 2019;2(2):108-114. [doi: [10.4103/EHP.EHP\\_14\\_19](https://doi.org/10.4103/EHP.EHP_14_19)]
11. Sokol K, Fackler J, Vogt JE. Artificial intelligence should genuinely support clinical reasoning and decision making to bridge the translational gap. *NPJ Digit Med*. Jun 10, 2025;8(1):345. [doi: [10.1038/s41746-025-01725-9](https://doi.org/10.1038/s41746-025-01725-9)] [Medline: [40494886](https://pubmed.ncbi.nlm.nih.gov/40494886/)]
12. Ghanbari S, Haghani F, Barekatin M, Jamali A. A systematized review of cognitive load theory in health sciences education and a perspective from cognitive neuroscience. *J Educ Health Promot*. 2020;9:176. [doi: [10.4103/jehp.jehp\\_643\\_19](https://doi.org/10.4103/jehp.jehp_643_19)] [Medline: [32953905](https://pubmed.ncbi.nlm.nih.gov/32953905/)]
13. Natali C, Marconi L, Dias Duran LD, Cabitza F. AI-induced deskilling in medicine: a mixed-method review and research agenda for healthcare and beyond. *Artif Intell Rev*. 2025;58(356). [doi: [10.1007/s10462-025-11352-1](https://doi.org/10.1007/s10462-025-11352-1)]
14. 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*. Oct 2025;10(10):896-903. [doi: [10.1016/S2468-1253\(25\)00133-5](https://doi.org/10.1016/S2468-1253(25)00133-5)] [Medline: [40816301](https://pubmed.ncbi.nlm.nih.gov/40816301/)]
15. Abdelwanis M, Alarafati HK, Tammam MM, Simsekler MC. Exploring the risks of automation bias in healthcare artificial intelligence applications: a Bowtie analysis. *J Saf Sci Resil*. Dec 2024;5(4):460-469. [doi: [10.1016/j.jnlssr.2024.06.001](https://doi.org/10.1016/j.jnlssr.2024.06.001)]
16. Wiig S, Aase K, Billett S, et al. Defining the boundaries and operational concepts of resilience in the resilience in healthcare research program. *BMC Health Serv Res*. Apr 19, 2020;20(1):330. [doi: [10.1186/s12913-020-05224-3](https://doi.org/10.1186/s12913-020-05224-3)] [Medline: [32306981](https://pubmed.ncbi.nlm.nih.gov/32306981/)]
17. Hollnagel E, Woods DD, Leveson N, editors. *Resilience Engineering: Concepts and Precepts*. CRC Press; 2006. [doi: [10.1201/9781315605685](https://doi.org/10.1201/9781315605685)]
18. Funer F, Wiesing U. Physician's autonomy in the face of AI support: walking the ethical tightrope. *Front Med (Lausanne)*. 2024;11:1324963. [doi: [10.3389/fmed.2024.1324963](https://doi.org/10.3389/fmed.2024.1324963)] [Medline: [38606162](https://pubmed.ncbi.nlm.nih.gov/38606162/)]
19. Grosser J, Düvel J, Hasemann L, Schneider E, Greiner W. Studying the potential effects of artificial intelligence on physician autonomy: scoping review. *JMIR AI*. Mar 13, 2025;4:e59295. [doi: [10.2196/59295](https://doi.org/10.2196/59295)] [Medline: [40080059](https://pubmed.ncbi.nlm.nih.gov/40080059/)]
20. Chen RJ, Wang JJ, Williamson DF, et al. Algorithmic fairness in artificial intelligence for medicine and healthcare. *Nat Biomed Eng*. Jun 2023;7(6):719-742. [doi: [10.1038/s41551-023-01056-8](https://doi.org/10.1038/s41551-023-01056-8)] [Medline: [37380750](https://pubmed.ncbi.nlm.nih.gov/37380750/)]
21. Nagendran M, Festor P, Komorowski M, Gordon AC, Faisal AA. Quantifying the impact of AI recommendations with explanations on prescription decision making. *NPJ Digit Med*. Nov 7, 2023;6(1):206. [doi: [10.1038/s41746-023-00955-z](https://doi.org/10.1038/s41746-023-00955-z)] [Medline: [37935953](https://pubmed.ncbi.nlm.nih.gov/37935953/)]
22. Augmented intelligence in medicine. American Medical Association. 2026. URL: <https://www.ama-assn.org/practice-management/digital-health/augmented-intelligence-medicine> [Accessed 2026-03-20]
23. Regulatory considerations on artificial intelligence for health. World Health Organization. 2023. URL: <https://www.who.int/publications/i/item/9789240078871> [Accessed 2026-01-20]
24. Clinical decision support software: guidance for industry and Food and Drug Administration staff. U.S. Food & Drug Administration. 2026. URL: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-decision-support-software> [Accessed 2026-02-22]
25. Chaparro JD, Beus JM, Dziorny AC, et al. Clinical decision support stewardship: best practices and techniques to monitor and improve interruptive alerts. *Appl Clin Inform*. May 2022;13(3):560-568. [doi: [10.1055/s-0042-1748856](https://doi.org/10.1055/s-0042-1748856)] [Medline: [35613913](https://pubmed.ncbi.nlm.nih.gov/35613913/)]

26. Ethics and governance of artificial intelligence for health: guidance on large multi-modal models. World Health Organization. 2025. URL: <https://www.who.int/publications/i/item/9789240084759> [Accessed 2026-01-16]
27. Sakamoto T, Harada Y, Shimizu T. Facilitating trust calibration in artificial intelligence-driven diagnostic decision support systems for determining physicians' diagnostic accuracy: quasi-experimental study. *JMIR Form Res*. Nov 27, 2024;8:e58666. [doi: [10.2196/58666](https://doi.org/10.2196/58666)] [Medline: [39602469](https://pubmed.ncbi.nlm.nih.gov/39602469/)]
28. Hager P, Jungmann F, Holland R, et al. Evaluation and mitigation of the limitations of large language models in clinical decision-making. *Nat Med*. Sep 2024;30(9):2613-2622. [doi: [10.1038/s41591-024-03097-1](https://doi.org/10.1038/s41591-024-03097-1)] [Medline: [38965432](https://pubmed.ncbi.nlm.nih.gov/38965432/)]
29. Sivarajkumar S, Kelley M, Samolyk-Mazzanti A, Visweswaran S, Wang Y. An empirical evaluation of prompting strategies for large language models in zero-shot clinical natural language processing: algorithm development and validation study. *JMIR Med Inform*. Apr 8, 2024;12:e55318. [doi: [10.2196/55318](https://doi.org/10.2196/55318)] [Medline: [38587879](https://pubmed.ncbi.nlm.nih.gov/38587879/)]
30. Wu K, Wu E, Wei K, et al. An automated framework for assessing how well LLMs cite relevant medical references. *Nat Commun*. 2025;16(1):3615. [doi: [10.1038/s41467-025-58551-6](https://doi.org/10.1038/s41467-025-58551-6)]
31. Ethics and governance of artificial intelligence for health: WHO guidance. World Health Organization. 2021. URL: <https://www.who.int/publications/i/item/9789240029200> [Accessed 2026-01-20]
32. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) (text with EEA relevance). European Union. URL: <https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng> [Accessed 2026-01-16]
33. WHO calls for safe and ethical AI for health. World Health Organization. 2023. URL: <https://www.who.int/news/item/16-05-2023-who-calls-for-safe-and-ethical-ai-for-health> [Accessed 2026-01-18]
34. Siegel RD, LeFebvre KB, Temin S, et al. Antineoplastic therapy administration safety standards for adult and pediatric oncology: ASCO-ONS standards. *JCO Oncol Pract*. Oct 2024;20(10):1314-1330. [doi: [10.1200/OP.24.00216](https://doi.org/10.1200/OP.24.00216)] [Medline: [38776491](https://pubmed.ncbi.nlm.nih.gov/38776491/)]
35. Weingart SN, Li JW, Zhu J, et al. US cancer center implementation of ASCO/Oncology Nursing Society chemotherapy administration safety standards. *J Oncol Pract*. Jan 2012;8(1):7-12. [doi: [10.1200/JOP.2011.000379](https://doi.org/10.1200/JOP.2011.000379)] [Medline: [22548004](https://pubmed.ncbi.nlm.nih.gov/22548004/)]
36. Martin DB, Kaemingk D, Frieze D, Hendrie P, Payne TH. Safe implementation of computerized provider order entry for adult oncology. *Appl Clin Inform*. 2015;6(4):638-649. [doi: [10.4338/ACI-2015-03-RA-0027](https://doi.org/10.4338/ACI-2015-03-RA-0027)] [Medline: [26767061](https://pubmed.ncbi.nlm.nih.gov/26767061/)]
37. Augmented intelligence in health care. American Medical Association. URL: <https://www.ama-assn.org/system/files/2019-01/augmented-intelligence-policy-report.pdf> [Accessed 2026-01-20]
38. Meeting the moment: addressing barriers and facilitating clinical adoption of artificial intelligence in medical diagnosis. National Academy of Medicine. 2022. URL: <https://nam.edu/perspectives/meeting-the-moment-addressing-barriers-and-facilitating-clinical-adoption-of-artificial-intelligence-in-medical-diagnosis/> [Accessed 2026-01-15]
39. Hospital survey on patient safety culture. Agency for Healthcare Research and Quality. 2019. URL: <https://www.ahrq.gov/sops/surveys/hospital/index.html> [Accessed 2026-01-20]
40. Chew KS, Durning SJ, van Merriënboer JJ. Teaching metacognition in clinical decision-making using a novel mnemonic checklist: an exploratory study. *Singapore Med J*. Dec 2016;57(12):694-700. [doi: [10.11622/smedj.2016015](https://doi.org/10.11622/smedj.2016015)] [Medline: [26778635](https://pubmed.ncbi.nlm.nih.gov/26778635/)]
41. Anderson JE, Ross AJ, Back J, et al. Implementing resilience engineering for healthcare quality improvement using the CARE model: a feasibility study protocol. *Pilot Feasibility Stud*. 2016;2:61. [doi: [10.1186/s40814-016-0103-x](https://doi.org/10.1186/s40814-016-0103-x)] [Medline: [27965876](https://pubmed.ncbi.nlm.nih.gov/27965876/)]
42. Embí PJ, Rhew DC, Peterson ED, Pencina MJ. Launching the Trustworthy and Responsible AI Network (TRAIN): a consortium to facilitate safe and effective AI adoption. *JAMA*. May 6, 2025;333(17):1481-1482. [doi: [10.1001/jama.2025.1331](https://doi.org/10.1001/jama.2025.1331)] [Medline: [39928518](https://pubmed.ncbi.nlm.nih.gov/39928518/)]
43. Reddy S. Generative AI in healthcare: an implementation science informed translational path on application, integration and governance. *Implement Sci*. Mar 15, 2024;19(1):27. [doi: [10.1186/s13012-024-01357-9](https://doi.org/10.1186/s13012-024-01357-9)] [Medline: [38491544](https://pubmed.ncbi.nlm.nih.gov/38491544/)]
44. Scaling innovations in public health systems: guidance and toolkit. World Health Organization. 2026. URL: <https://www.who.int/publications/i/item/9789240120761> [Accessed 2026-02-20]

## Abbreviations

**AI:** artificial intelligence

**ASoP:** automation scope of practice

**CDS:** clinical decision support

*Edited by Andrew Coristine; peer-reviewed by John Grosser, Kuan-Hsun Lin, Regina Poss-Doering; submitted 22.Nov.2025; final revised version received 11.May.2026; accepted 01.Jun.2026; published 24.Jun.2026*

*Please cite as:*

*Pan H, Liu J, Liu S*

*Enhancing Physician Resilience to Generative AI: Multilevel Framework for Shared Authority, Verification, and Skill Preservation*

*J Med Internet Res 2026;28:e88058*

URL: <https://www.jmir.org/2026/1/e88058>

doi: [10.2196/88058](https://doi.org/10.2196/88058)

© Hongxia Pan, Jialin Liu, Siru Liu. Originally published in the Journal of Medical Internet Research (<https://www.jmir.org>), 24.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in the Journal of Medical Internet Research (ISSN 1438-8871), is properly cited. The complete bibliographic information, a link to the original publication on <https://www.jmir.org/>, as well as this copyright and license information must be included.