

Artificial Intelligence Brain Fog and Artificial Intelligence Brain Fry in Healthcare and Their Implications for Clinical Reasoning

Enoch Chi Ngai Lim^{a, d} , Chi Eung Danform Lim^{a, b, c} 

Abstract

Artificial intelligence (AI) is being integrated into clinical documentation, decision support systems, healthcare education, and perioperative risk modelling. AI may enhance data synthesis, reduce administrative burden, and support individualized prediction, but it may also create cognitive risks that remain poorly defined in healthcare ethics. This narrative conceptual review distinguishes two emerging, non-diagnostic educational constructs. “AI brain fog” denotes diminished independent reasoning, recall, and clinical synthesis associated with habitual cognitive offloading to AI. “AI brain fry” denotes cognitive fatigue and overload caused by excessive AI outputs, alerts, monitoring, and verification work. The terms are not formal medical diagnoses; “brain fog” already has established clinical meanings, including in post-coronavirus disease 2019 (COVID-19) literature. The distinction is practical rather than absolute: fog is characterized primarily by under-engagement in reasoning, whereas fry is characterized by excessive cognitive demand, and both may occur together when overloaded clinicians default to shallow acceptance of AI output. Drawing on cognitive load theory, automation bias, cognitive offloading, alert fatigue, out-of-the-loop performance, healthcare education, and human–AI interaction, this review proposes observable markers, immediate remediation strategies, an interprofessional cognitive resilience framework, implementation responsibilities and metrics, and measurable hypotheses for prospective validation. Safe AI adoption requires technical evaluation alongside intentional protection of human judgment, diagnostic humility, attention, and professional accountability.

Keywords: AI brain fry; AI brain fog; Artificial intelligence; Clinical reasoning; Cognitive offloading; Healthcare education; Automation bias; Cognitive resilience

Introduction

Artificial intelligence (AI) is changing how clinicians and healthcare teams gather data, make assessments, document care, and convey their findings. Clinical AI tools can draft notes, summarize records, suggest possible diagnoses, prioritize risks, and recommend treatments. An ethical assessment of AI tools from an Australian healthcare perspective has asserted that unsettled AI technologies used in patient care must protect safety, remain transparent and accountable, and preserve professional responsibilities [1]. The World Health Organization guidance on large multi-modal models similarly warns that general-purpose AI in health can propagate false content, amplify bias, promote automation bias, compromise privacy, and be poorly governed in clinical decision-making [2].

The gap between the technology available to clinicians and clinicians’ preparedness to use it has widened. In a survey of 39 primary care clinicians in metropolitan Australia, average AI familiarity was just below 2 on a 5-point scale; 64% intended to use AI in their work to some extent, while concerns centered on privacy, AI-induced errors, hallucinations, and integration with existing workflows [3]. This suggests willingness to adopt AI in the absence of mature cognitive and ethical frameworks for safe use. Large language models have ingested extensive medical knowledge but retain gaps in safety, factuality, and clinical value [4]. GPT-4 has shown strong performance on clinical reasoning tasks compared with practicing physicians [5]. However, in a randomized clinical trial, access to a large language model that performed well in isolation did not improve clinicians’ diagnostic reasoning beyond conventional resources [6]. A 2026 Science study likewise reported that a large language model exceeded clinicians’ baseline performance on selected reasoning tasks and showed the value of clinical adjudication [7]. These findings create an empirical tension: clinicians may over-rely on AI and weaken independent reasoning, or fail to use AI when it could improve care.

“Brain fog” is used in biomedical literature to describe

Manuscript submitted June 20, 2026, accepted July 24, 2026
Published online July 30, 2026

^aTranslational Research Department, Specialist Medical Services Group, Earlwood, NSW 2206, Australia

^bNICM Health Research Institute, Western Sydney University, Westmead, NSW 2145, Australia

^cData Science Institute, University of Technology Sydney, Ultimo, NSW 2007, Australia

^dCorresponding Author: Enoch Chi Ngai Lim, Translational Research Department, Specialist Medical Services Group, Earlwood, NSW 2206, Australia. Email: enoch@smsg.au

doi: <https://doi.org/10.14740/aicm29>
AI in Clinical Medicine
2819-7437 (online)

subjective cognitive symptoms such as reduced attention, memory, language fluency, and mental clarity, although the term remains variably defined [8]. It is also prominent in post-coronavirus disease 2019 (COVID-19) literature as a patient-reported cluster of persistent cognitive symptoms [9]. In this article, “AI brain fog” is not a clinical diagnosis. It is an operational concept for reduced independent reasoning that may arise when clinicians repeatedly offload clinical synthesis, recall, and judgment to AI. “AI brain fry” has a more direct source: Bedard et al used the term for mental fatigue associated with extensive use or monitoring of AI tools in a workplace study of 1,488 United States workers [10]. Cognitive load theory provides a framework for this concern because reasoning may deteriorate when working-memory demands exceed human capacity [11].

In the literature reviewed, AI brain fog and AI brain fry are not established as a paired classification and are not routinely used interchangeably. AI brain fog is an author-proposed synthesis of literature on brain fog, cognitive offloading, dependency, automation, and deskilling; AI brain fry adapts a workplace term to healthcare workload and monitoring. Their distinction concerns the dominant direction of cognitive demand: too little active reasoning in fog versus too much information and oversight in fry. They may coexist, and both are conceptual educational tools intended to support observation, discussion, and research rather than diagnoses or validated syndromes. Differentiating them matters because the immediate response differs: fog calls for re-engagement with independent reasoning, whereas fry first calls for reduction and prioritization of cognitive load.

Methods

This article was designed as a narrative conceptual review rather than a systematic review or meta-analysis. The aim was not to provide pooled effect estimates, but to develop a transparent conceptual synthesis of emerging cognitive risks associated with AI in healthcare. The review focused on two non-diagnostic constructs: AI brain fog, referring to diminished independent clinical reasoning associated with excessive cognitive offloading to AI, and AI brain fry, referring to cognitive overload and fatigue arising from excessive AI-generated outputs, alerts, verification work, and monitoring demands.

A structured narrative search strategy was used to improve transparency and reproducibility. Literature was identified through PubMed/MEDLINE, Embase, Google Scholar, and relevant publisher websites. The searches combined terms related to AI and healthcare with terms related to cognition, clinical reasoning, education, and workload. Search terms included “artificial intelligence,” “healthcare,” “clinical reasoning,” “cognitive offloading,” “automation bias,” “alert fatigue,” “healthcare education,” “medical education,” “nursing education,” “large language models,” “AI dependency,” “clinical decision support,” “cognitive load,” “ambient AI scribes,” “diagnostic reasoning,” and “human-AI interaction.” Searches were supplemented by backward citation searching (backward snowballing) of relevant articles and by foundational cogni-

tive-science literature when it directly informed the conceptual framework [12].

The review prioritized peer-reviewed empirical studies, systematic reviews, narrative reviews, conceptual analyses, and major policy or governance documents relevant to AI in healthcare. Greater emphasis was placed on literature published between 2021 and 2026, reflecting the rapid development of generative AI and large language models in clinical and educational settings. Earlier studies were included when they provided foundational concepts, particularly cognitive offloading, memory, automation bias, cognitive load theory, out-of-the-loop performance, and clinical reasoning. Non-healthcare AI literature was considered only when it provided a mechanism plausibly transferable to healthcare, such as workplace AI-related fatigue or human-AI workload effects.

Studies and documents were eligible for inclusion if they addressed at least one of the following domains: AI-assisted clinical reasoning, cognitive offloading, automation bias, alert fatigue, healthcare education, clinician or healthcare-worker workload, AI-generated documentation, clinical decision support, human-AI interaction, or governance of AI in healthcare. Articles were excluded if they did not address healthcare or a directly transferable cognitive mechanism, did not examine cognitive or professional implications of AI use, or focused solely on technical model performance without implications for reasoning, education, workload, or governance.

The synthesis proceeded in three stages. First, relevant literature was grouped into thematic domains: human clinical cognition; cognitive offloading and dependency; automation bias and out-of-the-loop performance; cognitive load and alert fatigue; healthcare education; ambient AI documentation; clinical decision support; and healthcare AI governance. Second, these domains were mapped to the proposed constructs of AI brain fog and AI brain fry, including areas of overlap. Third, recurring risks were translated into a cognitive resilience framework for clinicians, educators, and health services, including human-first reasoning, AI as a second opinion, structured verification, cognitive-load control, AI-free competency assessment, and reflective accountability.

Because this was a narrative conceptual review, formal risk-of-bias assessment and meta-analysis were not undertaken. The review does not claim to exhaustively identify all studies, establish prevalence or causality, or validate AI brain fog or AI brain fry as diagnostic entities. Evidence drawn from non-clinical settings is used to generate testable mechanisms, not to imply that effects have already been demonstrated in healthcare. The proposed terms are conceptual and educational tools rather than formal clinical diagnoses.

Conceptual Synthesis

Human clinical cognition in AI-assisted practice

Clinical reasoning involves more than retrieval of facts. It includes problem representation, pattern recognition, management of uncertainty, formulation and testing of hypotheses, and reflection and adaptation. Dual-process models remain rel-

evant because clinicians move between rapid pattern recognition and slower analytical reasoning. Cognitive load can affect diagnostic performance, particularly in intensive care and other settings where data are incomplete and time is constrained [13]. AI alters this cognitive context by retrieving information, condensing records, and suggesting diagnostic possibilities, while also influencing what clinicians notice, remember, and question. Previous work on medical synthesis has described how AI may help integrate complex information while raising questions about standardization, interpretability, and professional judgment [14].

Automation can also change the clinician's role from active synthesis to passive monitoring. In other safety-critical domains, this "out-of-the-loop" condition has been associated with reduced situation awareness and slower recovery when automation fails [15]. Direct transfer to healthcare requires prospective testing, but the mechanism is relevant where a clinician must rapidly resume unaided reasoning after an erroneous output. The question is therefore not whether AI should assist cognition; its role will continue to expand. The more important question is whether AI use strengthens or weakens the reasoning habits needed when cases are ambiguous, data are incomplete, or AI output is wrong.

AI brain fog: cognitive offloading and diminished clinical synthesis

AI brain fog is defined here as a decline or underuse of independent clinical reasoning associated with habitual reliance on AI in situations where active synthesis would ordinarily be required. Observable manifestations may include an incomplete problem representation, inability to recall or reconstruct the reasoning path, a narrow differential diagnosis, inability to justify a clinical plan without reopening the AI output, and failure to identify disconfirming evidence. The term should be used cautiously because brain fog is an established but variably defined clinical descriptor, including in COVID-19 literature [8, 9]. AI brain fog does not describe a neurological syndrome; it is an educational and research concept.

Cognitive offloading underpins this concern. The "Google effect" describes how people may remember where information can be found while retaining less of the information itself [16]. Risko et al defined cognitive offloading as the use of an external system to reduce demands on internal cognitive resources [17]. Offloading can be efficient and adaptive; the risk arises when the external tool replaces the rehearsal, retrieval, or synthesis required to maintain competence, rather than removing low-value work while preserving judgment.

Clinical care may be especially vulnerable because clinicians routinely work with incomplete and contradictory information. A learner who asks AI to generate a differential diagnosis before framing the problem offloads the generative step. An emergency or primary care clinician who accepts an AI summary may miss a contradiction in the record. A radiology reader who sees an AI label before an independent search may narrow attention prematurely. A nurse, nurse practitioner, or physician assistant who follows a prompted assessment may

document individual elements without integrating the patient's trajectory. Conversely, a team can appropriately offload low-value retrieval while retaining responsibility for identifying the most dangerous diagnosis not to miss. The concern is therefore not offloading itself, but the transfer of cognitive authority for tasks that require professional synthesis [15, 17].

A survey of 319 knowledge workers found that greater confidence in generative AI was associated with less self-reported critical thinking, while the locus of effort shifted from information generation toward verification, integration, and task stewardship [18]. This non-clinical study does not establish AI brain fog in healthcare, but it supports a plausible mechanism for reduced engagement. Healthcare education is particularly relevant. Lim et al argued that AI dependency may diminish clinical intuition, empathy, and diagnostic confidence [19]. Learners require deliberate practice before their internal checks are sufficiently developed to use external supports safely; AI should therefore be sequenced to follow, rather than replace, initial problem representation and reasoning.

AI brain fry: overload, alerts, and verification work

AI brain fry is defined here as cognitive exhaustion caused by the volume, fragmentation, and monitoring demands of AI-assisted work. Bedard et al used the term for fatigue arising when people use or monitor AI tools beyond their cognitive capacity [10]. In healthcare, AI may reduce drafting or retrieval while increasing the need to review draft notes, reconcile multiple outputs, and verify recommendations. Cognitive load theory predicts that limited working-memory capacity can be consumed by verification, integration, and revision tasks [11]. Long, poorly prioritized summaries, low-value alerts, and disjointed or contradictory recommendations may therefore produce hasty decisions, missed signals, alert dismissal, and reliance on default options.

Clinical decision-support systems already demonstrate related workload effects. A systematic review found that alerts which disrupt workflow may be overridden, silenced, or ignored [20]. Combined with automation bias, fatigue can create risk in both directions: excessive trust in a recommendation and disengagement from useful warnings [21]. In a randomized clinical-vignette study involving hospitalist physicians, nurse practitioners, and physician assistants, standard AI modestly improved diagnostic accuracy, but systematically biased AI substantially reduced it; commonly used image-based explanations did not eliminate the harm [22]. This finding supports an interprofessional, rather than physician-only, account of AI-related cognitive risk.

Fatigued clinicians may accept a recommendation after only superficial confirmation—checking that the output is present, complete, or superficially plausible—rather than assessing its validity. In this review, structured verification means checking patient identity and context, evidentiary source, internal consistency, medication and dose accuracy, contradictory data, uncertainty, and fit with the patient's values and circumstances. Completing a review checkbox is not equivalent to validating a recommendation.

Ambient AI scribes illustrate the benefit-burden trade-off. Ma et al reported reductions of 6.89 min in average daily documentation time, 5.17 min in after-hours documentation, and 19.95 min in total daily electronic health-record time [23]. Generated notes still require review, especially where medication changes, symptom chronology, or patient preferences are concerned. Much of the current ambient-scribe evidence is physician-centered. A nursing commentary has highlighted risks of omission, hallucination, and bias when nurses are excluded from design and oversight [24]. The literature reviewed did not directly test whether early-career nurses who depend on clinical decision-support prompts have difficulty narrating an unprompted assessment. This is a plausible but currently unverified training concern and should be studied rather than assumed.

Clinical differentiation, co-occurrence, and immediate management

The practical distinction is based on the dominant observable problem, not on a formal diagnosis. AI brain fog is most likely when the clinician appears cognitively under-engaged: the output is accepted with little reconstruction, the plan cannot be explained without returning to AI, or important alternatives were never generated. AI brain fry is most likely when the clinician appears overloaded: output volume, interruptions, or monitoring demands lead to fatigue, skimming, delayed review, or indiscriminate dismissal. These patterns are consistent with the literature on brain fog, AI-related fatigue, cognitive load, cognitive offloading, out-of-the-loop performance, alert fatigue, and automation bias [8–11, 15, 17, 20–22]. A proposed bedside or supervisory probe is to ask the clinician to restate the problem, working diagnosis, dangerous alternative, and reasons for accepting or rejecting the AI output without reopening it. Difficulty doing so suggests fog; restoration of performance after reducing output and interruptions suggests fry.

The constructs can occur together. Drawing on cognitive-load, alert-fatigue, automation-bias, and cognitive-offloading mechanisms, a long, unprioritized AI-generated summary may first produce brain fry; the resulting fatigue may then prompt shallow acceptance of the recommendation, producing a fog-like failure to reason independently [11, 17, 20, 21]. Fog may reflect a gradually learned habit, whereas fry may present as an acute state, but neither distinction is absolute. Co-occurrence should therefore be managed in sequence: reduce the immediate cognitive burden, then re-establish independent reasoning and high-risk verification.

Based on these mechanisms, the authors propose that when fog is suspected, AI consultation should be paused while the clinician independently restates the problem representation, differential diagnosis, disconfirming evidence, and patient-specific risk. When fry is suspected, outputs should be triaged or limited, low-value alerts suppressed, review responsibility assigned, and protected review time provided. These responses are consistent with evidence on cognitive load, cognitive offloading, out-of-the-loop performance, alert fatigue, and automation bias, although they require prospective clinical validation [11, 15, 17, 20–22]. When both are present, cog-

nitive load should be reduced first, followed by an AI-free restatement and structured verification of medications, diagnoses, and follow-up. If safe independent review cannot be completed, the case should be escalated rather than processed through further AI output. The distinguishing features, areas of overlap, and corresponding remediation and prevention strategies for AI brain fog and AI brain fry are summarized in Table 1 [11, 15, 17, 20–22].

Impact across health professions and training stages

AI-related cognitive risks are unlikely to be distributed equally across the healthcare system, and they are not limited to physicians. Health professions students may be vulnerable because diagnostic, assessment, and illness scripts are still developing. Early-career clinicians who diagnose or prescribe—including junior doctors, nurse practitioners, and physician assistants—may be vulnerable to premature closure under time pressure. Experienced clinicians may face gradual deskilling in rare or complex tasks if AI replaces repeated practice. Nurses, allied health professionals, pharmacists, diagnostic professionals, and administrative staff may experience brain fry through alerts, documentation, standardized recommendations, or automation at scale. Evidence for the exact prevalence and gradient by profession or seniority remains limited; Table 2 is therefore hypothesis-generating. Its mechanisms are supported by research on out-of-the-loop performance, alert fatigue, automation bias across physicians, nurse practitioners and physician assistants, and AI in nursing education [15, 20, 22, 25].

Healthcare education literature supports this cautious, interprofessional approach. A systematic review of large language models in medical education identified opportunities for teaching, assessment, and feedback alongside concerns about accuracy, bias, integrity, and learner dependence [26]. A systematic review of 15 studies involving 1,464 nursing students and professionals found potential educational benefits but also emphasized the absence of longitudinal evidence and robust measurement [25]. Virtual-patient tools can support clinical reasoning when they require active engagement rather than passive retrieval of answers [27]. Across professions, a practical educational rule is that AI should create more reasoning, not less: learners should formulate an assessment or differential before viewing AI output and should be able to explain why they accepted, modified, or rejected it.

Potential benefits and the accuracy-dependency paradox

AI brain fog and AI brain fry should not be used to dismiss AI. The stronger argument is that benefits require safeguards. Several perioperative studies illustrate why clinicians may value AI-assisted modelling [28]. A causal machine-learning study of 3,017 adults undergoing non-cardiac surgery found heterogeneous renal effects associated with phenylephrine, with age, baseline renal function, and American Society of Anesthesiologists (ASA) physical status acting as important effect modifiers [29]. A study of 67,134 surgical procedures identified nonlin-

Table 1. Conceptual Distinction Between AI Brain Fog and AI Brain Fry

Domain	AI brain fog	AI brain fry
Core meaning	Reduced independent reasoning or clarity associated with excessive reliance on AI	Cognitive fatigue caused by excessive AI outputs, alerts, monitoring, or verification tasks
Dominant demand pattern	Under-engagement: too little active generation, retrieval, or synthesis	Overload: too much information, interruption, task switching, or oversight
Main mechanism	Cognitive offloading, automation dependency, and reduced effortful reasoning	Cognitive overload, alert fatigue, vigilance demands, and task switching
Typical setting	Healthcare education, diagnostic workup, imaging review, and literature synthesis	Electronic records, decision-support dashboards, ambient documentation, and messaging systems
Observable clinical marker	Clinician cannot explain or reconstruct a plan without returning to the AI output	Clinician skims, ignores, delays, or indiscriminately dismisses AI-generated material
Subjective or workflow cue	The task feels easy because AI has done the thinking, but understanding is shallow	The task feels mentally saturated; attention and review quality decline as outputs accumulate
Co-occurrence	May follow fry when fatigue triggers shallow acceptance and cognitive offloading	May precipitate fog when overload causes default acceptance of AI recommendations
Ethical risk	Deskilling, shallow understanding, premature closure, and diluted agency	Missed signals, alert dismissal, burnout, and diluted accountability
Immediate remediation strategy	Pause AI; restate the problem, differential, disconfirming evidence, and plan unaided	Reduce and prioritize outputs; suppress low-value alerts; protect review time and redistribute monitoring
Preventive strategy	Human-first reasoning, AI-free assessment, spaced retrieval, and reflective case review	Output prioritization, alert tuning, workflow limits, concise interfaces, and protected review time

The proposed distinctions synthesize literature on cognitive load, cognitive offloading, out-of-the-loop performance, alert fatigue and automation bias [11, 15, 17, 20–22]. The observable markers and remediation strategies are proposed conceptual applications for healthcare practice. AI: artificial intelligence.

ear fentanyl–propofol dose–response patterns associated with postoperative intensive care unit (ICU) admission risk [30]. Tabular foundation modelling has also shown promise in perioperative classification, although performance varied by sample size, outcome incidence, and calibration [31]. Related work has proposed TabPFN for real-time precision modelling in surgical health economics [32]. These studies support the claim that AI can reveal patterns that may be difficult to detect unaided.

The ethical risk emerges when accuracy creates dependence—the accuracy-dependency paradox. As AI becomes more useful, clinicians may grant it more cognitive authority; as authority shifts, independent reasoning may be practiced less often. The pattern may differ by specialty. In radiology, an accurate triage label can speed review but may narrow the independent visual search when it is wrong. In emergency and hospital medicine, an AI-generated differential may improve breadth but also anchor the team on an early frame. In perioperative and critical care, a risk model may reveal nonlinear associations while weakening attention to the physiological rationale if the score becomes the endpoint. In primary care, automated summaries may reduce inbox workload yet conceal contradictions or medication changes. In nursing and pharmacy, escalation or verification tools may speed recognition but also encourage template-driven assessment or superficial approval. These examples are mechanism-based illustrations, not evidence that every specialty will experience the same effect [15, 22].

There is no single correct organizational response. Depending on the task, evidence, and benefit–risk profile, a health ser-

vice may retain, redesign, restrict, pause, or withdraw an AI tool. Where the tool is retained, a proportionate goal is to preserve clinical agency rather than reflexively weaken useful capability. Interfaces can ask clinicians to record a working diagnosis before viewing advice, identify disconfirming evidence, state uncertainty, and name the patient-specific factor that could make a recommendation unsafe. The appropriate balance should be evaluated empirically for each role and workflow.

Cognitive resilience framework

Cognitive resilience is the ability of clinicians and teams to use AI while preserving independent reasoning, attention, accountability, and patient-centered judgment. It requires individual habits, interprofessional educational design, and organizational governance. Table 3 [17, 19–22, 25, 26, 33] presents a practical framework for embedding cognitive resilience in AI-assisted healthcare.

This framework reframes AI literacy. Healthcare professionals need to know not only how AI works, but how it changes their own thinking. AI literacy should include prompt formulation, bias recognition, uncertainty appraisal, structured verification, awareness of cognitive offloading and overload, and the ability to stop using AI when a task requires direct human attention. These competencies should be tailored to scope of practice and experience rather than delivered as a single physician-centered curriculum.

Table 2. AI-Related Cognitive Risks Across Healthcare Groups and Experience Levels

Group and stage	Likely AI exposure	Main cognitive risk	Protective strategy
Health professions students and trainees (medicine, nursing, nurse practitioner and physician assistant, pharmacy, allied health)	AI tutors, case simulators, exam tools, and assessment prompts	Learning an answer or template before developing the reasoning or assessment path	Require independent problem representation or assessment before AI consultation; retain AI-free examinations
Early-career diagnosing and prescribing clinicians (junior doctors, novice nurse practitioners and physician assistants)	Draft notes, differential-diagnosis tools, order suggestions, and risk scores	Premature closure, shallow verification, and reduced tolerance of uncertainty	Supervisor-led comparison of human and AI reasoning; explicit dangerous-alternative and disconfirming-evidence checks
Experienced diagnosing and prescribing clinicians (senior physicians, experienced nurse practitioners and physician assistants)	Risk models, diagnostic second opinions, literature synthesis, and complex decision support	Gradual deskilling in rare or complex cases and automation-induced anchoring	Periodic AI-free complex-case review, peer calibration, and audit of cases in which AI changed management
Early-career nurses and allied health clinicians	Assessment prompts, care-plan generation, escalation alerts, and documentation support	Prompt dependence, over-standardized assessments, and alert fatigue	Require unaided assessment summaries, role-specific supervision, and prioritized escalation pathways
Experienced nurses and allied health clinicians	Triage systems, longitudinal summaries, workload allocation, and treatment suggestions	Loss of situational awareness or patient-specific nuance when workflows become template-driven	Co-design alerts and templates; retain narrative handover, exception review, and patient-centered override pathways
Pharmacists and diagnostic professionals	Medication verification, interaction alerts, imaging or laboratory prioritization, and report drafting	High-volume verification burden, false-positive fatigue, and anchoring on AI labels	Risk-tiered queues, independent first reads for selected cases, and audit of false-positive and false-negative consequences
Administrative and operational staff	Scheduling, recall systems, coding, billing, and capacity management	Automation of errors across large patient groups and exception overload	Audit trails, sampling of automated actions, exception review, and clear human escalation pathways

These profession- and experience-specific risks are hypothesis-generating rather than estimates of established prevalence. The proposed mechanisms are informed by literature on out-of-the-loop performance, alert fatigue, automation bias across physicians, nurse practitioners and physician assistants, and AI in nursing education [15, 20, 22, 25]. AI: artificial intelligence.

System design must complement individual vigilance. Health services should determine whether AI reduces workload or redistributes it into less visible verification and monitoring. Potential metrics include time spent reviewing generated notes, alert volume, output length, changes in diagnoses or actions after AI consultation, override and dismissal patterns, near misses, corrected documentation errors, clinician confidence without AI support, and disparities across professions, settings, and patient groups. These measures can distinguish genuine assistance from a shift of cognitive work to verification.

Operationalizing the framework

Operational responsibility should be explicit, consistent with governance guidance emphasising patient safety, transparency, accountability, and preservation of professional responsibility [1, 2]. Each AI tool should have an executive sponsor, a named clinical service owner, and oversight by a multidisciplinary AI governance group. Membership should include physicians, nurses, nurse practitioners, physician assistants, pharmacists, allied health professionals, informaticians, educators, human-factors and quality-safety specialists, privacy and legal rep-

resentatives, and patient or consumer representatives where appropriate. The clinical owner should define safe use and escalation; educators should define competency requirements; information technology and vendors should provide telemetry and change control; and quality-safety teams should conduct independent audits. This distribution of responsibility also aligns with implementation frameworks that emphasize stakeholder involvement and organizational context [33].

Compliance should be monitored through workflow data and sampled case review, not solely through self-attestation. This is important because alert fatigue [20], automation bias and erroneous AI recommendations [21, 22], and errors or omissions in AI-generated documentation [23, 24] may not be captured by simple attestations. Before deployment, organizations should establish baseline workload, error, and performance measures, while staged pilots should use predefined success criteria and stop rules [2, 33]. Early monitoring may be monthly and can become less frequent only after stable performance is demonstrated. Audits should examine review time, alert burden, unreviewed outputs, changes in management, accepted incorrect recommendations, corrected scribe errors, near misses, and the ability to complete selected tasks without AI [20–24]. Monitoring should support learning rather than punitive surveillance,

Table 3. Cognitive Resilience Framework for AI-Assisted Healthcare

Principle	Clinical practice	Health professions education	Organizational safeguard
Human-first reasoning	Clinician records an initial assessment before viewing AI advice	Learners submit a problem representation, assessment, or differential before AI use	Interface design preserves the clinician's first impression and time-stamps AI exposure
AI as second opinion	AI output is compared explicitly with the human plan	Learners identify agreement, disagreement, and uncertainty	Audit cases in which AI changed diagnosis, treatment, escalation, or follow-up
Verification discipline	High-risk outputs undergo source, consistency, medication, dose, and patient-fit checks	Teach hallucination and bias detection using profession-specific cases	Require review for medications, diagnoses, handovers, and follow-up plans; define accountable reviewer
Cognitive load control	Limit low-value alerts, duplicated outputs, and unnecessarily long generated text	Teach concise prompting, output triage, and stopping rules	Monitor alert volume, review time, override patterns, and output length; tune thresholds by role
AI-free competence	Maintain selected sessions and high-risk rehearsals without AI support	Use AI-free objective structured clinical examinations, simulations, and case discussions	Credentialing and continuing competence include unaided reasoning or assessment tasks
Role- and stage-sensitive use	Match AI autonomy and explanation to experience, role, and clinical risk	Increase scaffolding gradually while preserving independent performance	Co-design with physicians, nurses, nurse practitioners, physician assistants, pharmacists, allied health, and administrative users
Reflective accountability	Clinician documents why consequential AI advice was accepted, modified, or rejected	Case reflection includes ethical, cognitive, and patient-centered analysis	Governance committees review AI-related near misses, adverse events, and inequitable effects

This framework is proposed by the authors based on the synthesis of literature concerning AI dependency, cognitive offloading, automation bias, alert fatigue, healthcare education and implementation complexity [17, 19–22, 25, 26, 33]. AI: artificial intelligence.

with rapid feedback to frontline users [2, 33].

Implementation barriers include acquisition and maintenance costs, workflow disruption, interoperability, training time, unequal access to support, vendor lock-in, and the burden of ongoing evaluation; concerns about workflow integration have also been reported among frontline primary care clinicians [3]. The Nonadoption, Abandonment, Scale-up, Spread, and Sustainability (NASSS) framework emphasizes that sustained adoption depends on interacting complexity across the condition, technology, value proposition, adopters, organization, wider system, and change over time [33]. Practical mitigation includes high-risk-first pilots, co-design with each affected profession, protected training time, fallback workflows for downtime or uncertainty, procurement rights to audit logs and performance data, and a recurring budget for evaluation [2, 33]. A tool should be paused, reconfigured, or withdrawn when benefit cannot be demonstrated or when cognitive burden, inequity, or safety risk worsens [1, 2, 33].

Future research and prospective validation hypotheses

The proposed constructs require prospective validation rather than rhetorical acceptance. Studies should test whether they have discriminant validity, whether they predict clinically important outcomes, and whether targeted safeguards reduce risk. Table 4 converts the conceptual claims into measurable hypotheses suitable for educational, simulated, and real-world studies.

Controlled educational studies can test H1 by allowing

one group to access AI before forming a diagnosis or assessment and requiring another group to form an independent problem representation first. Outcomes should extend beyond immediate accuracy to explanation quality, delayed retention, transfer to new cases, and performance after AI is removed. Longitudinal cohorts should include medical, nursing, nurse practitioner, physician assistant, pharmacy, and allied health trainees, while measuring baseline knowledge, sleep, burnout, workload, and AI exposure to avoid attributing all cognitive change to AI.

AI brain-fry research should quantify the number, type, length, timing, and priority of outputs received during real work. Validated workload measures such as the NASA Task Load Index can be combined with system logs, review omissions, decision time, eye-tracking where feasible, and qualitative accounts [34]. Ambient-scribe studies should include nurses and other professions and test whether repeated use changes the quality of unprompted assessment and handover. Clinical decision-support trials should compare displays of uncertainty, opposing evidence, and alternative diagnoses, and should measure both appropriate reliance when AI is correct and resistance when it is wrong.

Construct-validation studies should determine whether AI brain fog and AI brain fry are distinguishable from ordinary fatigue, burnout, sleep deprivation, low knowledge, and general digital overload. A credible measure of fog should predict loss of unaided reasoning or retention after controlling for workload; a credible measure of fry should predict acute review failure after controlling for baseline competence. Real-world studies should

Table 4. Proposed Measurable Hypotheses for Prospective Validation

Hypothesis	Suggested design	Primary measures
H1—Ordering effect (brain fog): Viewing AI before independent problem representation will reduce delayed unaided reasoning and explanation quality compared with viewing AI after an initial assessment.	Randomized parallel or crossover study in health professions learners and clinicians; AI-first versus human-first sequence.	Immediate and 2–4-week unaided diagnostic or assessment accuracy; explanation rubric; differential breadth; recall of case rationale.
H2—Output-load effect (brain fry): Verbose, unprioritized outputs will increase workload and clinically important omissions compared with concise, prioritized outputs containing the same facts.	Randomized interface trial using controlled output length, alert count, and prioritization.	NASA Task Load Index [34]; decision time; review omissions; alert dismissal; error detection.
H3—Discriminant validity: A brain-fog measure will relate more strongly to unaided reasoning and retention, whereas a brain-fry measure will relate more strongly to acute workload, skimming, and missed review.	Scale development with cognitive interviews, factor analysis, and external validation across professions.	Factor structure; reliability; correlations with unaided performance, workload, sleep, and burnout; incremental validity.
H4—Co-occurrence interaction: High output load combined with an AI-first workflow will produce more uncritical acceptance of incorrect recommendations than either exposure alone.	Two-by-two factorial experiment: low/high output load × human-first/AI-first sequence.	Acceptance of incorrect AI advice; commission errors; omitted alternatives; interaction effect size.
H5—Remediation effect: Human-first interfaces plus a disconfirming-evidence prompt will reduce erroneous AI acceptance without a clinically unacceptable delay or loss of benefit when AI is correct.	Randomized non-inferiority or superiority trial comparing standard and resilience-oriented interfaces.	Diagnostic or treatment accuracy; time to decision; appropriate reliance; override quality; near misses.
H6—Role and experience moderation: Brain-fog and brain-fry effects will differ by profession, training stage, AI familiarity, and baseline workload.	Multisite, stratified cohort or pragmatic trial including physicians, nurses, nurse practitioners, physician assistants, pharmacists, allied health, and administrative staff.	Profession-by-exposure interactions; role-specific performance; workload; retention; safety events.

AI: artificial intelligence.

also test the proposed sequence for co-occurrence: reduce load first, then restore independent reasoning and structured verification. These studies would establish whether the terms add explanatory and practical value beyond existing constructs.

Conclusions

AI brain fog and AI brain fry are conceptual educational tools for naming different cognitive risks associated with AI-assisted healthcare. AI brain fog describes under-engagement and possible erosion of independent reasoning when AI supplies the first-pass assessment or synthesis. AI brain fry describes overload and fatigue produced by excessive outputs, alerts, monitoring, and verification. They are not formal diagnoses, do not have a clinical ontology, and may occur together. Their practical value lies in linking observable problems to different immediate responses: re-engage reasoning for fog, reduce load for fry, and do both sequentially when they coexist.

The central ethical question is not whether AI is proper in healthcare in the abstract, but whether a particular system and workflow preserve the cognitive capabilities required for safe care. Highly accurate AI may influence documentation, prediction, education, and professional practice in subtle ways. When a clinician cannot explain a diagnosis without the AI output, or when a team is desensitized by repeated alerts, responsibility lies not only with the individual but also with the design, implementation, and governance of the system.

Human-first reasoning, AI as a second opinion, structured verification, AI-free competence assessment, role-sensitive education, and workload-aware system design should be evaluated as core components of AI governance. Implementation requires named owners, interprofessional oversight, measurable performance indicators, staged deployment, and willingness to reconfigure or withdraw a tool when harms exceed benefits. Healthcare AI should be judged not only by model performance and time saved, but by whether it helps professionals maintain clear, critical, patient-centered thinking when AI is absent, uncertain, or wrong.

Acknowledgments

None to declare.

Financial Disclosure

None to declare.

Conflict of Interest

None of the authors declared conflicts of interest regarding this manuscript.

Author Contributions

Conceptualization: ECNL, CEDL. Data curation: ECNL. Formal analysis: ECNL, CEDL. Funding acquisition: CEDL. Investigation: ECNL. Methodology: ECNL. Project administration: ECNL. Supervision: CEDL. Visualization: ECNL, CEDL. Writing – original draft: ECNL, CEDL. Writing – review and editing: ECNL, CEDL.

Data Availability

The authors declare that data supporting the findings of this study are available within the article.

AI Use Declaration

The authors acknowledge the use of Grammarly solely for language editing, grammar checking, spelling correction, and stylistic refinement of the manuscript. The authors take full responsibility for the content, interpretation, and conclusions presented in this work. No AI tool was used to generate scientific content, analyse data, or formulate the manuscript's conclusions.

References

1. Lim ECN, Lim CED. Responsible use of AI in healthcare: An Australian perspective on promise, perils, and professional duties. *AI Ethics*. 2026;6:10. [doi](#)
2. Ethics and governance of artificial intelligence for health: Guidance on large multi-modal models. World Health Organization. 2025. <https://www.who.int/publications/item/9789240084759>.
3. Lim ECN, Lim CED. Artificial intelligence use in primary care: Attitudes, concerns, and readiness among health professionals in metropolitan Australia. *Journal of Computer and Communications*. 2025;13:63-71. [doi](#)
4. Singhal K, Azizi S, Tu T, Mahdavi SS, Wei J, Chung HW, Scales N, et al. Large language models encode clinical knowledge. *Nature*. 2023;620(7972):172-180. [doi](#) [pubmed](#)
5. Cabral S, Restrepo D, Kanjee Z, Wilson P, Crowe B, Abdunnour RE, Rodman A. Clinical reasoning of a generative artificial intelligence model compared with physicians. *JAMA Intern Med*. 2024;184(5):581-583. [doi](#) [pubmed](#)
6. Goh E, Gallo R, Hom J, Strong E, Weng Y, Kerman H, Cool JA, et al. Large language model influence on diagnostic reasoning: a randomized clinical trial. *JAMA Netw Open*. 2024;7(10):e2440969. [doi](#) [pubmed](#)
7. Brodeur PG, Buckley TA, Kanjee Z, Goh E, Ling EB, Jain P, Cabral S, et al. Performance of a large language model on the reasoning tasks of a physician. *Science*. 2026;392(6797):524-527. [doi](#) [pubmed](#)
8. Denno P, Hampshire A. Defining brain fog: A transdiagnostic narrative review. *European Psychiatry*. 2025;68(S1):S602-S603. [doi](#)
9. Graham EL, Clark JR, Orban ZS, Lim PH, Szymanski AL, Taylor C, DiBiase RM, et al. Persistent neurologic symptoms and cognitive dysfunction in non-hospitalized Covid-19 "long haulers". *Ann Clin Transl Neurol*. 2021;8(5):1073-1085. [doi](#) [pubmed](#)
10. Bedard J, Kropp M, Hsu M, Karaman OT, Hawes J, Rosen Kellerman G. When using AI leads to "brain fry." *Harvard Business Review*. 2026. <https://hbr.org/2026/03/when-using-ai-leads-to-brain-fry>.
11. Wang P, Liu H, Zou L, Paas F. Overloaded minds and machines: A cognitive load framework for human-AI symbiosis. *Artificial Intelligence Review*. 2026;59:109. [doi](#)
12. Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ*. 2005;331(7524):1064-1065. [doi](#) [pubmed](#)
13. Harris E, Santhosh L. Dual process theory and cognitive load: how intensivists make diagnoses. *Crit Care Clin*. 2022;38(1):27-36. [doi](#) [pubmed](#)
14. Lim ECN, Cheng NCL, Lim CED. The art of medical synthesis: Where Chinese medical wisdom intersects with artificial intelligence. *Journal of Traditional Chinese Medical Sciences*. 2026;13(1):51-59. [doi](#)
15. Endsley MR, Kiris EO. The out-of-the-loop performance problem and level of control in automation. *Human Factors*. 1995;37(2):381-394. [doi](#)
16. Sparrow B, Liu J, Wegner DM. Google effects on memory: cognitive consequences of having information at our fingertips. *Science*. 2011;333(6043):776-778. [doi](#) [pubmed](#)
17. Risko EF, Gilbert SJ. Cognitive offloading. *Trends Cogn Sci*. 2016;20(9):676-688. [doi](#) [pubmed](#)
18. Lee HP, Sarkar A, Tankelevitch L, Drosos I, Rintel S, Banks R, Wilson N. The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery. 2025;Article 1121:1-22. [doi](#)
19. Lim ECN, Lim CED. Uncertain hands: Ethical implications of AI dependency in medical education. *AI Ethics*. 2026;6:219. [doi](#)
20. Olakotan OO, Mohd Yusof M. The appropriateness of clinical decision support systems alerts in supporting clinical workflows: a systematic review. *Health Informatics J*. 2021;27(2):14604582211007536. [doi](#) [pubmed](#)
21. Khera R, Simon MA, Ross JS. Automation bias and assistive AI: risk of harm from AI-driven clinical decision support. *JAMA*. 2023;330(23):2255-2257. [doi](#) [pubmed](#)
22. Jabbour S, Fouhey D, Shepard S, Valley TS, Kazerooni EA, Banovic N, Wiens J, et al. Measuring the impact of AI in the diagnosis of hospitalized patients: a randomized clinical vignette survey study. *JAMA*. 2023;330(23):2275-2284. [doi](#) [pubmed](#)
23. Ma SP, Liang AS, Shah SJ, Smith M, Jeong Y, Devon-Sand A, Crowell T, et al. Ambient artificial intelligence

- scribes: utilization and impact on documentation time. *J Am Med Inform Assoc.* 2025;32(2):381-385. [doi](#) [pubmed](#)
24. Topaz M. Invisible scribes: can nurses trust ambient AI for clinical documentation? *J Contin Educ Nurs.* 2025;56(9):358-359. [doi](#) [pubmed](#)
 25. Ma J, Wen J, Qiu Y, Wang Y, Xiao Q, Liu T, Zhang D, et al. The role of artificial intelligence in shaping nursing education: a comprehensive systematic review. *Nurse Educ Pract.* 2025;84:104345. [doi](#) [pubmed](#)
 26. Lucas HC, Upperman JS, Robinson JR. A systematic review of large language models and their implications in medical education. *Med Educ.* 2024;58(11):1276-1285. [doi](#) [pubmed](#)
 27. Plackett R, Kassianos AP, Mylan S, Kambouri M, Raine R, Sheringham J. The effectiveness of using virtual patient educational tools to improve medical students' clinical reasoning skills: a systematic review. *BMC Med Educ.* 2022;22(1):365. [doi](#) [pubmed](#)
 28. Lim ECN, Lim CED. Machine learning in perioperative medicine: a comparative review of predictive, causal, and foundation model approaches in surgical data science. *AI in Clinical Medicine.* 2026;2:e27. [doi](#)
 29. Cui J, Lim ECN, Wu X, Lim CED. Phenylephrine's heterogeneous renal effects in non-cardiac surgery: a causal machine learning study. *Cureus Journal of Computer Science.* 2025;2:es44389-025-00018-2. [doi](#)
 30. Cui J, Li W, Lim ECN, Wu X, Lim CED. Stratified causal inference for intensive care unit risk prediction: informatics-based modeling of anesthetic drug combinations. *JMIR Form Res.* 2026;10:e80294. [doi](#) [pubmed](#)
 31. Cui J, He Z, Lim ECN, Lim CED. Precision vs recall in surgical AI: Performance and limitations of tabular foundation model for perioperative classification and risk prediction. *Digit Health.* 2026;12:20552076261443585. [doi](#) [pubmed](#)
 32. Lim ECN, Lim CED. Redefining surgical health economics: The potential of TabPFN for real-time precision modelling. *Journal of Computer and Communications.* 2025;13:30-40. [doi](#)
 33. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, Hinder S, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *J Med Internet Res.* 2017;19(11):e367. [doi](#) [pubmed](#)
 34. Hart SG, Staveland LE. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In: Hancock PA, Meshkati N. (Eds.) *Human mental workload.* North-Holland. 1988. p. 139-183. [doi](#)