Clinical Reasoning and Medical Decision-Making: Science, Art, and Challenges in the Era of Artificial Intelligence

Moshe Y. Flugelman MD[1,2] and Arieh Riskin MD PhD[2,3]

[1]Department of Cardiovascular Medicine, Carmel Medical Center, Haifa, Israel
[2]Rappaport Faculty of Medicine, Technion–Israel Institute of Technology, Haifa, Israel
[3]Department of Neonatology, Bnai Zion Medical Center, Haifa, Israel

ABSTRACT Clinical reasoning, the cognitive process by which physicians gather, interpret, and synthesize information to reach diagnostic and therapeutic decisions, is the cornerstone of medical practice. However, it is rarely taught explicitly in medical curricula and the consequences of this omission are substantial. In this narrative review, we integrate theoretical models of clinical cognition, the history and limitations of evidence-based guidelines, and emerging data on artificial intelligence (AI)-assisted clinical reasoning. Clinical reasoning operates through dual cognitive pathways: rapid, pattern-based System 1 and deliberate, analytical System 2 thinking. This thinking is vulnerable to well-characterized cognitive biases responsible for most diagnostic errors. Guidelines, while representing a systematic approach to codifying evidence, carry inherent limitations when applied to individual patients. Large language models (LLMs) demonstrate diagnostic capabilities comparable to experienced clinicians in structured settings; however, they introduce new risks including automation bias and hallucination. The medical interview remains the foundational act of clinical reasoning and the primary vehicle for building the therapeutic trust that underpins all clinical decisions. In the era of AI, strong clinical reasoning skills are not diminished in importance; they are indispensable. Situation awareness, the integration of perception, comprehension, and projection, is the cognitive capacity that distinguishes effective clinical decision-making, and it is a capacity that neither algorithms nor guidelines can replace. Formal, longitudinal education in clinical reasoning must become a core component of medical training at all levels.

IMAJ 2026; 28: 724–729

KEY WORDS: artificial intelligence (AI), clinical reasoning, cognitive bias, history taking, medical decision-making

Clinical reasoning, the process through which physicians gather and interpret information to reach diagnostic and therapeutic decisions, is the intellectual core of medicine, yet it is rarely taught explicitly. The assumption that reasoning competence emerges naturally through clinical experience is not supported by evidence.

In a national survey of internal medicine clerkship directors, many students were perceived to begin clinical training with limited knowledge of core clinical reasoning concepts [1]. Diagnostic errors contribute up to 15% of adverse events, and the majority are attributable to cognitive rather than knowledge failures [2,3].

The introduction of artificial intelligence (AI)-based diagnostic tools has intensified this challenge: evaluating algorithmically generated or model-generated recommendations calls for the same metacognitive and analytical skills that clinical reasoning training develops [4,5].

In this review, we examined the theoretical foundations of clinical reasoning, the role and limitations of clinical practice guidelines, the emerging influence of AI, and the irreplaceable importance of the patient encounter as the locus of trust and clinical judgment. This review carries the theoretical account of dual-process cognition on which the other reviews in this issue of the *Israel Medical Association Journal (IMAJ)* build.

THEORETICAL MODEL AND CLINICAL REASONING

The hypothetico-deductive method

Elstein and colleagues [6] established that clinicians employ a hypothetico-deductive strategy, generating a small set of hypotheses early and systematically seeking evidence to confirm or refute them. The distinction between expert and novice lies not in the process itself but in the quality of hypotheses generated at the outset: the expert

identifies the most probable and serious diagnoses more rapidly and precisely.

Illness scripts

Complementing the hypothetico-deductive framework, illness script theory proposes that experienced clinicians store structured mental representations of diseases in long-term memory [7]. Each script encodes enabling conditions, pathophysiological faults, and clinical consequences. As experience accumulates, scripts become richer, enabling faster and more accurate pattern recognition. Deliberate construction of illness scripts during training can accelerate the development of clinical expertise.

Dual process theory

The most influential contemporary model of clinical cognition is dual process theory, which distinguishes between two fundamentally different modes of thinking [8-10]. System 1 is fast, is automatic, and requires minimal cognitive effort. It operates through pattern recognition and the rapid retrieval of illness scripts and is the dominant mode in routine clinical encounters. System 2 is slow, deliberate, and analytically demanding. It involves the explicit, stepwise application of medical knowledge and is engaged when cases are complex, atypical, or diagnostically uncertain.

CLINICAL REASONING IS THE INTELLECTUAL CORE OF MEDICINE BUT IS RARELY TAUGHT EXPLICITLY.

Neither system is inherently superior. Cognitive errors arise not from System 1 per se but from its deployment in circumstances requiring System 2 analysis [8]. Expert clinicians are distinguished by their metacognitive capacity: monitoring their own reasoning and shifting deliberately to analytical thinking when intuition is likely to mislead, a capacity that can be taught [11]. This framework is the common theoretical foundation of the present special issue: the mapping of specific heuristics onto specific biases is developed in the review of cognitive biases and heuristics [12], and the monitoring function that governs the switch between the two systems is analyzed in the review of metacognitive effects [13].

Cognitive biases

Analysis of diagnostic errors consistently implicates cognitive biases as a primary causative factor. Graber and colleagues [2] demonstrated that cognitive failures were present in 74% of a series of 100 internal medicine diagnostic errors, and that these errors occurred predominantly in common, not exotic, clinical presentations. The most clinically significant errors are anchoring bias, premature closure, availability bias, confirmation bias, and representativeness bias. Each is defined, mechanistically explained, and illustrated in the dedicated review of cognitive biases and heuristics in this issue of *IMAJ* [12], and is therefore not re-derived here.

Recognition of these biases is necessary but insufficient. Formal instructional strategies including structured case-based reflection, *think aloud* exercises, and explicit metacognitive training are required to reduce their clinical impact [11,14,15].

MEDICAL DECISION-MAKING

The probabilistic nature of clinical decisions

A fundamental principle of medical decision-making is that clinical certainty is rarely achievable. Clinicians operate under conditions of uncertainty and work with incomplete information and probabilistic reasoning. The goal is not the elimination of uncertainty, but rather its skilled management [16].

Bayes' theorem provides the formal framework for updating diagnostic probabilities as new information is acquired. The post-test probability of a diagnosis is a joint function of pre-test probability and the likelihood ratio of the test. The same result carries profoundly different interpretive weight depending on clinical context.

The threshold model of decision-making

Pauker and Kassirer's threshold model provides a practical framework for clinical decisions under uncertainty [17]. The model defines two thresholds: a treatment threshold above which the physician proceeds to treat without further diagnostic workup; and a testing threshold below which the diagnosis is considered sufficiently unlikely to be excluded. Between these thresholds lies the zone of diagnostic uncertainty in which further investigation is warranted. The precise location of these thresholds is not fixed but varies as a function of the clinical stakes, specifically the probability of the diagnosis, the benefit of effective treatment, the harm of unnecessary treatment, and the risk of the diagnostic test itself.

The components of a clinical decision

A rigorous clinical decision integrates four dimensions: the magnitude of benefit if treatment is given to a patient with the target condition; the magnitude of harm if given to a patient without it; the cost of withholding treatment

from a patient who does have it; and the patient's values, preferences, and life circumstances, which may profoundly alter the optimal decision for a given individual.

Patient values are the element most likely to differentiate what is optimal for a specific individual from what a guideline algorithm or AI-based model would recommend [18]. Shared decision-making, which is a structured process in which the physician presents the available options, communicates the relevant evidence, and incorporates the patient's informed preferences into a jointly constructed plan, is examined in detail in the review of the physician–patient relationship in a review in the January 2027 issue of *IMAJ*.

Decision making under pressure

Clinical decisions made under time pressure, fatigue, or cognitive overload shift the balance toward System 1 reasoning [10,19]. Many diagnostic errors represent a failure to deploy the appropriate reasoning strategy under adverse conditions rather than a failure of knowledge. Structured diagnostic pauses and explicit prompts to consider alternative diagnoses improve accuracy even with minimal additional time investment.

GUIDELINES AND ARTIFICIAL INTELLIGENCE IMPROVE CARE BUT CARRY SERIOUS LIMITATIONS, MAKING STRONG CLINICAL REASONING MORE ESSENTIAL THAN EVER.

Endsley's model of situation awareness (SA) provides a useful framework for understanding why pressure degrades clinical performance [20]. SA comprises three hierarchical levels: perception of relevant elements in the environment (Level 1), comprehension of their integrated meaning considering the clinician's goals (Level 2), and projection of the patient's likely future state (Level 3). Under time pressure, fatigue, or cognitive overload, clinicians lose access to the higher levels of SA first. Level 2 comprehension and Level 3 projection collapse, leaving only fragmented pattern recognition. This mechanism produces tunnel vision and premature closure. The clinician perceives individual data points but fails to integrate them into a coherent situational picture.

CLINICAL PRACTICE GUIDELINES: AN ESSENTIAL BUT IMPERFECT TOOL

Origins and rationale

Modern clinical practice guidelines and policy-based approaches to care were strongly shaped by David Eddy's work on practice policies and explicit decision analysis [18]. Sackett and colleagues [21] later codified the principles of evidence-based medicine. The subsequent development of the GRADE system provided a structured framework for evaluating evidence quality and recommendation strength [22].

The value of guidelines

Guidelines synthesize complex evidence into actionable recommendations, reducing cognitive burden and promoting consistency of care. They are particularly valuable in high-stakes, time-sensitive situations and, for trainees, provide a scaffold for reasoning by making the evidential basis for practice explicit.

The implementation of guidelines is associated with quality improvement by reducing practice variation and improving patient care standards. In addition, guidelines facilitate evidence-based decisions, optimize efficiency, enhance cost-effectiveness considerations, and support the integration of the latest scientific discoveries.

The limitations of guidelines in individual clinical decisions

Despite their value at the population level, clinical practice guidelines carry inherent limitations when applied to individual patients, which must be explicitly recognized [21]. First, guidelines are derived from aggregate data, typically from randomized controlled trials with well-defined inclusion and exclusion criteria. The patients recruited to these trials are therefore a selected population that may differ substantially from the patient presenting in clinical practice in age, co-morbidity burden, concomitant medications, functional status, and personal values. This gap between trial populations and real-world patients makes it extremely difficult to apply guideline recommendations without careful judgment about their individual applicability. Thus, even when a guideline is evidence-based and current, it addresses a defined and idealized patient, not the person sitting across from the clinician with their complexity of illness narrative, values, and social context.

Second, guidelines necessarily lag behind the evolving evidence base. When a recommendation is issued, it reflects the data available at the time of writing. In rapidly advancing fields, current practice may already have moved beyond what a recently published guideline recommends.

As Eddy observed, the clinician who applies a guideline without engaging in independent reasoning about its appropriateness for their patient has substituted rule-following for clinical judgment [18].

ARTIFICIAL INTELLIGENCE AND CLINICAL REASONING

Emerging capabilities

Large language models (LLMs) have demonstrated diagnostic capabilities comparable or superior to experienced physicians on validated instruments [4,5]. In a randomized trial, researchers found that clinician access to GPT-4 recommendations improved diagnostic accuracy from 47% to 65% and from 63% to 80% across two cohorts, without introducing demographic bias [23]. This result will be expanded in a review in the January 2027 issue of *IMAJ*, which will describe a second randomized trial by an overlapping group in which physicians who were given access to an LLM did not diagnose standardized cases significantly better than those using conventional resources, although the model working alone outperformed both groups. In that review, physicians were given access to an LLM. The physicians did not diagnose standardized cases significantly better than those using conventional resources, although the model working alone outperformed both groups. Taken together, the two trials indicate that the benefit of AI assistance depends on how physicians and tools are integrated rather than on model capability alone. Randomized controlled trials of AI in cardiovascular care and AI-assisted clinical practice have documented the expanding evidence base for AI-augmented medicine [24,25].

THE MEDICAL INTERVIEW REMAINS THE FOUNDATION OF CLINICAL REASONING AND THERAPEUTIC TRUST; NO ALGORITHM CAN FULLY REPLACE THE CLINICIAN'S JUDGMENT.

The paradox: greater AI capability demands stronger clinical reasoning

The diagnostic performance of AI systems has led some researchers to suggest that clinical reasoning may become less important. This view misreads the evidence. The integration of AI into clinical practice does not diminish the need for clinical reasoning; rather, it creates new and more demanding requirements for it. Just as computed tomography did not make clinical reasoning redundant but demanded a new layer of interpretive skill, AI-generated recommendations require a clinician who can assess their plausibility in context, weigh the quality of the underlying evidence, and integrate the output with the patient's full clinical picture.

Automation bias and hallucination

Two risks deserve emphasis. The first risk is automation bias, which is the tendency to defer to algorithmically or model-generated recommendations even when they are incorrect. In a randomized clinical vignette survey study, Jabbour and colleagues [26] showed that exposure to systematically biased AI predictions reduced clinician diagnostic accuracy. The second risk is hallucination. LLMs generate factually incorrect outputs that carry the surface confidence of well-reasoned analysis [4], a property of current architectures rather than incidental errors. Both phenomena, together with algorithmic bias and the propagation of fabricated evidence, are analyzed in a review of the clinical risks of artificial intelligence in this issue of *IMAJ* [27].

Both risks point in the same direction. Safe AI use in clinical medicine demands precisely the skills that clinical reasoning training provides: the capacity to interrogate a recommendation, the ability to recognize when it does not fit the patient at hand, and the understanding to maintain independent judgment in the face of an authoritative-seeming algorithmic output. Automation bias is disproportionately prevalent among less experienced clinicians who have not yet developed the clinical reasoning foundation necessary to critically evaluate algorithmic outputs [26].

THE PATIENT ENCOUNTER: HISTORY TAKING AND TRUST

The medical interview as a diagnostic instrument

The medical interview remains the most powerful diagnostic instrument available to the clinician [28,29]. History taking, when conducted attentively, yields the diagnosis in most clinical encounters without recourse to laboratory or imaging data, a finding that still holds in the contemporary environment. Angina pectoris, major depression, chronic obstructive pulmonary disease, and many prevalent conditions are diagnosed primarily through the clinical narrative. Despite this situation, the medical interview is systematically undervalued in clinical training and practice. Time pressure, reliance on investigation, and the absence of formal instructional frameworks have resulted in a widespread tendency to abbreviate or bypass careful history taking. The time pressure so frequently cited as justification for a truncated interview is, in most instances, artificial. Studies consistently show that a skilled, focused history does not require more time than a fragmented one. What is lacking is not time but technique, the ability to prioritize the clinically relevant narrative and a professional attitude that regards attentive listening as a core clinical act rather than an optional courtesy. The observation that the average clinician interrupts the patient within seconds of the opening

of the interview [28] is not merely a communication failure, it is a diagnostic failure since the information that the interrupted patient would have volunteered may be precisely the information required for an accurate diagnosis.

Trust as a clinical instrument

Data collection, understanding the patient as a person, and the creation of trust are not separable goals of the medical interview. They are achieved together or not at all [28,29]. Trust is a clinical instrument, not merely a relational aspiration. Clinical evidence demonstrates that trust in a referring physician reduces patient anxiety during community-to-hospital transitions with effects that persist in the trusted physician's absence [29]. Decisions made without this trust carry an elevated risk of being inappropriate or futile. Building trust requires specific, learnable skills, including showing genuine interest in the patient as a person, which is expressed through purposeful inquiry into their life and values, open questioning, and deliberate leveling of the asymmetric power relationship inherent in the clinical encounter [28]. These soft skills are not ancillary to clinical reasoning; they are core components of the information-gathering process on which accurate diagnosis depends.

The case for formal education in clinical reasoning

Consistent findings that clinical reasoning is poorly taught yet important show the need for fundamental reforms in medical education. The Clinical Reasoning in Medical Education (CReME) consensus group [14] identified 30 core topics across five domains, from foundational concepts (metacognition, illness scripts, dual process thinking, cognitive bias) to history taking, diagnostic test interpretation, problem representation, and shared decision-making, as the required content of a formal undergraduate curriculum.

Effective instructional approaches, identified through systematic review, include case-based learning with structured feedback, illness script construction exercises, *think aloud* methods in which instructors make their reasoning process explicit and observable, and metacognitive reflection using structured debriefing following clinical encounters [11,14]. Generic rule-based instruction without clinical content specificity has not been shown to produce meaningful improvements. This education must extend beyond the undergraduate years. Postgraduate programs should embed supervised AI use within a framework of critical appraisal. Clinical teachers require specific training in making their reasoning processes explicit [1].

DISCUSSION

Clinical reasoning occupies a paradoxical position in contemporary medicine: universally recognized as the defining intellectual activity of the physician, yet rarely the subject of systematic instruction. The practical consequences are well documented. Diagnostic errors account for up to 15% of adverse events in healthcare systems, and their principal cause is consistently attributed to cognitive rather than knowledge failure [2,3].

Two developments raise the stakes. First, evidence-based guidelines have reduced unwarranted practice variation and made the evidential basis of recommendations explicit, but their value depends on clinicians who can judge when a population-level recommendation applies to a particular patient, a judgment the guidelines cannot make [21]. Second, LLM-based tools show real diagnostic capability, but automation bias, hallucination, and algorithmic discrimination are documented hazards [27]. Clinicians who can use AI without deferring to it uncritically have mastered the foundations of clinical reasoning.

CONCLUSIONS

Clinical reasoning and medical decision-making are learnable skills, not capacities that emerge automatically from clinical exposure. The theoretical framework we reviewed, including hypothetico-deductive reasoning, illness scripts, dual process theory, situation awareness, cognitive bias, and probabilistic decision-making, is unified. Each construct points to the same practical conclusion: the skilled clinician is distinguished not by knowledge alone but by the capacity to maintain integrated situation awareness across all levels of complexity, to recognize when that awareness is incomplete or failing, and to act on that recognition. Guidelines and AI tools are valuable adjuncts, but neither can substitute for this capacity. Guidelines cannot apply themselves to individual patients and AI cannot critically appraise its own output. The medical interview, and the therapeutic trust it creates, remains the irreplaceable foundation of clinical medicine: a diagnostic, ethical, and relational achievement that no algorithm can replicate.

Correspondence

Prof. M.Y. Flugelman
Dept. of Cardiovascular Medicine, Carmel Medical Center, Haifa 34362, Israel
Phone: (972-4) 825-0575
Fax: (972-4) 825-0936
Email: myf@technion.ac.il

References

1. Rencic J, Trowbridge RL Jr, Fagan M, Szauter K, Durning S. Clinical reasoning education at US medical schools: results from a national survey of internal medicine clerkship directors. *J Gen Intern Med* 2017; 32 (11): 1242-6.
2. Graber ML, Franklin N, Gordon R. Diagnostic error in internal medicine. *Arch Intern Med* 2005; 165 (13): 1493-9.
3. Singh H, Meyer AN, Thomas EJ. The frequency of diagnostic errors in outpatient care: estimations from three large observational studies involving US adult populations. *BMJ Qual Saf* 2014; 23 (9): 727-31.
4. Cabral S, Restrepo D, Kanjee Z, et al. Clinical reasoning of a generative artificial intelligence model compared with physicians. *JAMA Intern Med* 2024; 184 (5): 581-3.
5. Eriksen AV, Moller S, Ryg J. Use of GPT-4 to diagnose complex clinical cases. *N Engl J Med AI* 2024; 1 (1): AIoa2300031.
6. Elstein AS, Shulman LS, Sprafka SA. Medical Problem Solving: An Analysis of Clinical Reasoning. Cambridge, MA: Harvard University Press, 1978.
7. Charlin B, Boshuizen HP, Custers EJ, Feltovich PJ. Scripts and clinical reasoning. *Med Educ* 2007; 41 (12): 1178-84.
8. Norman GR, Monteiro SD, Sherbino J, et al. The causes of errors in clinical reasoning: cognitive biases, knowledge deficits, and dual process thinking. *Acad Med* 2017; 92 (1): 23-30.
9. Croskerry P. A universal model of diagnostic reasoning. *Acad Med* 2009; 84 (8): 1022-8.
10. Kahneman D. Thinking, Fast and Slow. New York: Farrar, Straus and Giroux, 2011.
11. Mamede S, Schmidt HG, Penaforte JC. Effects of reflective practice on the accuracy of medical diagnoses. *Med Educ* 2008; 42 (5): 468-75.
12. Mintz A, Riskin A. Cognitive biases and heuristics in medical decision-making: mechanisms, clinical manifestations, and debiasing strategies. *IMAJ* 2026; 28 (11): 690-700.
13. Riskin Y, Riskin A. Metacognitive effects on clinical decision-making. *IMAJ* 2026; 28 (11): 701-8.
14. Cooper N, Bartlett M, Gay S, et al. Consensus statement on the content of clinical reasoning curricula in undergraduate medical education. *Med Teach* 2021; 43 (2): 152-9.
15. Croskerry P. The importance of cognitive errors in diagnosis and strategies to minimize them. *Acad Med* 2003; 78 (8): 775-80.
16. Sox HC, Blatt MA, Higgins MC, Marton KI. Medical Decision Making. Stoneham, MA: Butterworths, 1988.
17. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. *N Engl J Med* 1980; 302 (20): 1109-17.
18. Eddy DM. Practice policies: where do they come from? *JAMA* 1990; 263 (9): 1265-75.
19. Croskerry P. Achieving quality in clinical decision making: cognitive strategies and detection of bias. *Acad Emerg Med* 2002; 9 (11): 1184-204.
20. Endsley MR. Toward a theory of situation awareness in dynamic systems. *Hum Factors* 1995; 37 (1): 32-64.
21. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ* 1996; 312 (7023): 71-2.
22. Guyatt GH, Oxman AD, Vist GE, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008; 336 (7650): 924-6.
23. Goh E, Bunning B, Khoong EC, et al. Physician clinical decision modification and bias assessment in a randomized controlled trial of AI assistance. *Commun Med (Lond)* 2025; 5 (1): 59.
24. Hadida Barzilai D, Sudri K, Goshen G, et al. Randomized controlled trials evaluating artificial intelligence in cardiovascular care: a systematic review. *JACC Adv* 2025; 4 (11 Pt 1): 102152.
25. Tessler I, Wolfovitz A, Livneh N, et al. Advancing medical practice with artificial intelligence: ChatGPT in healthcare. *IMAJ* 2024; 26 (2): 80-5.
26. Jabbour S, Fouhey D, Shepard S, et al. Measuring the impact of AI in the diagnosis of hospitalized patients: a randomized clinical vignette survey study. *JAMA* 2023; 330 (23): 2275-84.
27. Glebov M, Mintz A. Clinical risks of artificial intelligence in medicine. *IMAJ* 2026; 28 (11): 717-23.
28. Flugelman MY. History-taking revisited: simple techniques to foster patient collaboration, improve data attainment, and establish trust with the patient. *GMS J Med Educ* 2021; 38 (6): Doc109.
29. Flugelman MY, Jaffe R, Luria G, Yagil D. Trust in the referring physician reduces anxiety in an integrated community-to-hospital care system. *Isr J Health Policy Res* 2020; 9 (1): 7.

It is the ability to take a joke, not make one, that proves you have a sense of humor.

Max Eastman (1883–1969), American writer, poet, and prominent political activist

Capsule

Sugarcoating the fate of γδ T cells

γδ T cells play an important role in maintaining tissue homeostasis but become dysregulated in inflammatory conditions. **Santos-Pereira** and co-authors revealed that the γδ T cell glycome plays an important role in γδ T17 cell-fate commitment, promoting a hyperactivated proinflammatory state that exacerbates immune-mediated inflammatory diseases. Analysis of human transcriptomic data and murine models revealed that the absence of β1,6 *N*-acetylglucosamine–branched *N*-glycans promoted type-3 interleukin-17A–producing γδ T cells and loss of tolerance. Loss of *N*-glycan branching in glycoengineered mouse models increased susceptibility to psoriasis, lupus, and inflammatory bowel disease. Supplementing γδ T cells with *N*-acetylglucosamine reduced disease pathology and γδ T17 cell activation, revealing the importance of glycobiology in γδ T cell behavior and strategies for disease intervention.

Sci Immunol 2026; 11 (124): eaec9553

Eitan Israeli