

This article may be reprinted free of charge provided 1) that there is clear attribution to the Orthomolecular Medicine News Service, and 2) that both the OMNS free subscription link <http://orthomolecular.org/subscribe.html> and also the OMNS archive link <http://orthomolecular.org/resources/omns/index.shtml> are included.

FOR IMMEDIATE RELEASE

Orthomolecular Medicine News Service, October 12, 2026



The Medical AI Bias Trap: Are ChatGPT and Other AI Systems Making Us Smarter - or Just More Confidently Misled?

More data cannot cure flawed medical assumptions. Why evidence fit, root-cause reasoning, and human conscience must shape the next generation of medical AI.

Richard Z. Cheng, M.D., Ph.D.

Orthomolecular Medicine News Service | Scientific commentary

Still asking ChatGPT, Doubao (in China), or another general-purpose AI what to do about your health? Here is the uncomfortable question: What if you are being fed the blind spots of modern medicine-and do not even know it? The answer may arrive in seconds, cite studies, and sound reassuringly authoritative. None of that makes it the whole truth.

AI can make a medical answer sound definitive even when the underlying question is poorly framed. Confidence is not truth; a citation is not a complete causal explanation. If the literature AI relies on reflects selective funding, institutional habits, under-studied interventions, or inappropriate assumptions, the resulting response can carry those weaknesses into the consultation room-at machine speed.

Is this literal "brainwashing"? No. AI is not necessarily trying to manipulate anyone. The risk is subtler: a polished answer can make inherited assumptions feel like settled fact. Nor does identifying this risk mean that modern medicine is a conspiracy, that pharmaceuticals are ineffective, or that under-studied alternatives must work. It means users can internalize an inherited bias precisely because an answer appears comprehensive, neutral, and evidence-based. An AI recommendation must be examined, not worshipped.

Our September 2026 preprint, *Beyond Evidence Hierarchies: An Intervention-Aware, Contextual Systems Framework for Medical Artificial Intelligence*, addresses a foundational part of this challenge: What model of medicine is AI learning-and is that model adequate for the biological question and the individual patient? [1]

1. Who shapes the medical knowledge AI learns?

Before asking whether an AI answer is evidence-based, ask a prior question: who got to produce the evidence? General-purpose AI draws on publications, guidelines, textbooks, clinical trials, and other human-produced material. These sources are not created outside society. Research agendas reflect funding availability, institutional priorities, commercial incentives, publication practices, and public-health needs. Such influences can affect which questions receive attention and which remain underinvestigated.

Money does not automatically corrupt a result. Yet financial incentives can shape which questions are asked in the first place. A commercially promising intervention may attract extensive investment and large trials. By comparison, preventive strategies, nutrition, lifestyle change, or combinations of nonpatentable interventions may face different funding incentives and methodological challenges. This imbalance is a concern worth investigating-not proof that pharmaceutical studies are unreliable or that underfunded interventions are effective.

The number of papers is not a referendum on truth. More publications do not automatically mean less bias. Fewer publications do not automatically mean no benefit. But a smaller evidence base also cannot be treated as positive proof of an intervention's efficacy. Absence of evidence, evidence of no clinically meaningful effect, and evidence generated by a poorly matched trial are distinct categories.

Medical AI needs to recognize these distinctions. It should identify conflicts of interest and limitations in the evidence where relevant, without assuming bad faith or drawing conclusions unsupported by the data.

2. More data can mean more authority-not necessarily more truth

Here is the paradox: the more fluent medical AI becomes, the harder its blind spots may be to notice. AI may inherit the blind spots of medical literature and its conceptual vocabulary. Training on more publications does not automatically repair gaps in research funding, causal models, or clinical representation. A fluent answer can collapse uncertainty, omit a neglected mechanism, or present a population-average recommendation as if it were individually definitive.

This is not the same as saying that AI cannot correct biases. Careful retrieval, comparative evidence appraisal, independent evaluation, and transparent reasoning can help. But those safeguards must be deliberately designed and tested. **Scaling a model is not a substitute for questioning its assumptions.**

The underlying concern is simple: an AI trained within an incomplete framework may become extraordinarily good at practicing an incomplete model of medicine. The solution requires a better model of medical reasoning, not merely a faster answer engine.

3. A diagnosis is not a causal explanation: consider type 2 diabetes

Consider the difference between naming a disease and understanding a person. Type 2 diabetes is a useful example. Diagnosis identifies a clinically important state. Evidence-based medications can lower glucose and, in appropriate patients, reduce important complications. Yet naming the disease does not by itself explain why this particular patient developed it.

The clinician may need to explore insulin resistance, visceral adiposity, dietary patterns, energy balance, physical activity and skeletal muscle, sleep, medications, genetics, stress, and other relevant factors. Their relative importance varies between patients. Not every proposed contributor will be causal or modifiable in any given case.

A disease label is necessary for many purposes. **It is not a complete model of disease causation.** AI that merely maps "type 2 diabetes" to a guideline may miss clinically relevant trajectories and modifiable processes. Conversely, root-cause reasoning must never become an excuse to discontinue established treatment or substitute a plausible mechanism for demonstrated benefit.

4. Evidence hierarchy is not the same as evidence fit

We should defend rigorous trials-not worship their labels. Randomized controlled trials and systematic reviews remain indispensable. Randomization protects against many forms of confounding, and well-conducted trials can reveal effects that mechanistic speculation cannot. But high methodological rank does not automatically establish that a study tested the appropriate biological question. [\[1-3\]](#)

Consider a nutrient trial. Unlike a new drug, an essential nutrient is already present in the body and diet. Participants enter a trial with different baseline levels, intakes, absorption, physiological requirements, and interacting nutrient exposures. If treatment does not achieve a meaningful biological separation-or the trial mostly enrolls participants who are already sufficient-a null result may not answer whether correcting deficiency is beneficial. [\[2\]](#)

This does not invalidate randomization. Achieved-status analyses can also introduce confounding, and favorable mechanistic stories cannot rescue a poorly supported intervention. It does mean investigators should assess baseline status, assigned dose, adherence, achieved exposure, outcome selection, duration, and prespecified treatment-effect heterogeneity.

Two questions must be asked together: Was the study conducted rigorously? And was it designed to test the relevant biological effect?

Nutrients, pharmaceuticals, diets, exercise, procedures, behavioral interventions, environmental changes, and multimodal programs differ in their causal and exposure structures. Rigorous evidence

appraisal must therefore consider both methodological quality and intervention-specific fit, while retaining a clear distinction between mechanistic plausibility and clinical outcome evidence.

5. From evidence retrieval to contextual systems reasoning

Our preprint proposes **Contextual Systems Reasoning for Medical AI (CSR-MAI)**-a seven-domain research framework intended to move AI beyond the oversimplified sequence "question → literature search → evidence ranking → recommendation." [1]

1. Patient context and goals: individual circumstances, preferences, values, and constraints.
2. Phenotype and trajectory: what is happening clinically, and how it is changing.
3. Physiological process mapping: which biological processes plausibly explain the observations.
4. Causal and contributing-driver analysis: what is established, likely, plausible, or speculative.
5. Intervention ontology and evidence-method fit: whether dose, route, timing, exposure, and study design address the mechanism and question.
6. Decision synthesis, safety, and feasibility: benefits, harms, interactions, alternatives, cost, and patient priorities.
7. Longitudinal verification and updating: what measurable response would confirm, weaken, or refute the working model.

In practical terms, AI should ask not only "What does the guideline recommend?" but also "What is driving this patient's condition? What competing explanations are plausible? Which interventions have evidence of meaningful benefit? What are the risks? And how will we know whether our model was right?"

The intended loop is **observe → hypothesize → intervene → measure → learn → update**. This is a conceptual framework and research agenda, **not a validated clinical decision system**. Prospective evaluation is essential before AI based on CSR-MAI can be relied upon in patient care. [1]

6. Intelligence is not medical conscience

A patient is not a spreadsheet, and medicine is not simply the application of a prediction score to a human body. It involves suffering, uncertainty, dignity, competing goals, and ethical responsibility. AI can summarize survival statistics but cannot legitimately decide, on its own, how a patient should weigh an additional month of survival against toxicity, caregiving burdens, treatment costs, or time at home.

Contemporary AI does not possess the human professional and moral accountability required for autonomous medical decisions. Whether AI could ever possess subjective consciousness or anything analogous to a "soul" is unsettled; this article does not claim to resolve that philosophical question.

Intelligence is not wisdom. Fluency is not compassion. And a medical algorithm-however impressive-is not a medical conscience. Physicians remain responsible for clinical judgment within their roles, and patients must retain meaningful autonomy through informed, shared decision-making.

7. Root-cause-focused medicine-without replacing one dogma with another

Integrative Orthomolecular Systems Medicine (I-OM Systems Medicine) is our proposed approach to bringing physiology, nutrition, metabolism, environmental exposures, established treatment, and longitudinal patient responses into a unified reasoning framework. It asks clinicians to look beyond a disease label toward the interacting upstream contributors and downstream biological processes that may shape illness.

Those contributors may include diet, metabolic dysfunction, nutrient inadequacy, inactivity, sleep, stress, oral pathology, hormonal disturbance, environmental exposures, medication effects, and genetic susceptibility. They may converge on inflammation, redox imbalance, metabolic inflexibility, mitochondrial or microvascular dysfunction, immune dysregulation, and impaired repair. These are candidate mechanisms, not universal causes proven in every patient.

For orthomolecular interventions, biological definition matters. For example, dietary vitamin C, oral supplementation, and high-dose intravenous vitamin C are not interchangeable: their exposure, pharmacokinetics, indications, evidence, and safety considerations differ. Treating all three as a single undifferentiated "vitamin C intervention" impairs scientific interpretation.

The goal is not to replace conventional dogma with I-OM dogma. We should challenge pharmaceutical, nutritional, lifestyle, and I-OM claims by the same scientific principles. Established therapies deserve recognition when supported by evidence. Root-cause hypotheses deserve testing-not automatic acceptance. Conflicts of interest, including our own institutional interests, should be disclosed.

We are developing I-OM.ai as an AI-enabled clinical reasoning and education platform inspired by this approach. It is under development; its conceptual aspirations must not be confused with demonstrated clinical effectiveness or regulatory authorization for autonomous medical practice.

8. The real danger-and the real opportunity

We could use AI simply to automate existing medicine: summarize records faster, retrieve guidelines faster, draft notes and support prescribing decisions faster, and match diagnostic labels to treatments faster. Those are useful applications. But **faster repetition of an incomplete model is not necessarily progress.**

The more ambitious task is to build AI that can recognize the limitations of its information, expose uncertainty, compare competing causal explanations, appraise evidence with appropriate biological context, and update its reasoning when patient outcomes disagree with predictions.

Medical AI should help physicians think more rigorously and patients ask better questions-not encourage either group to accept an eloquent answer as unquestionable truth.

The most dangerous AI medical answer may not be an obvious hallucination. It may be the beautifully written, citation-rich answer that leaves out the question that matters most. An incomplete or biased answer can be delivered so convincingly that no one thinks to challenge it.

Medicine needs more than bigger AI. It needs better science, clearer causal reasoning, accountable clinicians, informed patients, and a deeper commitment to human dignity. The future of medical AI should not merely automate what medicine already does. It should help medicine reason better-and care better.

References

1. Cheng RZ, Cheng EK. Beyond Evidence Hierarchies: An Intervention-Aware, Contextual Systems Framework for Medical Artificial Intelligence. Preprints. 2026;2026091652. doi: [10.20944/preprints202609.1652.v1](https://doi.org/10.20944/preprints202609.1652.v1).
2. Heaney RP. Guidelines for optimizing design and analysis of clinical studies of nutrient effects. Nutr Rev. 2014;72(1):48-54. doi: [10.1111/nure.12090](https://doi.org/10.1111/nure.12090).
3. Schwingshackl L, et al. Most methodological characteristics do not exaggerate effect estimates in nutrition RCTs: findings from a meta-epidemiological study. J Clin Epidemiol. 2026;189:112038. doi: [10.1016/j.jclinepi.2025.112038](https://doi.org/10.1016/j.jclinepi.2025.112038).

About the author and disclosure

Richard Z. Cheng, MD, PhD, is Editor-in-Chief of the Orthomolecular Medicine News Service (OMNS) and works in integrative orthomolecular and systems medicine. The source preprint was coauthored with Eric K. Cheng, who works in artificial intelligence and is co-developing the medical-AI reasoning framework described in that paper.

Disclosure: Richard Z. Cheng is associated with I-OM.ai, an initiative dedicated to developing medical AI grounded in **Integrative Orthomolecular Systems Medicine (I-OM Systems Medicine)**. Its mission is to advance scientifically rigorous, root-cause-focused, patient-centered medicine by integrating orthomolecular nutrition, metabolic health, systems biology, lifestyle interventions, and conventional medical knowledge. I-OM.ai aims to address the potential biases and limitations of

general-purpose medical AI through independent evidence appraisal, intervention-aware causal reasoning, individualized biological assessment, and continuous evaluation of clinical outcomes. The goal is not merely to reproduce existing medical guidelines, but to develop AI that critically evaluates evidence, identifies underlying drivers of disease, and prioritizes patient safety, long-term health, and human well-being. This approach remains under development and requires independent clinical validation.

The underlying manuscript was posted on September 20, 2026, as a non-peer-reviewed preprint. This commentary extends its conceptual analysis to institutional incentives, potential knowledge-base bias, and the ethics of medical AI; these extensions are editorial interpretations, not empirical findings established by the preprint. This article is for education and does not provide individualized medical advice or validate any AI system for independent clinical decision-making.