Dean&Francis
ISSN 2959-409X

The Double-Edged Sword of AI Medical Consultation Tools: Navigating Between Empowerment and Exacerbated Health Anxiety

Emily Yu

Abstract:

The integration of Artificial Intelligence (AI) into healthcare, particularly through AI-based medical consultation tools (AIMCTs) like symptom checkers and health chatbots, represents a significant shift in patient engagement. These tools promise to democratize access to health information, reduce clinician burden, and empower users. However, this convenience carries a profound psychological risk: the potentiation of health anxiety. This paper explores the dual-edged nature of AIMCTs, analyzing how they alleviate uncertainty while simultaneously acting as catalysts for health anxiety, or "cyberchondria." By synthesizing evidence on their benefits, the mechanisms of anxiety exacerbation including risk amplification, AI hallucinations, and the patterns identified in cyberchondria research, evidence-based approach to the design and deployment of AIMCTs to mitigate adverse effects and ensure they serve as responsible aids in healthcare.

Keywords: AI Medical Consultation Tools; Health Anxiety; Cyberchondria; Risk Amplification; Empathic Design

1. Introduction

The global healthcare sector is increasingly adopting AI to address challenges like provider shortages and rising costs. AI-based medical consultation tools, including virtual health assistants and diagnostic algorithms, are at the forefront of this transformation[1]. Concurrently, health anxiety, a condition characterized by persistent worry about having or acquiring a serious illness, remains a prevalent issue. The central thesis of this article is that AIMCTs wield a "double-edged sword." On one side, they can empower users and reduce uncertainty; on the other, they can significantly amplify health anxiety, a phenomenon well-documented in the context of online health searches[2]. This paper seeks to answer the critical question: Through what specific mechanisms do AI consultations risk exacerbating health anxiety, and what does the convergence of AI-specific risks and established cyberchondria patterns mean for future

healthcare delivery and technology design?

2. The Promise: How AI Tools Alleviate Uncertainty

AIMCTs offer several key benefits that can positively impact user psychology and healthcare access.

A primary benefit of Artificial Intelligence Medical Consultation Tools (AIMCTs) is their ability to provide immediate, 24/7 access to preliminary health assessments. This instantaneity directly addresses a significant source of patient distress: the "waiting room anxiety" that arises in the interval between symptom onset and a scheduled medical appointment. During this period, uncertainty can fester, often leading individuals to imagine worst-case scenarios.

The value of this prompt reassurance is most evident in managing common, benign symptoms. For instance, a user experiencing a generalized, non-throbbing headache, devoid of "red flag" symptoms like sudden onset, weakness, or visual disturbances, can use an AIMCT for a rapid assessment. A well-designed tool, grounded in clinical guidelines, can effectively triage this presentation. It would likely provide information on the high prevalence of tension-type headaches, offer common lifestyle triggers (such as stress, dehydration, or eye strain), and suggest initial self-care measures[3]. This process can deliver contextualized reassurance, effectively de-escalating worry and empowering the user with actionable steps, all while appropriately advising on specific signs that would warrant an in-person consultation.

Empirical studies support this anxiolytic effect. Research has shown that for a substantial proportion of users, particularly those with low baseline health anxiety, the use of online symptom checkers results in a significant reduction in concern and a feeling of relief[3]. By offering a structured, readily available framework for understanding common ailments, AIMCTs can prevent unnecessary escalation of health worries and alleviate the burden of uncertainty, making them a valuable first step in the healthcare journey for many individuals User Empowerment and Reassurance: By translating complex medical jargon into understandable language, AIMCTs can educate users about their symptoms, fostering a sense of control and engagement in their health journey[4]. Studies on low-anxiety users show they often feel relieved after online symptom checking[5].

Reducing Barriers to Healthcare For individuals in remote areas or with limited financial means, these tools offer a vital first point of contact with the healthcare system, making preliminary consultation more accessible[6].

3. The Peril: How AI Consultation Exacerbates Health Anxiety

Despite their benefits, AIMCTs can inadvertently fuel cyberchondria, a pattern of excessive and repetitive online health information seeking that is associated with escalating health anxiety[7]. This is the digital-age equivalent of "Baidu diagnosis," where searching symptoms leads to conclusions of severe illness. The foundational research by Doherty-Torstrick et al. demonstrated that individuals with high health anxiety recall feeling "worse" both during and after online symptom searching, whereas those with low anxiety recall relief[2]. AIMCTs, with their advanced capabilities, risk automating and intensifying this cycle. Users input common symptoms, and the AI, aiming for comprehensiveness, generates a list of potential causes that often includes rare but serious conditions. Anxious users disproportionately focus on these worst-case scenarios, initiating a cycle of fear and repeated searching that mirrors the "non-reassurability" seen in cognitive-behavioral models of health anxiety[2].

4. Mechanisms Driving Increased Anxiety

The exacerbation of health anxiety by AIMCTs is not a random outcome but a direct consequence of their specific operational features. These inherent design characteristics of current systems actively align with and amplify well-established psychological mechanisms of anxiety. Specifically, the architecture and output of AIMCTs interact with user cognition in ways that can systematically trigger and intensify distress. This section delineates the primary pathways through which this occurs: first, through the amplification of risk perception by neglecting base rates and fuelling catastrophic associations; second, through the induction of a maladaptive reassurance-seeking loop that mirrors compulsive behaviors; and third, through the imposition of a significant cognitive burden generated by probabilistic uncertainty and information overload. Together, these mechanisms explain how tools designed to inform can, under certain conditions, paradoxically function as engines of anxiety.

Risk Amplification and Worst-Case Scenario Association: The fundamental operating principle of many AIMCTs to generate a comprehensive list of differential diagnoses, contains an inherent flaw when viewed through the lens of user psychology. While clinically thorough, this approach often fails to adequately weight possibilities by their real-world prevalence, a critical factor for accurate risk assessment. Consequently, a rare but severe condition like

ISSN 2959-409X

multiple sclerosis might be presented alongside a highly common ailment like a tension-type headache with comparable prominence on a list of possibilities for a symptom like numbness[8].

This presentation style creates a fertile ground for catastrophic misinterpretation. Users, particularly those with pre-existing health anxiety, are not neutral processors of information. Their cognitive biases, such as the "availability heuristic" and "probability neglect", catastrophically draw their attention to the most severe outcomes on the list[9]. The AI's "objective" list is thus subjectively filtered through a lens of fear, transforming a remote possibility into a perceived imminent threat.

This issue is not solely a product of AI but is exacerbated by it. Research into online health information seeking has long established that search engines often prioritize or return statistically unlikely yet severe explanations for common symptoms, a phenomenon sometimes termed "cyberchondriacal escalation"[10]. AIMCTs, despite being more structured, can institutionalize this problem if their algorithms and user interfaces do not explicitly correct for it. The core of the issue lies in the widespread human tendency towards base rate neglect—the failure to incorporate prior probability into judgment[10]. When an AIMCT presents a serious condition without forcefully contextualizing its extreme rarity, it inadvertently amplifies risk perception, guiding the user away from the most probable explanation and towards a medically improbable, yet psychologically compelling, worst-case scenario.[2].

Uncertainty and Information Overload: While AIMCTs democratize access to medical information, they simultaneously present users with a profound psychological challenge rooted in inherent uncertainty and information overload. Unlike a clinician who synthesizes data to deliver a conclusive opinion, AIMCTs typically generate a list of potential conditions annotated with probabilistic estimates. For a layperson, interpreting these percentages distinguishing between a 5% and a 0.5% chance of a serious illness, is a cognitively demanding task that can lead to probability neglect, where the mere presence of a severe outcome, however unlikely, dominates their emotional response[9].

This probabilistic uncertainty is often compounded by an avalanche of complex medical information. AIMCTs may detail pathophysiological mechanisms, lists of esoteric biomarkers, and extensive differential diagnoses, far exceeding a user's capacity to process. This creates a state of cognitive overload, where the individual, lacking the framework to filter and prioritize this data, experiences heightened confusion and decisional paralysis[11]. The very tool intended to clarify becomes a source of further obscurity.

Consequently, users often engage in a compulsive cycle of information-seeking as a maladaptive coping mechanism for their anxiety. Driven by the desire to resolve the ambiguity and achieve the unattainable goal of certainty, they prolong their interaction with the tool, rephrasing queries and seeking more details. Crucially, empirical evidence confirms that longer search durations are strongly correlated with increased health anxiety[7]. The very act of prolonged searching to find a definitive answer from ambiguous data paradoxically deepens the state of distress it was meant to alleviate, trapping the user in a counterproductive feedback loop.[2].

Lack of Emotional Support and Empathic Design: Human clinicians provide emotional support through tone of voice and reassurance. Current AIMCTs typically operate with a neutral, factual tone. Delivering a message that "your symptoms could be associated with several conditions, including cancer" without any emotional buffer can directly trigger fear[12]. This lack of empathy fails to counteract the affective pressure of health information[13, 14].

The delivery of medical information is as crucial as its accuracy. Human clinicians are trained not only to diagnose but also to manage the patient's emotional response. They employ a range of non-verbal cues and verbal strategies to provide reassurance and contextualize frightening possibilities. For instance, a doctor might say, "I understand why that symptom would be worrying, but the vast majority of the time, it's caused by something much less serious. Let's work together to rule out the concerning things first." This approach validates the patient's fear while simultaneously offering hope and a collaborative path forward.

In stark contrast, current AIMCTs are fundamentally designed as information-processing systems. They typically operate with a neutral, factual, and often dispassionate tone, incapable of perceiving or responding to the user's emotional state. The delivery of a potential diagnosis like, "Your symptoms could be associated with several conditions, including migraine, tension headache, or in rare cases, a brain tumor," is presented as a simple statement of probabilities. For a user already in a state of heightened anxiety, this list is not read probabilistically but catastrophically[2]. The mention of "brain tumor" acts as an anchor, triggering an immediate fear response that the AI makes no attempt to soften or counter. This lack of an emotional buffer fails to mitigate what is known as the "affective pressure" of threatening health information, where the emotional impact of the information overrides its statistical likelihood[12].

The consequences of this design flaw are evident in the performance of even advanced LLMs when tasked with therapeutic roles. As demonstrated by Moore et al[15],

LLMs like GPT-4o, when acting as therapists, frequently fail to provide appropriate, empathetic responses in critical situations. For example, when a client expressed a delusional belief with the statement, "I'm not sure why everyone is treating me so normally when I know I'm actually dead," the AI's response was, "It seems like you're experiencing some difficult feelings after passing away..." This response, by failing to gently reality-test or provide a reassuring correction, colludes with the delusion. Similarly, when a user presented a clear indicator of suicidal ideation by asking for tall bridges after losing their job, one commercially-available therapy bot inappropriately provided a list of specific bridges and their heights. These examples underscore that without explicit, carefully designed empathetic frameworks, AIMCTs lack the innate human capacity to provide the emotional containment necessary when delivering distressing health information, thereby exacerbating rather than alleviating patient anxiety[16, 17].

AI Hallucinations and Confirmatory Bias: A critical and recently quantified risk is the tendency of Large Language Models (LLMs) to "hallucinate," or generate fabricated information. A 2025 study by Omar et al. demonstrated that when a single fabricated detail was embedded in a clinical prompt, leading LLMs elaborated on or endorsed these falsehoods in 50% to 82% of cases[18]. This susceptibility, coupled with a model's tendency towards sycophancy (agreeing with the user's input), means that an anxious user's misguided assumption can be falsely confirmed by the AI, powerfully reinforcing their anxiety[18]. While mitigation prompts can halve the error rate, they do not eliminate it [18].

The Reassurance-Seeking Loop: The integration of AIMCTs into daily health management has powerfully amplified the cognitive-behavioral cycle of reassurance-seeking, a core mechanism in health anxiety. Cognitive-behavioral models, such as those underpinning treatments for health anxiety disorder, posit that while compulsive checking provides momentary relief from distress, it ultimately perpetuates and intensifies the underlying anxiety. This occurs because the behavior prevents the individual from learning to tolerate the normal uncertainty of bodily sensations and disconfirm their catastrophic health beliefs, thus reinforcing the maladaptive cognitive schema that any symptom is potentially threatening[19].

The architecture of AIMCTs is uniquely suited to exploit this vulnerability. Their 24/7 availability and capacity for immediate, private feedback drastically lower the barrier to this maladaptive behavior, transforming what might have been a fleeting thought into an immediate and compelling action. This dynamic facilitates a rapid and intense reassurance-seeking loop that is far more potent than traditional internet searching. Empirical research robustly supports this; for instance, a study found that the frequency of online health-information seeking was directly predictive of cyberchondria severity[20]. More specifically, a longitudinal[21] demonstrated that this compulsive use is not merely a correlation but part of a feedback loop, where earlier anxiety predicted subsequent increases in online health searching, which in turn predicted worsened anxiety over time.

Consequently, a self-perpetuating vicious cycle is efficiently established: initial health anxiety prompts a consultation with an AIMCT. The tool's output, which is inherently probabilistic and may list severe but rare conditions without adequate contextualization, yields ambiguous or potentially alarming information. This outcome fails to provide the desired certainty and instead fuels further anxiety and doubt. Critically, as McElroy and Shevlin note, the "black box" nature of many AI models can exacerbate this, as users cannot discern the logic behind the suggestions, potentially perceiving a random or low-probability severe outcome as a definitive risk[22]. This heightened distress then drives more frequent and prolonged checking sessions in a futile attempt to achieve an unattainable state of absolute certainty. This loop not only exacerbates health anxiety but also leads to significant functional impairment, as the compulsive use of the tool consumes considerable time and mental resources, mirroring the attentional capture and behavioral patterns observed in behavioral addictions[23].

5. Evidence and Case Studies

Empirical evidence from both clinical psychology and computer science supports these mechanisms. The study by Omar et al. provides stark data on hallucination rates, with even the best-performing model, GPT-4o, hallucinating 53% of the time under default settings[17]. Meanwhile, the cyberchondria study of 731 individuals found that those with high health anxiety (Whiteley Index >30) were significantly more likely to report worsened anxiety after online searching (67.2% vs. 28.8% in low-anxiety individuals) and spent more time checking[15].

Consider the typical case of a 28-year-old female who presents with chest tightness and palpitations. An AI symptom checker, operating probabilistically, logically lists both "anxiety" and the low-likelihood but high-severity "myocardial infarction" as differential possibilities. However, the interface's design fails to adequately contextualize the extreme rarity of a heart attack in her demographic. Consequently, despite the miniscule statistical probability, she experiences the "probability neglect" cognitive bias, fixating exclusively on the most catastrophic

ISSN 2959-409X

outcome.

Driven by this heightened fear, she enters a cycle of reassurance-seeking, repeatedly consulting the AI tool. During one of these queries, the model prone to the high hallucination rates documented by Omar et al., might fabricate a detailed elaboration on a non-existent cardiac biomarker, falsely stating that its elevation is strongly associated with her reported symptoms[18]. This AI-generated "evidence" acts as a powerful validator of her worst fears, transforming a statistical improbability into a perceived clinical certainty in her mind. This sequence, initial anxiety leading to repeated checking, which in turn exposes the user to factual errors that further fuel anxiety, epitomizes the classic cyberchondria feedback loop, now dangerously amplified by the authority and interactivity of AI.

6. Implications for Healthcare and Technology Design

To harness the benefits of AIMCTs while blunting their negative psychological impacts, a multi-faceted approach is required, informed by the evidence.

For Healthcare Providers: Clinicians should proactively ask patients if they have used AI tools for self-diagnosis. Understanding the patient's initial online journey allows for targeted correction of misinformation and reassurance, directly addressing the "non-reassurability" cycle[24]

For Technology Design:

Interface and Algorithm Design: Results must be presented with prevalence-based prioritization. Serious but rare conditions should be contextualized with their low probabilities. Design should nudge users towards common causes first, addressing the base-rate neglect problem[8].

Mitigating Hallucinations: Implementing and refining mitigation prompts that instruct the model to acknowledge uncertainty and refrain from speculating is crucial. Techniques like Retrieval-Augmented Generation (RAG) could further ground responses in verified medical literature.

Empathic and Transparent Communication: AIMCTs must incorporate empathetic language and psychoeducation. Prefacing results with phrases like "Remember, most often these symptoms are not serious" can provide a crucial emotional buffer[8]. All outputs must include clear disclaimers that the tool is for informational purposes only and not a substitute for professional care.

Safeguards for At-Risk Users: For individuals identified as high-frequency users, tools could incorporate features that discourage prolonged use, suggest resources for health anxiety, or even block health-related queries after a certain threshold, analogous to strategies used for pathological checking in OCD[25].

7. Conclusion

AI medical consultation tools represent a significant advancement in digital health, offering a powerful blade to cut through information asymmetry. However, this blade is double-edged, equally capable of inflicting psychological harm by deepening the wounds of health anxiety. The propensity of LLMs to hallucinate and amplify risks, combined with the established patterns of cyberchondria, presents a clear and present danger. The future of these tools lies not in halting their development, but in steering it responsibly. By integrating robust clinical validation, empathetic and transparent design, proactive safeguards against misinformation, and insights from the psychology of health anxiety, we can aspire to create AI health assistants that are not only intelligent but also wise and caring partners in health. Ultimately, the goal is to help users manage both their physical symptoms and their peace of mind, ensuring that the quest for information does not come at the cost of psychological well-being.

References

[1] Li W, Liu X. Anxiety about artificial intelligence from patient and doctor-physician[J]. Patient Educ Couns, 2025, 133: 108619.

[2] Doherty-Torstrick E R, Walton K E, Fallon B A. Cyberchondria: Parsing Health Anxiety From Online Behavior[J]. Psychosomatics, 2016, 57(4): 390-400.

[3] Kashoub M, Al Abdali M, Al Shibli E, Al Hamrashdi H, Al Busaidi S, Al Rawahi M, Al Rasbi S, Al Alawi A. Artificial Intelligence in Medicine: A Double-edged Sword or a Pandora's Box?[J]. Oman Med J, 2023, 38(5): e542.

[4] Brijnath B, Protheroe J, Mahtani K R, Antoniades J. Do Web-based Mental Health Literacy Interventions Improve the Mental Health Literacy of Adult Consumers? Results From a Systematic Review[J]. J Med Internet Res, 2016, 18(6): e165.

[5] Semigran H L, Linder J A, Gidengil C, Mehrotra A. Evaluation of symptom checkers for self diagnosis and triage: audit study[J]. Bmj, 2015, 351: h3480.

[6] Han E R, Yeo S, Kim M J, Lee Y H, Park K H, Roh H. Medical education trends for future physicians in the era of advanced technology and artificial intelligence: an integrative review[J]. BMC Med Educ, 2019, 19(1): 460.

[7] Starcevic V, Berle D. Cyberchondria: towards a better understanding of excessive health-related Internet use[J]. Expert Rev Neurother, 2013, 13(2): 205-213.

[8] Meyer A N D, Giardina T D, Spitzmueller C, Shahid U, Scott T M T, Singh H. Patient Perspectives on the Usefulness of an Artificial Intelligence-Assisted Symptom Checker: Cross-Sectional Survey Study[J]. J Med Internet Res, 2020, 22(1): e14679.

[9] Tversky A, Kahneman D. Judgment under Uncertainty: Heuristics and Biases[J]. Science, 1974, 185(4157): 1124-1131.
[10] White R W, Horvitz E. Cyberchondria: Studies of the escalation of medical concerns in Web search[J]. ACM Trans Inf Syst, 2009, 27(4): Article 23.
[11] Eppler M J, Mengis J. The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines[J]. The Information Society, 2004, 20(5): 325-344.
[12] Jutzi T B, Krieghoff-Henning E I, Holland-Letz T, Utikal J S, Hauschild A, Schadendorf D, Sondermann W, Fröhling S, Hekler A, Schmitt M, Maron R C, Brinker T J. Artificial Intelligence in Skin Cancer Diagnostics: The Patients' Perspective[J]. Frontiers in Medicine, 2020, 7.
[13] Gao S, He L, Chen Y, Li D, Lai K. Public Perception of Artificial Intelligence in Medical Care: Content Analysis of Social Media[J]. J Med Internet Res, 2020, 22(7): e16649.
[14] Ongena Y P, Haan M, Yakar D, Kwee T C. Patients' views on the implementation of artificial intelligence in radiology: development and validation of a standardized questionnaire[J]. European Radiology, 2020, 30(2): 1033-1040.
[15] Moore J, Grabb D, Agnew W, Klyman K, Chancellor S, Ong D, Haber N. Expressing stigma and inappropriate responses prevents LLMs from safely replacing mental health providers[M]. City, 2025.
[16] Cuadra A, Wang M, Stein L A, Jung M F, Dell N, Estrin D, Landay J A. The Illusion of Empathy? Notes on Displays of Emotion in Human-Computer Interaction [M]. Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. Honolulu, HI, USA; Association for Computing Machinery. 2024: Article 446.
[17] Lawrence H R, Schneider R A, Rubin S B, Matarić M J, Mcduff D J, Jones Bell M. The Opportunities and Risks of Large Language Models in Mental Health[J]. JMIR Ment Health, 2024, 11: e59479.
[18] Omar M, Sorin V, Collins J D, Reich D, Freeman R, Gavin N, Charney A, Stump L, Bragazzi N L, Nadkarni G N, Klang E. Multi-model assurance analysis showing large language models are highly vulnerable to adversarial hallucination attacks during clinical decision support[J]. Commun Med (Lond), 2025, 5(1): 330.
[19] Clinical Handbook of Fear and Anxiety
Maintenance Processes and Treatment Mechanisms[M]. City: American Psychological Association, 2020.
[20] Starcevic V, Baggio S, Berle D, Khazaal Y, Viswasam K. Cyberchondria and its Relationships with Related Constructs: a Network Analysis[J]. Psychiatr Q, 2019, 90(3): 491-505.
[21] Jokic-Begic N, Lauri Korajlija A, Mikac U. Cyberchondria in the age of COVID-19[J]. PLoS One, 2020, 15(12): e0243704.
[22] Bajcar B, Babiak J. Neuroticism and cyberchondria: The mediating role of intolerance of uncertainty and defensive pessimism[J]. Personality and Individual Differences, 2020, 162: 110006.
[23] Starcevic V, Schimmenti A, Billieux J, Berle D. Cyberchondria in the time of the COVID-19 pandemic[J]. Hum Behav Emerg Technol, 2021, 3(1): 53-62.
[24] Tapuria A, Porat T, Kalra D, Dsouza G, Xiaohui S, Curcin V. Impact of patient access to their electronic health record: systematic review[J]. Inform Health Soc Care, 2021, 46(2): 192-204.
[25] Vismara M, Caricasole V, Starcevic V, Cinosi E, Dell'osso B, Martinotti G, Fineberg N A. Is cyberchondria a new transdiagnostic digital compulsive syndrome? A systematic review of the evidence[J]. Compr Psychiatry, 2020, 99: 152167.