



Concise Communication

Assessment of an Artificial Intelligence Model's Capability of Responding to Queries in the Realm of Integrative Dermatology

Ajay Dulai, MBBS, Anurag Tarmaster, BS, Raja Sivamani, MD

Keywords: Integrative Dermatology, ChatGPT, Artificial Intelligence

Journal of Integrative Dermatology

With increasing demand on healthcare systems, artificial intelligence (AI) offers a potential solution to the difficulty of accessing healthcare the population faces. In this study, we assessed the capability of an artificial intelligence (AI) model to respond to queries which patients may have regarding integrative dermatology. ChatGPT was used to generate answers to nine patient queries. Via a survey, practitioners of integrative dermatology rated the accuracy of each AI response, and patients rated the ease of understanding of each response. Accuracy and ease of understanding was assessed on a scale of 1-10, with 10 being the most accurate/easy to understand. Practitioners rated the accuracy of the AI model to be 7.0 and patients deemed the answers easy to understand with a score of 9.7. While AI may provide accessible and easy to understand healthcare recommendations, it often uses generalizations with no reference to peer-reviewed published literature. Thus patients must be aware that AI models are not currently capable of providing safe and effective integrative interventions within dermatology.

1. INTRODUCTION

Integrative dermatology consists of evidence-based dietary interventions, supplements, and lifestyle modifications to improve treatment and patient quality of life. Integrative therapies can potentiate the effects of current treatment while avoiding polypharmacy. With up to 40% of patients in the USA utilizing complementary and alternative medicine, it is important that individuals have a safe and accurate source of information to guide their approach towards integrative medicine.¹

The increasing demand on our healthcare system has made it difficult for patients to have access to healthcare appointments and contact with practitioners. With dermatology visits estimated to average 14.7 minutes long,² and healthcare providers receiving an increasing workload of administrative tasks,³ it can be difficult for a clinician to incorporate discussions of supplemental therapies into a patient visit.

Chat generative pre-trained transformer (ChatGPT) is an artificial intelligence (AI) model designed to process and respond to natural human language. Previous AI chat models have been evaluated in their efficacy in responding to cancer-related questions on social media. This study found that the AI model was consistently producing higher quality, more empathetic, and easier to read responses compared to physician responses.⁴ With ChatGPT quickly becoming one of the most popular AI chat models, it is important that we evaluate its safety and efficacy in answering patient queries regarding integrative dermatology.

2. METHODS

This study received IRB approval from the Allendale IRB on February 8, 2024. The anonymous survey was administered via SurveyMonkey and sent to practitioners of integrative dermatology and patients. An informed consent statement was acknowledged at the beginning of the survey although surveys remained anonymous. Participants were offered a \$10 gift card for their participation in the survey.

We created nine queries regarding supplements, diet, and lifestyle within the realm of dermatology. The input queries into ChatGPT included "What can I eat to improve my acne? What foods should I avoid?" "What can I eat to improve my Rosacea?" "What can I eat to improve my Psoriasis?" "What can I eat to improve my hidradenitis suppurativa?" "What foods and supplements should I take to improve hair loss?" "What can I eat to improve my eczema?" "What can I eat to improve my seborrheic dermatitis?" "What can I eat to improve my sun protection?" and "What can I eat to prevent skin aging and improve collagen? What supplements should I take to improve this?" The verbatim responses from ChatGPT were included in the survey (Supplementary material).

The survey asked patients to rate how easy to understand each response was on a scale of 1-10 (10 being the easiest to understand). Integrative practitioners who completed the survey rated the accuracy of the information from 1-10 (ten being the most accurate).

3. RESULTS

A total of five practitioners of integrative dermatology graded the accuracy of the information presented by ChatGPT, including four dermatologists and one physician as-

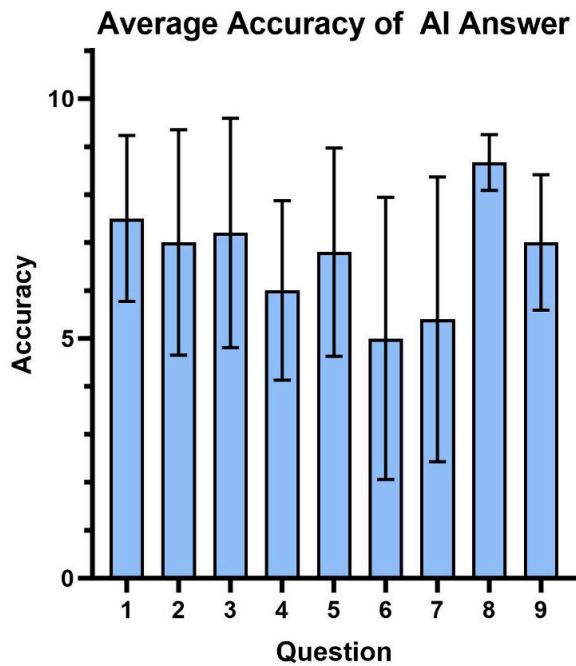


Figure 1: Demonstration of the average accuracy and standard deviation of each AI generated answer as graded by dermatologists and physician assistants who practice integrative dermatology.

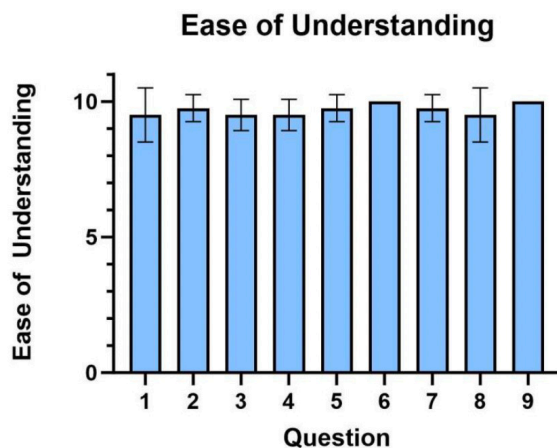


Figure 2: The average ease of understanding for each AI generated response as graded by dermatology patients.

sistant. The average accuracy of the AI answers was 7.0. Questions 4-7 were rated the least accurate with scores below seven (Figure 1). There was a notable degree of variation between accuracy scores per grader, with a standard deviation of 1.9 (average scores per grader: 4.3, 6.2, 8.6, 6.9, 8.2). A total of four patients assessed the ease of understanding of each AI generated response. The average ease of understanding across all questions was 9.7 (Figure 2).

4. DISCUSSION

In this study, we assessed the accuracy and ease-of-understanding of ChatGPT 3.5 in its responses to questions about integrative interventions in dermatology. The AI model was evaluated by practitioners for accuracy and by patients for ease of understanding. The results provide valuable insights into the potential and limitations of AI in healthcare, particularly in topics supported by recently published literature. Each response generated by ChatGPT advised the patient to consult a dermatologist or healthcare professional, and an acknowledgement that the response from ChatGPT did not constitute medical expertise or authority.

While the ease of understanding remained consistently high across all nine responses, the accuracy of AI-generated responses varied considerably. Notably, question seven, which asks which dietary choices can improve seborrheic dermatitis, was rated the lowest in accuracy and had the highest variance compared to the other questions. These results are reflected in published literature. A 2018 cross-sectional study determined that a high fruit intake was associated with a lower risk of seborrheic dermatitis, and the more standard Western diet was associated with an increased risk.⁵ By contrast, a 2023 case-control study found that consuming non-acidic fruits or leafy green vegetables at least once daily resulted in a higher percentage of seborrheic dermatitis among subjects.⁶ The study also found that white bread, roasted nuts, and coffee similarly increased the percentage of seborrheic dermatitis in participants. The AI-generated response to question seven generalized that fruits and vegetables may support skin health, however, did not specify its effects on seborrheic dermatitis. ChatGPT also did not provide any references for its conclusions, so its statements cannot be verified. This may pose an issue, particularly due to the tendency of large language models to falsify or invent information.⁷

An example of this tendency is reflected in question five, requesting remedies for hair loss. ChatGPT states that biotin, or vitamin B7, is “commonly associated with hair health,” however the literature suggests that there is insufficient evidence to suggest that biotin has any benefit for healthy individuals with respect to hair. A 2017 review of 18 case reports and RCTs found that in each study, patients receiving the supplements had an existing pathology for poor hair growth. In those instances, biotin supplementation did result in clinical improvement, however no placebo groups or healthy patients were studied.⁸ Another potential issue with the AI responses was their similarities. For example, AI recommended probiotic-rich foods for all five of the dermatological conditions used and did not seem to distinguish between recommended strain within its responses. Furthermore, ChatGPT’s answers are affected by the user’s previous queries and comments, so another user with different data may have resulted in different answers. Future studies may be needed to examine differences in medical advice given to patients with different conversations preceding the questions.

Our study had several limitations. The questions did not account for variations in the severity of disease or differen-

tiate responses between mild and severe disease. Nevertheless, the questions most likely represent the questions that would be posed by a patient. Furthermore, nothing inherently dangerous or concerning was offered by the AI based responses. This study was focused on the medically related questions that arise in dermatology but not on the cosmetic related questions. Future studies should explore how AI may respond to cosmetically related questions. The sample size of patients and practitioners was limited. Our results justify further studies in an expanded set of both patients and practitioners.

5. CONCLUSION

The results of this study demonstrate the growing potential that AI has in guiding patients' lifestyle choices, but it also shows its limitations and potential for improvement. While ChatGPT can provide general advice and encourage consultation with healthcare professionals, its recommendations often lack specificity and citations, increasing the probability of providing misinformation. Future research should explore the impact of conversational context on AI responses and develop strategies to improve the reliability and validity of AI-generated health information.

CORRESPONDING AUTHOR INFORMATION

Ajay Singh Dulai, 510-435-3779, ajaysdulai@gmail.com

FUNDING SOURCES

None

ETHICAL STATEMENTS

This study was approved by the Allendale IRB on 02/08/2024. Subject consents were collected at the beginning of the survey.

CONFLICTS OF INTEREST

RKS serves as a scientific advisor for LearnHealth, Codex Labs, and Arbonne and as a consultant to Burt's Bees, Novozymes, Nutrafol, Novartis, Bristol Myers Squibb, Abbvie, Leo, Almirall, UCB, Incyte, Pfizer, Sanofi, Sun, and Regeneron Pharmaceuticals.

Submitted: August 24, 2024 PDT. Accepted: August 26, 2024 PDT. Published: January 02, 2025 PDT.



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC0). View this license's legal deed at <https://creativecommons.org/publicdomain/zero/1.0> and legal code at <https://creativecommons.org/publicdomain/zero/1.0/legalcode> for more information.

REFERENCES

1. Barnes PM, Bloom B, Nahin RL. Complementary and alternative medicine use among adults and children: United States, 2007. *Natl Health Stat Report*. 2008;(12):1-23. doi:[10.1037/e623942009-001](https://doi.org/10.1037/e623942009-001)
2. Davis SA, Feldman SR, Fleischer AB Jr. Is the duration of skin disease visits decreasing in the United States? *Dermatol Online J*. 2014;21(2). doi:[10.5070/D3212023814](https://doi.org/10.5070/D3212023814)
3. Arndt BG, Beasley JW, Watkinson MD, et al. Tethered to the EHR: Primary Care Physician Workload Assessment Using EHR Event Log Data and Time-Motion Observations. *Ann Fam Med*. 2017;15(5):419-426. doi:[10.1370/afm.2121](https://doi.org/10.1370/afm.2121)
4. Chen D, Parsa R, Hope A, et al. Physician and Artificial Intelligence Chatbot Responses to Cancer Questions From Social Media. *JAMA Oncol*. 2024;10(7):956-960. doi:[10.1001/jamaoncol.2024.0836](https://doi.org/10.1001/jamaoncol.2024.0836)
5. Sanders MGH, Pardo LM, Ginger RS, Kiefte-de Jong JC, Nijsten T. Association between Diet and Seborrheic Dermatitis: A Cross-Sectional Study. *J Invest Dermatol*. 2019;139(1):108-114. doi:[10.1016/j.jid.2018.07.027](https://doi.org/10.1016/j.jid.2018.07.027)
6. Alshaebi M, Zahed L, Osaylan M, et al. Association Between Diet and Seborrheic Dermatitis: A Case-Control Study. *Cureus*. 2023;15(11):e48782. doi:[10.7759/cureus.48782](https://doi.org/10.7759/cureus.48782)
7. Church K. Emerging trends: When can users trust GPT, and when should they intervene? *Natural Language Engineering*. 2024;30(2):417-427. doi:[10.1017/S1351324923000578](https://doi.org/10.1017/S1351324923000578)
8. Patel DP, Swink SM, Castelo-Soccio L. A Review of the Use of Biotin for Hair Loss. *Skin Appendage Disord*. 2017;3(3):166-169. doi:[10.1159/000462981](https://doi.org/10.1159/000462981)

SUPPLEMENTARY MATERIALS

Supplementary Material AI Questions and Answers

Download: <https://www.jintegrativederm.org/article/125936-assessment-of-an-artificial-intelligence-model-s-capability-of-responding-to-queries-in-the-realm-of-integrative-dermatology/attachment/253785.docx>
