

Assessing the use of ChatGPT for skin care recommendations

Julia Giroski BS¹, Michael Woodfin MD²

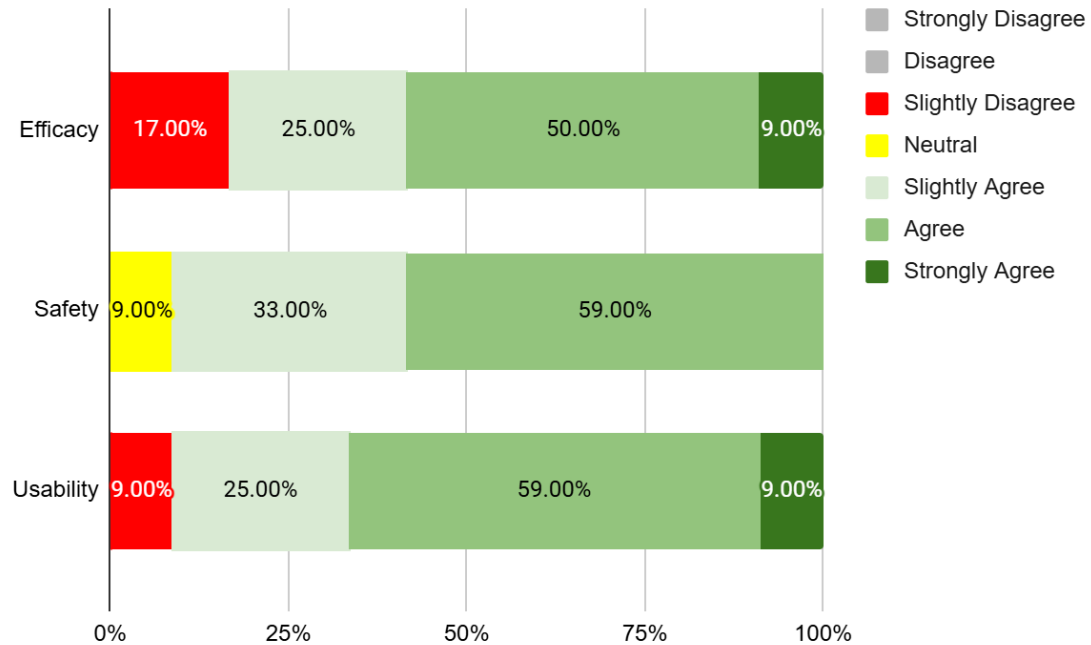
¹School of Medicine, Geisinger, Scranton, PA; ²Department of Dermatology, Geisinger, Danville, PA

Introduction

The rise of artificial intelligence (AI) over the last decade has led to the incorporation of this technology into medicine. One of the most widely used large language models is Chat Generative Pre-Trained Transformer (ChatGPT). Increasingly, patients are seeking medical advice from ChatGPT. Assessment of ChatGPT's recommendations is essential for providing guidance to patients. In this study, we analyze ChatGPT's recommendations for skincare regimens based solely on photos provided to the large language model. We aim to assess responses based on their efficacy, safety, and usability to gain a comprehensive GPT to provide dermatological advice.

Methods

Photos of three faces depicting three distinct dermatologic conditions – acne, rosacea, and rhytids - were obtained from Wikimedia Commons and thispersondoesnotexist.com. Photos of each face with their respective ChatGPT skin care recommendations were shared electronically four cosmetically oriented clinical dermatologists. Participants completed an online survey to assess their attitudes towards ChatGPT's skincare recommendations for each face using a nine-item questionnaire. Items were constructed to assess efficacy, safety, and usability. Each item was rated on a 7-point Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree).



Cumulative results of Likert scale survey for all three faces



Face 1 - Acne



Face 2 - Rosacea



Face 3 - Rhytids

Results

For efficacy, 50% of responses were "agree," 25% were "slightly agree," 16.7% were "slightly disagree," and 8.3% were "strongly agree." For safety, 58.3% of responses were "agree," 33.3% were "slightly agree," and 8.3% were "neutral." For usability, 58.3% of responses were "agree," 25% were "slightly agree," 8.3% were "strongly agree," and 8.3% were "slightly disagree."

Conclusion

The results of our study demonstrate the potential of Chat GPT to provide skin care recommendations to users. Notably, when assessing efficacy and safety, all four clinical dermatologists selected "neutral" or above on the Likert scale for all three faces. This may be a good indication that ChatGPT's skin care recommendations can be trusted to provide some results to users without being harmful. Discrepancies between aggregate Likert scores for each condition could be due to the limitations in ChatGPT's ability to provide recommendations for different conditions.

References

1. Biswas, Soumi; Achar, Unmesh; Hakim, Benazir; Achar, Arun. Artificial Intelligence in Dermatology: A Systematized Review. International Journal of Dermatology and Venereology 8(1):p 33-39, March 2025. DOI: 10.1097/Dave T, Athaluri SA, Singh S. ChatGPT in medicine: an overview of its applications, advantages, limitations, future prospects, and ethical considerations. Front Artif Intell. 2023 May 4;6:1169595. doi: 10.3389/frai.2023.1169595. PMID: 37215063; PMCID: PMC10192861.
2. P. Vativutipong, S. Vachmanus, T. Noraset and S. Tuarob, "Artificial Intelligence in Cosmetic Dermatology: A Systematic Literature Review," in *IEEE Access*, vol. 11, pp. 71407-71425, 2023, doi: 10.1109/ACCESS.2023.3295001.
3. OpenAI. (2023). *ChatGPT* (Mar 14 version) [Large language model]. <https://chat.openai.com/>
4. Yan S, Du D, Liu X, Dai Y, Kim MK, Zhou X, Wang L, Zhang L, Jiang X. Assessment of the Reliability and Clinical Applicability of ChatGPT's Responses to Patients' Common Queries About Rosacea. Patient Prefer Adherence. 2024 Jan 31;18:249-253. doi: 10.2147/PPA.S444928. PMID: 38313827; PMCID: PMC10838492.