

Research Letter

Readability of Information Generated by ChatGPT for Hidradenitis Suppurativa

Lauren Gawey¹, MD; Caitlyn B Dagenet², BS; Khiem A Tran³, MD, PhD; Sarah Park⁴, BA; Jennifer L Hsiao⁵, MD; Vivian Shi³, MD

¹University of Arkansas for Medical Sciences, Little Rock, AR, United States

²College of Medicine, The University of Arizona, Tucson, AZ, United States

³Department of Dermatology, University of Arkansas for Medical Sciences, Little Rock, AR, United States

⁴David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, United States

⁵Department of Dermatology, University of Southern California, Los Angeles, CA, United States

Corresponding Author:

Vivian Shi, MD

Department of Dermatology

University of Arkansas for Medical Sciences

4301 W Markham St

#576

Little Rock, AR, 72205

United States

Phone: 1 8148022747

Email: vivian.shi.publications@gmail.com

(*JMIR Dermatol* 2024;7:e55204) doi: [10.2196/55204](https://doi.org/10.2196/55204)

KEYWORDS

hidradenitis suppurativa; ChatGPT; Chat-GPT; chatbot; chatbots; chat-bot; chat-bots; machine learning; ML; artificial intelligence; AI; algorithm; algorithms; predictive model; predictive models; predictive analytics; predictive system; practical model; practical models; deep learning; patient resources; readability

Introduction

ChatGPT is an artificial intelligence (AI) language model that has emerged as a resource for patient education, with over 100 million general users worldwide [1]. Despite its popularity, the readability of information provided by ChatGPT on dermatological conditions, such as hidradenitis suppurativa (HS), has yet to be explored. Patients with HS wait an average of 7 years after their initial symptoms appear to seek medical attention, which is largely attributed to insufficient awareness of the condition [2]. Effective patient education is vital for informed decision-making and self-management of medical conditions. The American Medical Association and the National Institutes of Health recommend that patient educational materials should be written at a sixth- and eighth-grade reading level, respectively [3]. This study aimed to assess the readability of ChatGPT-generated responses in comparison to established HS educational materials and web-based resources.

Methods

We compared the readability of responses to frequently asked questions from the HS Foundation (HSF), HS Patient Guide

(HSPG) [4], and ChatGPT-3.5, along with HS-related websites (Google, Yahoo, and Bing were searched using the term “hidradenitis suppurativa”). The top 50 web pages from each search engine were reviewed, of which, 55 met inclusion criteria for further analysis. Readability was determined by average readability grade level and Flesch Reading Ease, which is scored from 0 to 100, with a higher score indicating that the material is easier to read. These readability formulas take into account the number of characters, syllables, words, and sentences to determine their score. Lexical density—a measure of linguistic complexity—and other text readability metrics were also recorded. While reviewers did not directly participate in the scoring process, the use of standardized software from online-utility.org facilitated objective evaluations aligned with established criteria for readability assessment. The 2-tailed Student *t* test was used for bivariate analysis, with significance set at $P < .05$.

Results

ChatGPT-generated responses had an average readability grade level of 15.0, which was significantly higher than that of the HSF (8.0), the HSPG (11.0), and HS-related websites (12.0; $P < .001$). Flesch Reading Ease was significantly lower for

ChatGPT-generated responses (28.7) than for the HSF (66.1), the HSPG (49.2), and HS-related websites (40.9; $P < .001$; Figure 1). Both ChatGPT and HS-related websites had a higher lexical

density of 58.0 and 57.47 respectively, indicating higher linguistic complexity than that for the HSF (49.1) and the HSPG (52.6; Figure 2).

Figure 1. Readability of information for patients with hidradenitis suppurativa (HS) based on the average readability grade level, Flesch Reading Ease, and lexical density. The average readability grade level is calculated by averaging the Flesch Kincaid Grade Level, Gunning Fog Index, Simple Measure of Gobbledygook index, Coleman–Liau index, and automated readability index scores. Flesch Reading Ease is scored between 0 and 100, with a higher score indicating that the article is easier to read. Lexical density estimates linguistic complexity in a composition from the functional words (grammatical units) and content words (lexical units), calculated by comparing the ratio of lexical items to the total number of words.

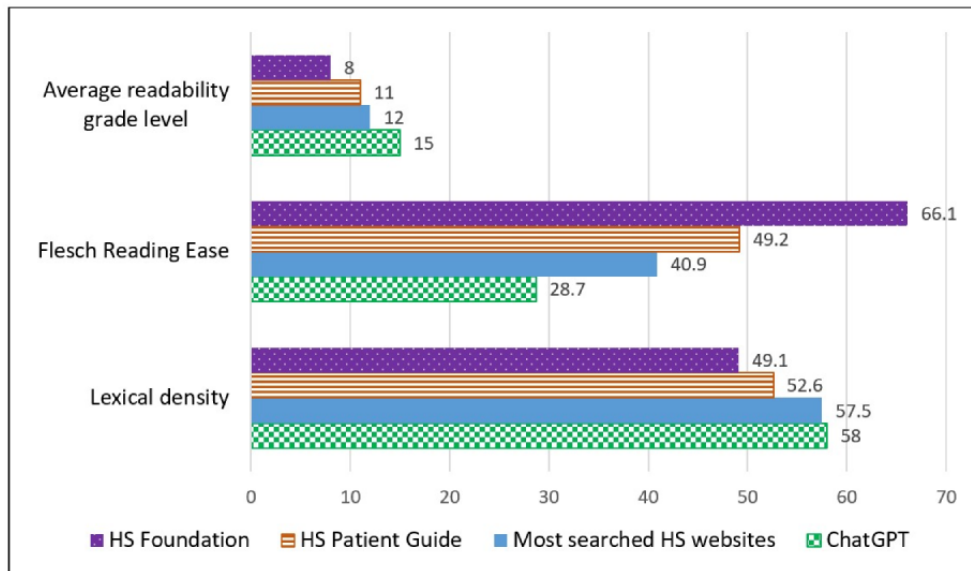
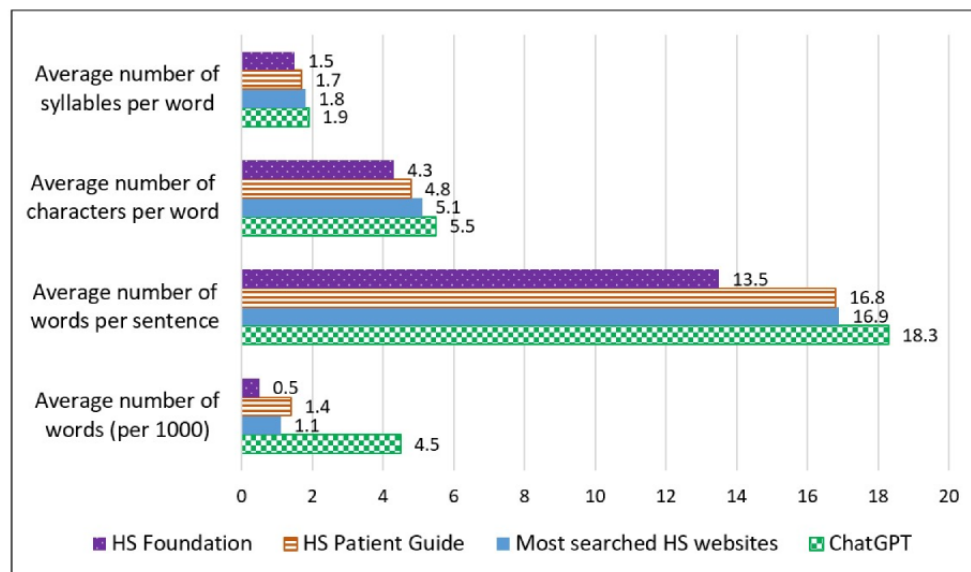


Figure 2. Text readability metrics of information for patients with hidradenitis suppurativa (HS). These values represent an average of text readability metrics for each specified source.



Discussion

Our results show that ChatGPT-generated responses were 7-9 grade levels above the recommended reading level and had a higher linguistic complexity than other HS-related web-based resources. These findings underscore the limitations of ChatGPT as a patient resource for HS, as the higher reading level and linguistic complexity of ChatGPT could hinder patient comprehension. The potential of AI-driven resources, such as ChatGPT, to transform health care communication hinges on their ability to align with recommended readability standards.

One study showed that when prompting AI to convert patient educational material to an easier grade level, AI could improve the readability of input material [5]. However, without prompting, the baseline reading level of ChatGPT-generated information is much higher than is recommended for patient educational materials. It is important to note that the practice of prompting AI systems for readability adjustments is currently not commonplace among the general public user base. As AI integration becomes more commonplace, future studies can explore and compare the effectiveness of prompting strategies to make consistent adjustments in readability. Educating health

care providers about the availability of options to prompt ChatGPT responses for enhanced readability can allow them to counsel their patients on adjusting readability levels that are most suitable for their preferences.

While the readability formulas used in this study offer a useful quantitative measure of text complexity, they focus primarily on surface-level features such as sentence length and syllable count, neglecting the structural complexity of texts, such as

coherence, organization, and language context, which also influence readability. Additionally, AI-generated texts may exhibit variations in tone, style, and content that traditional readability formulas may struggle to evaluate accurately.

Future directions should work toward improving not only the readability of AI, but also the quality and accuracy of generated information. The findings of this study serve as a foundational reference for future AI resource development in dermatology.

Conflicts of Interest

VS is on the board of directors for the Hidradenitis Suppurativa Foundation; is a shareholder in Learn Health; and has served as an advisory board member, investigator, speaker, and received research funding from Genzyme (Sanofi), Regeneron Pharmaceuticals, AbbVie, Eli Lilly, Novartis, Sun Pharmaceutical Industries Limited, LEO Pharma Inc, Pfizer, Incyte Corporation, Boehringer Ingelheim, Aristeia Therapeutics, VYNE Therapeutics (formerly Menlo Therapeutics), Dermira, Inc (Eli Lilly), Burt's Bees, Galderma, Kiniksa Pharmaceuticals, UCB, TARGET PharmaSolutions, Altus Lab/cQuell, MYOR, Polyfins Technology, GPSkin, Skin Actives Scientific, and the National Eczema Association. JLH is on the board of directors for the Hidradenitis Suppurativa Foundation, a consultant for Novartis, and speaker for AbbVie. LG, CBD, KAT, and SP have nothing to declare.

References

1. Mesko B. The ChatGPT (generative artificial intelligence) revolution has made artificial intelligence approachable for medical professionals. *J Med Internet Res*. Jun 22, 2023;25:e48392. [FREE Full text] [doi: [10.2196/48392](https://doi.org/10.2196/48392)] [Medline: [37347508](https://pubmed.ncbi.nlm.nih.gov/37347508/)]
2. Saunte DM, Boer J, Stratigos A, Szepletowski JC, Hamzavi I, Kim KH, et al. Diagnostic delay in hidradenitis suppurativa is a global problem. *Br J Dermatol*. Dec 03, 2015;173(6):1546-1549. [doi: [10.1111/bjd.14038](https://doi.org/10.1111/bjd.14038)] [Medline: [26198191](https://pubmed.ncbi.nlm.nih.gov/26198191/)]
3. Rooney MK, Santiago G, Perni S, Horowitz DP, McCall AR, Einstein AJ, et al. Readability of patient education materials from high-impact medical journals: a 20-year analysis. *J Patient Exp*. Mar 03, 2021;8:2374373521998847. [FREE Full text] [doi: [10.1177/2374373521998847](https://doi.org/10.1177/2374373521998847)] [Medline: [34179407](https://pubmed.ncbi.nlm.nih.gov/34179407/)]
4. Hidradenitis Suppurativa Patient Guide. URL: <https://hspatientguide.com/> [accessed 2023-07-20]
5. Kirchner GJ, Kim RY, Weddle JB, Bible JE. Can artificial intelligence improve the readability of patient education materials? *Clin Orthop Relat Res*. Apr 28, 2023;481(11):2260-2267. [doi: [10.1097/corr.0000000000002668](https://doi.org/10.1097/corr.0000000000002668)]

Abbreviations

AI: artificial intelligence
HS: hidradenitis suppurativa
HSF: Hidradenitis Suppurativa Foundation
HSPG: Hidradenitis Suppurativa Patient Guide

Edited by J Solomon, I Brooks; submitted 05.12.23; peer-reviewed by G Farid, S Grigaliunas, H Sun; comments to author 15.04.24; revised version received 12.05.24; accepted 29.06.24; published 14.08.24

Please cite as:

Gawey L, Dagenet CB, Tran KA, Park S, Hsiao JL, Shi V
Readability of Information Generated by ChatGPT for Hidradenitis Suppurativa
JMIR Dermatol 2024;7:e55204
URL: <https://derma.jmir.org/2024/1/e55204>
doi: [10.2196/55204](https://doi.org/10.2196/55204)
PMID:

©Lauren Gawey, Caitlyn B Dagenet, Khiem A Tran, Sarah Park, Jennifer L Hsiao, Vivian Shi. Originally published in *JMIR Dermatology* (<http://derma.jmir.org>), 14.08.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Dermatology*, is properly cited. The complete bibliographic information, a link to the original publication on <http://derma.jmir.org>, as well as this copyright and license information must be included.