



Introduction

- Patient education materials (PEMs) are essential tools for aiding healthcare workers in providing patients with high-quality, up-to-date information about their diagnosis, treatment, and recovery.
- AMA, NIH, and CDC recommend that PEMs not exceed a 6th-grade reading level¹
- Many orthopedic PEMs, including those from the American Academy of Orthopedic Surgeons (AAOS) and the American Society for Surgery of the Hand (ASSH), significantly exceed this threshold.²
- English-first-language patients with limited health literacy face barriers to understanding such materials, contributing to poorer outcomes and reduced engagement in care.³
- Simplifying PEMs using AI may bridge the readability gap, reduce health disparities, and foster patient autonomy.⁴
- Our study explores the ability of two popular large language models to translate PEMs for Carpometacarpal (CMC) Arthritis, a highly prevalent hand condition in adults with a complex diagnosis with several treatment and recovery pathways.

Materials & Methods

- PEMs on CMC Arthritis were obtained from AAOS and ASSH websites.
- Original text from each webpage was copied in plain text format into individual Microsoft Word 2025 (Version 16.100.1 (25081721)) documents. Pictures, picture captions, and extraneous content unrelated to the main text were removed. Copyright notes, date stamps, citations, hyperlinks, tables were removed, and numerical digits were converted to their written equivalents (e.g., “7” as “seven”).
- Each website’s PEM was inserted into OpenAI’s ChatGPT-5.0 and Google’s Gemini 2.5:

PROMPT

“Simplify the following patient education material about carpometacarpal (CMC) thumb arthritis so that it is written at approximately a sixth-grade reading level. Keep the medical accuracy but make the language clear and easy for patients to understand. Output as a complete patient education handout: [‘Original PEM here’].”

- Three methods of readability were assessed: Flesch-Kincaid Grade Level (FKGL), Flesch Reading Ease (FRE), and Simple Measure of Gobbledygook (SMOG). FKGL emphasizes sentence length and syllable count, FRE emphasizes overall reading ease, and SMOG emphasizes polysyllabic words.
- ASSH and AAOS PEMs and their translated outputs from ChatGPT-5.0 and Google Gemini 2.5 were analyzed.
- The FKGL provides a standardized U.S. academic grade level, while the FRE provides an overall measure of reading ease, ranging from 0 (very difficult) to 100 (very easy), with higher scores indicating easier-to-read text. Both scores together characterize the readability of the PEMs.
- SMOG was calculated using the online SMOG readability calculator (<https://charactercalculator.com/smog-readability/>). The same standardized PEM text used for FKGL and FRE was entered into the calculator, which automatically counts sentences and identifies words with three or more syllables to compute a grade-level score representing the U.S. school level required to understand the text.

Figures

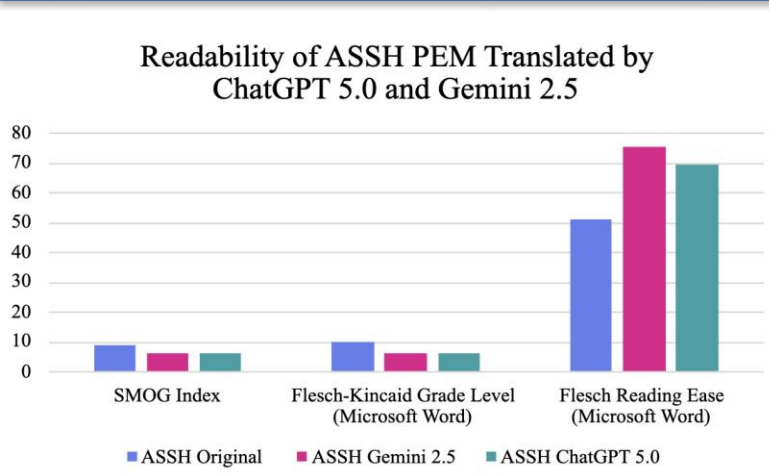


Figure 1. Comparison of SMOG, FKGL, and FRE for original and AI-translated ASSH PEM.

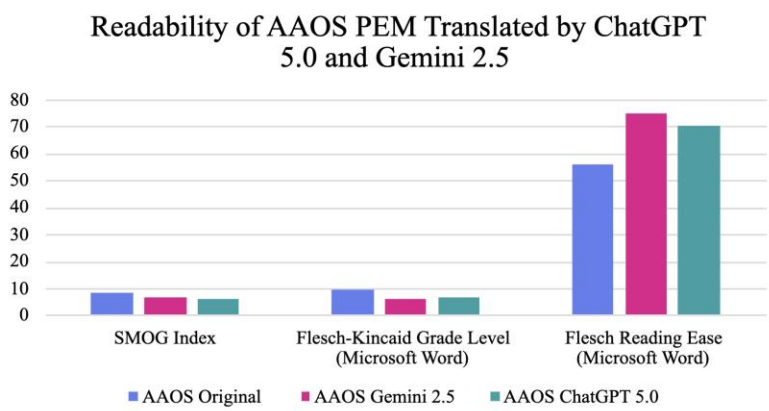


Figure 2. Comparison of SMOG, FKGL, and FRE for original and AI-translated AAOS PEM

Results

- Readability analysis of the original PEMs showed that the ASSH and AAOS texts were written at approximately the 9th-10th grade and 8th-9th grade levels, respectively.
- Gemini 2.5 and ChatGPT-5.0 ASSH PEMs were scored at approximately grade 6, reducing SMOG by 2.59 and 2.88, and increased FRE by 24 and 18.1, respectively.
- Gemini 2.5 and ChatGPT-5.0 AAOS PEMs were scored at approximately grade 6, reducing SMOG by 1.56 and 1.93, and increased FRE by 19 and 14.4, respectively.

Conclusions

- This study demonstrates that large language models, specifically ChatGPT-5.0 and Gemini 2.5, can effectively simplify complex orthopedic patient education materials on CMC thumb arthritis to the recommended sixth-grade reading level and improve Flesch Reading Ease and Simple Measure of Gobbledygook scores
- Using AI language models may prove especially beneficial, as they can promote patient autonomy, support more informed decision-making, and help reduce anxiety.
- Understanding and comparing the strengths and weaknesses of AI models in this context can assist healthcare providers in distributing high-quality, easy-to-read PEMs to English-speaking patients.
- Clinically, this has important implications for reducing anxiety, improving adherence to treatment plans, and advancing health equity among patients with limited health literacy.

References

• Clear & simple [Internet]. U.S. Department of Health and Human Services; 2025 [cited 2025 Aug 25]. Available from: <https://www.nih.gov/institutes-nih/ni-h-office-director/office-communications-public-liaison/clear-communication/clear-simple>

• Eltorai AE, Shama P, Wang J, Daniels AH. Most American Academy of Orthopaedic Surgeons' online patient education material exceeds average patient reading level. Clin Orthop Relat Res. 2015 Apr;473(4):1181-6. doi: 10.1007/s11999-014-4071-2. Epub 2014 Dec 5. PMID: 25475715; PMCID: PMC4353543.

• Shahid R, Shoker M, Chu LM, Frehlick R, Ward H, Pahwa P. Impact of low health literacy on patients' health outcomes: a multicenter cohort study. BMC Health Serv Res. 2022 Sep 12;22(1):1148. doi: 10.1186/s12913-022-08527-9. PMID: 36096793; PMCID: PMC9465902.

• Kirchner GJ, Kim RY, Weddle JB, Bible JE. Can Artificial Intelligence Improve the Readability of Patient Education Materials? Clin Orthop Relat Res. 2023 Nov 1;481(11):2260-2267. doi: 10.1097/CORR.0000000000002668. Epub 2023 Apr 28. PMID: 37116006; PMCID: PMC10566892.