



Abstract: 4513

Scientific Abstracts > Education

USING GENERATIVE AI TO IMPROVE PATIENT EDUCATION MATERIALS FOR INTERVENTIONAL PAIN PROCEDURES

Mohan Ravi, Annabel Chen, Michael Leong, Vafi Salmasi
Stanford University

Introduction

Internet-based education materials are becoming increasingly important funds of knowledge for patients [1]. However, most online education materials have been found to exceed the recommended 6th grade reading level [2]. In this study, we find that generative AI can be used to create patient education documents about dorsal root ganglion stimulation. These documents are more accessible than human-authored counterparts and can be created in a fraction of the time without burdening clinician's already heavy workloads.

Materials and Methods

To obtain available education materials, a Google search was performed using a browser without prior search history for the terms "dorsal root ganglion stimulation patient education". To create new materials, the publicly available research preview of ChatGPT was prompted with "Write a patient education document for dorsal root ganglion stimulation. Make it simple enough that a third/sixth/ninth grader could read it". Each prompt was provided five times, to which generated fifteen unique outputs. Readability was assessed using five previously validated algorithms: Flesch-Kincaid Grade Level, Coleman-Liau Index, Simple Measure of Gobbledygook Index, Automated Readability Index, and Linsear-Write Formula. Comprehensiveness of each article was rated from 0-5 points predetermined metrics, with one point given if the article contained information on indications, contraindications, anatomic targets, procedure description, and complications. As the case report is devoid of patient identifiable information, it is exempt from IRB review requirements as per Stanford policy.

Results/Case Report

Human-authored patient education materials were written at a mean reading level of 12.6 (mean [SD], 12.6 [1.9]; range, 6.2-16.01), with only one of fifty resources (2%) meeting the target reading level. We found that ChatGPT wrote at a higher grade level than prompted. When asked to write at a 6th grade reading level, ChatGPT-authored documents were within the target reading level 50% of the time, with a mean 8th grade reading level (7.98 [1.4]). Readability scores were significantly different between ChatGPT and human-authored documents ($p < 0.005$). When writing at a 3rd grade level, ChatGPT produced documents within the target reading level 90% of the time, with a mean reading level of 7th grade (6.8 [0.7]). There was no statistical difference in comprehensiveness between human and ChatGPT authored texts. Human-authored documents had mean comprehensiveness of 2.76 (0.91), while

ChatGPT authored texts at the 3rd grade level had comprehensiveness scores of 2.8 (1.095).

Discussion

We found that ChatGPT writes at a higher grade level than prompted. Considering current GPT relies heavily on extant human authored literature, it is no surprise that the generative model—like the human text it is based on—both provide higher reading levels than desired. The authors recommend use of a deliberately lower reading level to compensate for this observed effect.

References

1. Rainie L, Fox S. The online health care revolution. Pew Research Center website. November 26, 2000. <http://www.pewinternet.org/2000/11/26/the-online-health-care-revolution/>
2. Ad Hoc Committee on Health Literacy for the Council on Scientific Affairs, American Medical Association. Health Literacy: Report of the Council on Scientific Affairs. *JAMA*. 1999;281(6):552–557. doi:10.1001/jama.281.6.552

Disclosures

No

Tables / Images

