Application of OpenAI Chat-GPT for Patient Accessible Lay Radiology Reports

Hanssen Li, MD, Resident Physician, Emory University
Zach Zaiman, BS; John Moon, MD; Imon Banerjee, PhD; Judy Gichoya, MD, MS; Hari Trivedi, MD

Introduction

The 21st Century Cures Act’s mandate for immediate release of electronic health information allows patients to view their radiology reports often ahead of physicians explaining their significance. These reports can contain complex verbiage beyond the 8th grade reading level of the average American and thus can generate significant patient confusion and anxiety. We explored the application of a natural language processing (NLP) model called ChatGPT (OpenAI, California, USA, https://chat.openai.com/chat), which recently garnered widespread attention on the internet, to simplify radiology reports to layman’s terms.

Hypothesis

OpenAI’s ChatGPT can accurately simplify radiology reports to layman’s terms and can significantly reduce the reading difficulty.

Methods

We extracted 20 CT abdomen and pelvis radiology reports from our institution’s database with varying common pathologies using search terms “appendicitis, colitis, cyst, diverticulosis, IPMN, portal venous thrombosis, renal atrophy, steatosis, cholecystitis, and no acute abnormalities.” These were processed by ChatGPT using the prompt “Simplify this radiology report to layman’s terms at the middle school level” in addition to the full report. Flesch-Kincaid reading level (0 to 18+) and Flesch reading ease scores (0-100, where 100 is the easiest to understand) were calculated with Microsoft Word’s Editor Tools (Microsoft, Washington, USA) for each original report and simplified ChatGPT output. Student’s t-test was used to determine the significance of the difference in the scores.

Results

The 20 extracted radiology reports consisted of an average of 239 words (standard deviation - SD 105) with mean Flesch reading ease score of 9.2 (SD 10.6) and mean Flesch-Kincaid reading level of 16.7 (SD 2.7). The simplified ChatGPT output consisted of an average of 105 words (SD 26) with mean Flesch reading ease score of 66.1 (SD 9.9) and mean Flesch-Kincaid reading level of 7.7 (SD 1.8). This reflects a mean decrease of 133 words (p < 0.001), increase in Flesch reading ease score of 53.9 (p < 0.001), and decrease in Flesch-Kincaid reading level of 9.0 (p < 0.001).

Conclusion

OpenAI’s ChatGPT can significantly reduce the reading level of radiology reports by simplifying reports to a concise summary statement at the 8th grade reading level of the average American.
Figure(s)

This radiology report describes the results of a CT scan of the abdomen and pelvis. The scan was done with a contrast agent to make the details of the organs and tissues easier to see. The scan showed that the patient's lungs were clear, but that they had some bibasilar atelectasis (collapsed air sacs in the lower part of the lungs). The liver, spleen, and pancreas were normal, but the scan showed a well-circumscribed, hypodense lesion (an area of decreased density) in the left liver lobe that is likely a hepatic cyst or biliary hamartoma (a non-cancerous growth in the liver). The gallbladder, kidneys, and bladder were normal. The patient had air-fluid levels in their proximal colon, which may be due to diarrhea. No other abnormalities were found.

Figure 1. Sample ChatGPT output of a CT abdomen and pelvic with incidental findings as well as evidence of a diarrheal illness.

Keywords
Applications; Artificial Intelligence/Machine Learning; Emerging Technologies