WhatGPT?
Assessing Medical Student Use of Language Models During Clinical Clerkships

Rafael C Schulman[1], Ali S Razavi[1], Jonathan Chung[1], Jason Grullon[1]
[1]Norton College of Medicine, SUNY Upstate Medical University, 750 East Adams Street, Syracuse NY, 13210

Background.

Recent innovations in machine learning model architecture and training approaches have resulted in significant advancements in large language models (LLMs). These developments, highlighted by generative models, are poised to have a profound impact on the medical field. Emerging tools will affect how we integrate medical information, interact with patients, and maintain medical records. In addition to new opportunities, this rapid innovation brings an impetus to consider the ethical, legal, and logistical implications of emerging technologies. To do this effectively, it will be critical to understand how exactly new tools are being adopted by medical practitioners.

Medical students on clinical clerkships are in the earliest stage of their clinical career. These students take regular clerkship exams, which motivates them to seek out the most effective and up to date learning tools. Additionally, students on clerkships are constantly under evaluation and are therefore uniquely incentivized to adopt tools and techniques that benefit their performance in the clinical setting. These two reasons make it likely that students will be early adopters of LLMs adapted for clinical application.

We conducted a survey of third year medical students to evaluate familiarity with ChatGPT. ChatGPT has been one of the most publicized LLMs released to date, launched as a prototype in November of 2022. Although the model was trained on a general corpus of data, it can successfully discuss clinical vignettes, explain pathophysiology of disease, and generate differential diagnoses. The students surveyed have completed two years of preclinical medical coursework, and are currently on clinical rotations in internal medicine, surgery, and other specialties.

Methods.

We distributed a survey to third year students at a SUNY Upstate Medical University in Syracuse, New York. The survey consisted of seven questions: two about preferred clinical and study software, four about ChatGPT, and one open ended question inviting participants to comment. The survey was shared with the class on March 31st 2023, and responses were accepted for one week. 46 students, approximately one third of the third-year class, completed the survey.

Results.

Familiarity with ChatGPT:

Of the respondents, 74% (34) were familiar with ChatGPT. Of the remaining 12 students, 8 heard of ChatGPT but didn't know what it was, and 4 had never heard of it.

Use of ChatGPT:

Only 34% (20) of students reported having ever used ChatGPT. Although most students had never used ChatGPT, others had found ways to integrate it into their clinical education. Among the 20 students who had used ChatGPT, five students said that they utilized ChatGPT for studying and exam preparation. Four reported having used ChatGPT during work in the clinic.

Preferred tools:

Additionally, we asked students what tools they use as sources of information during clerkships, both when in the clinic and when studying for exams. For information while on the wards, students prefer UpToDate (used by 84%), Amboss (50%), and Google (45%). For exam preparation, the most popular tools were Amboss (67%), UWorld (50%), and UpToDate(30%).

Comments from Students:

At the end of the survey, students were asked if they had any comments, and their responses revealed an array of attitudes towards these emerging technologies as well as some interesting applications of LLMs. Comments included "ChatGPT could be very helpful with elaborating on practice question explanations", "I've experimented with having ChatGPT generate practice questions", "AI scares me", and "Chat GPT has no sense on its own uncertainty".

Conclusion.

Large language models present exciting new opportunities for the medical field. With the rapidly evolving landscape it will be critical to understand where and how emerging tools are being applied. Surveying practitioners, including students, may be an effective way to gain insights to what the future looks like.

Here, a cross-sectional survey of one medical class shows that students are familiar with ChatGPT but most have not used it. However, a small number of students have begun to integrate the tool, indicating that more tailor-made LLM applications might find fast adoption by some in the medical community.