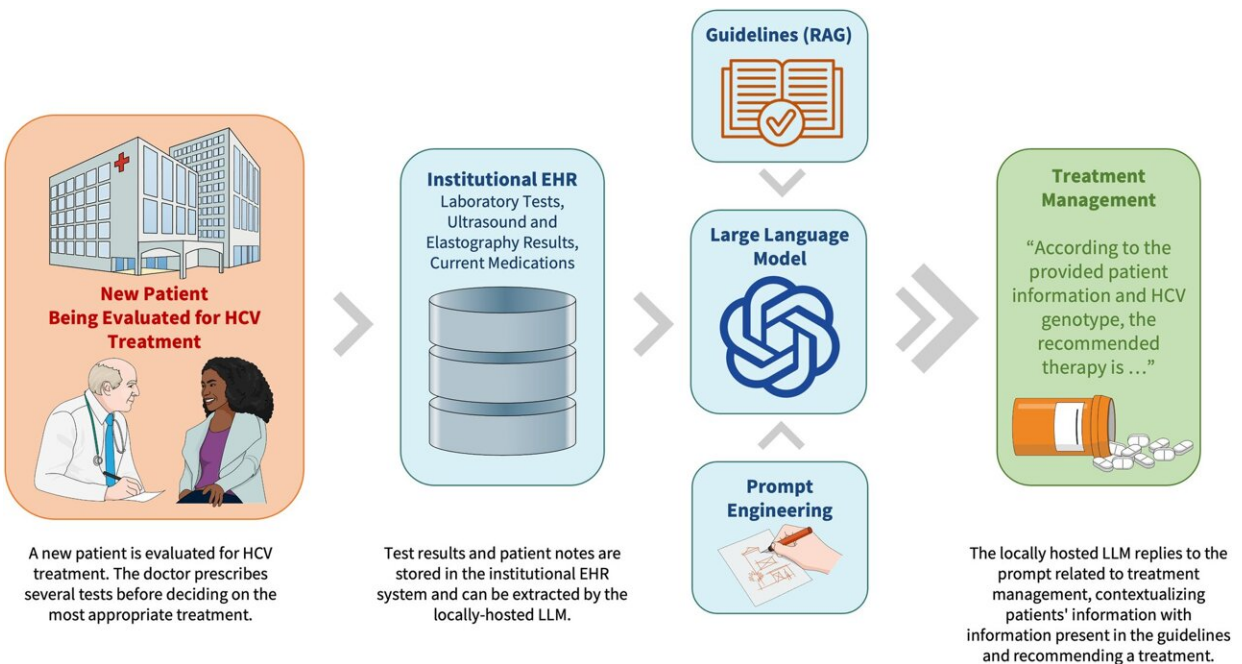


Making clinical guidelines work for large language models

July 19 2024, by Rachel Martin



Example of a clinical decision support system integrated with large language models. Credit: *npj Digital Medicine* (2024). DOI: 10.1038/s41746-024-01091-y

Clinical guidelines are essential to the practice of evidence-based medicine, but they are long and complex, which makes it hard for busy doctors to quickly and easily find the information they need to care for each patient.

Faculty members across Yale's Department of Internal Medicine are researching different approaches to make [clinical guidelines](#) more accessible to clinicians by incorporating them into existing tools and workflows. Tools that rely on [large language models](#) (LLMs) to generate responses to clinical questions are promising in their easy usability and ability to respond to physician queries.

"While we were developing an LLM [tool](#) to help clinicians answer questions about hepatology and gastrointestinal conditions, we realized that LLM companies often automatically convert clinical guidelines from a PDF to a text document," said Dennis Shung, MD, Ph.D., assistant professor of medicine (digestive diseases). "But when you automatically convert these guidelines, you lose important data that is essential for clinical reasoning."

Dennis and his team noticed data loss was especially dramatic in tables, graphics, and flowcharts. LLM accuracy rates were about 80% if a table contained only text. If a guideline included a graphic, accuracy dropped to about 16%. If the guideline included a flowchart, accuracy dropped to nearly zero.

"This is especially concerning because sometimes the most important information in a guideline is in a flowchart," said Mauro Giuffrè, MD, postdoctoral associate (digestive disease). "For LLM tools to be helpful to clinicians, they must be capable of understanding all the data—not just pieces of it."

Shung and Giuffrè [published](#) a paper titled "Optimizing Large Language Models for Medical Guidelines Interpretation: A Framework Based on Study on Hepatitis C Virus Guidelines Using Retrieval Augmented Generation" in *njp Digital Medicine* that looked at how different ways of formatting the text of the Hepatitis C Virus guideline improved the accuracy of the LLM tool so it could be more useful to clinicians.

Simone Kresevic, postgraduate fellow (digestive diseases) and Ph.D. candidate at the University of Trieste, spearheaded the software engineering effort to test the different LLM configurations. Other collaborators included Milos Ajcevic and Agostino Accardo, professors of biomedical engineering at the University of Trieste, and Lory S. Crocè, professor of gastroenterology and hepatology at the University of Trieste.

"We started with no formatting at all—just a straight PDF to text, formatted with all its problems," Giuffrè said. "Then we added more labels and specificity to give the LLM more information about each figure or table, and then asked the LLM to answer questions about caring for patients with Hepatitis C."

They found that appropriately formatting text, figures, and tables in the clinical guidelines allowed the LLM model to reason more easily over the data, and accuracy improved dramatically.

Shung and Mauro say these findings could apply to other guidelines and specialties.

"LLMs are only as good as the information they are trained on," said Shung. "Using LLM-friendly versions of clinical guidelines could help us more quickly develop point-of-care tools that provide highly accurate and [relevant information](#) for clinicians."

Ultimately, Shung and Mauro hope medical societies will create versions of LLM-friendly guidelines so LLMs can easily ingest the guidelines without the need for reformatting.

"Medical societies want their members to practice evidence-based medicine, which requires access to the best information at the right time for each patient," said Shung.

"By creating LLM-friendly clinical guidelines, medical societies can help us create LLM tools that are up-to-date, complete, and use reputable sources. Clinicians need to have confidence in the information to make the best decisions with the patient in front of them."

More information: Simone Kresevic et al, Optimization of hepatological clinical guidelines interpretation by large language models: a retrieval augmented generation-based framework, *npj Digital Medicine* (2024). [DOI: 10.1038/s41746-024-01091-y](https://doi.org/10.1038/s41746-024-01091-y)

Provided by Yale University

Citation: Making clinical guidelines work for large language models (2024, July 19) retrieved 19 July 2024 from <https://medicalxpress.com/news/2024-07-clinical-guidelines-large-language.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.