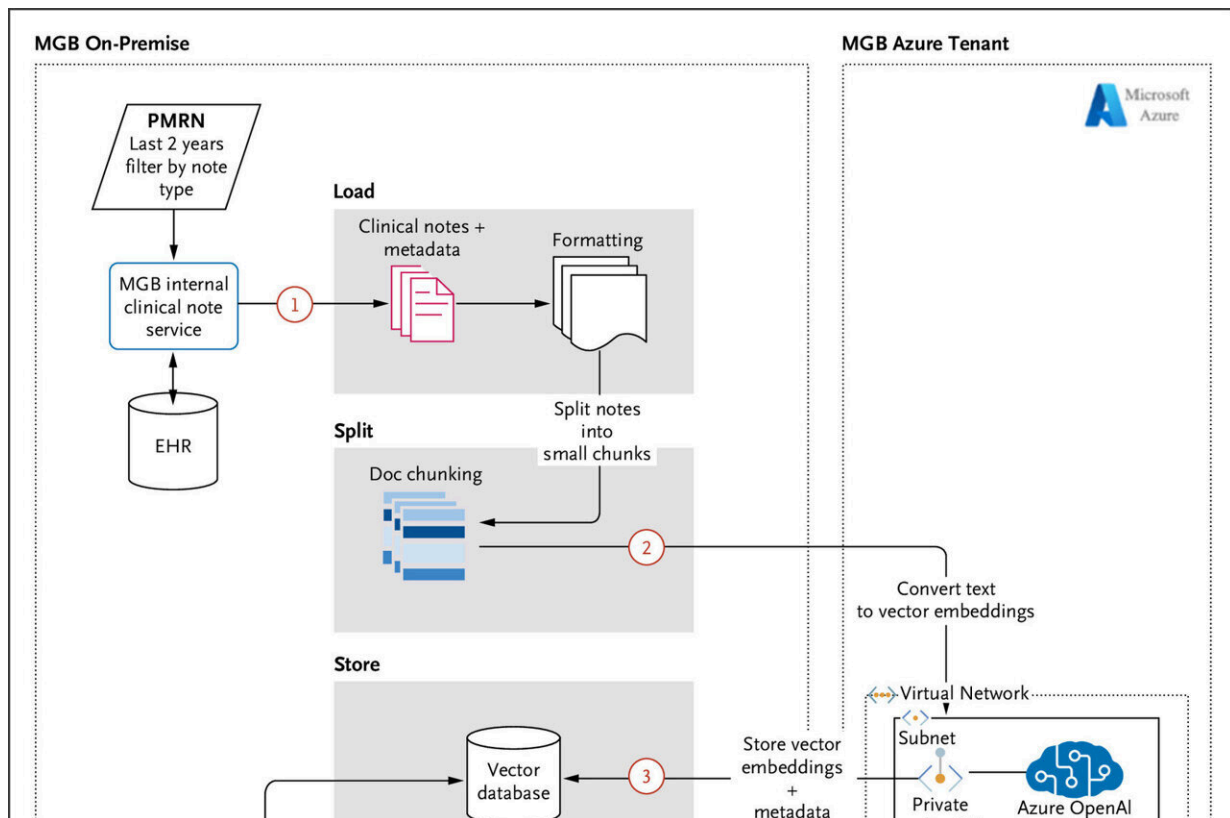# Researchers say AI accurately screens heart failure patients for clinical trial eligibility

June 17 2024

Workflow of the clinical note–based question-answering system powered by the rag architecture. Credit: *NEJM AI* (2024). DOI: 10.1056/AIoa2400181

Generative Artificial Intelligence (Gen AI) can rapidly and accurately screen patients for clinical trial eligibility, according to a new study from

Mass General Brigham researchers. Such technology could make it faster and cheaper to evaluate new treatments and, ultimately, help bring successful ones to patients.

Investigators assessed the accuracy and cost of a Gen AI process they named RAG-Enabled Clinical Trial Infrastructure for Inclusion Exclusion Review (RECTIFIER), that identifies patients who meet criteria for enrollment in a heart failure trial based on their medical records. For criteria that require reviewing patient notes, they found that RECTIFIER screened patients more accurately than disease-trained research coordinators who typically conduct screening—and for a fraction of current costs.

The results of the study were [published](link) in *NEJM AI*.

"We saw that large language models hold the potential to fundamentally improve clinical trial screening," said co-senior author Samuel (Sandy) Aronson, ALM, MA, executive director of IT and AI Solutions for Mass General Brigham Personalized Medicine and senior director of IT and AI Solutions for the Accelerator for Clinical Transformation.

"Now the difficult work begins to determine how to integrate this capability into real world trial workflows in a manner that simultaneously delivers improved effectiveness, safety, and equity."

Clinical trials enroll people that meet specific criteria, such as age, diagnoses, key health indicators and current or past medications. These criteria help researchers ensure they enroll participants who are representative of those expected to benefit from the treatment. Enrollment criteria also help researchers avoid including patients who have unrelated health problems or are taking medications that could interfere with the results.

Co-lead author Ozan Unlu, MD, a fellow in Clinical Informatics at Mass General Brigham and a fellow in Cardiovascular Medicine at Brigham and Women's Hospital, said, "Screening of participants is one of the most time-consuming, labor-intensive, and error-prone tasks in a clinical trial."

The research team, part of the Mass General Brigham Accelerator for Clinical Transformation, tested the ability of the AI process to identify patients eligible for The Co-Operative Program for Implementation of Optimal Therapy in Heart Failure (COPILOT-HF) trial, which recruits patients with symptomatic heart failure and identifies potential participants based on electronic health record (EHR) data.

The researchers designed 13 prompts to assess clinical trial eligibility. They tested and tweaked those prompts using medical charts of a small group of patients, before applying them to a dataset of 1,894 patients with an average of 120 notes per patient. They then compared the screening performance of this process to those of study staff.

The AI process was 97.9% to 100% accurate, based on alignment with an expert clinician's "gold standard" assessment of whether the patients met trial criteria. In comparison, study staff assessing the same medical records were slightly less precise than AI, with accuracy rates between 91.7% and 100%.

The researchers estimated the AI model costs about $0.11 to screen each patient. The authors explain that this is orders of magnitude less expensive compared to traditional screening methods.

Co-senior author Alexander Blood, MD, a cardiologist at Brigham and Women's Hospital and associate director of the Accelerator for Clinical Transformation, noted that using AI in clinical trials could speed up the time it takes to determine whether a therapy is effective. "If we can

accelerate the clinical trial process, and make trials cheaper and more equitable without sacrificing safety, we can get drugs to patients faster and ensure they are helping a broad population," Blood said.

The researchers noted that AI can have risks that should be monitored when it is integrated to routine workflows. It could introduce bias and miss nuances in medical notes. Plus, a change in how data is captured in a health system could significantly impact the performance of AI.

For these reasons, any study using AI to screen patients needs to have checks in place, the authors concluded. Most trials have a clinician who doublechecks the participants whom the study staff deem eligible for a trial, and the researchers recommended that this final check continue with AI screening.

"Our goal is to prove this works in other disease areas and use cases while we expand beyond the walls of Mass General Brigham," added Blood.

In addition to Unlu, Blood and Aronson, Mass General Brigham authors include co-first author Jiyeon Shin (BWH, MGB), Charlotte Mailly (BWH, MGB), Michael Oates (BWH, MGB), Michaela Tucci (BWH), Matthew Varugheese (BWH), Kavishwar Wagholikar (BWH), Fei Wang (BWH, MGB), and Benjamin Scirica (BWH).

Citation: Researchers say AI accurately screens heart failure patients for clinical trial eligibility (2024, June 17) retrieved 26 June 2024 from https://medicalxpress.com/news/2024-06-ai-accurately-screens-heart-failure.html