

Accelerating Clinical Trials with Large Language Models and Agentic AI



Danielle Bitterman, MD

Radiation Oncologist, Brigham and Women's Hospital; Assistant Professor of Radiation Oncology, Harvard Medical School

dbitterman@bwh.harvard.edu



Curating clinical trial outcomes is costly, inefficient, and error-prone. These challenges inflate clinical trial expenses, delay important trial-related decisions, compromise study validity, undermine trust in clinical outcomes, and represent a large proportion of a trial's costs. Improving outcome reporting efficiency in real-time could multiply the number of trials run across healthcare sites and speed the delivery of high-quality data, bringing better treatments to our patients, faster.

Generative artificial intelligence (AI), including large language models (LLMs), provides a new approach to address the challenge of clinical trial adverse event abstraction and reporting. However, LLMs in isolation cannot do this with the performance and reliability required by clinical trials.

To address this, we have developed an agentic AI system, where multiple LLMs collaboratively conduct clinical reasoning over patients' clinical notes to automate adverse event detection, severity grading, and attribution. Our system uses guidelines and study protocols to output standardized outcomes that adhere to regulatory and compliance standards.

Compared to expert oncologist determination, our agentic AI system detects adverse events with an accuracy of 98%. Our system is more accurate than clinical research coordinators (66% vs. 98% accuracy, respectively). Our system can process a note in approximately 10 seconds, compared to 10 minutes required for fully manual abstraction. When used to assist clinical research coordinators, they require >40% less time to collect trial data. In qualitative interviews, clinical research coordinators report that using the AI agentic system is acceptable, usable, and improves their experience abstracting adverse events. A real-time pilot of our system is underway at Mass General Brigham. To our knowledge, this is the first agentic AI system fully implemented in a healthcare system for real-time trial reporting.

Our promising results and demonstrated success implementing our agentic AI system for real-time clinical trial reporting shows the significant potential of our system to accelerate clinical trials. In next steps, we will expand our system to other clinical trial outcomes to improve the cost, efficiency, and quality of clinical trials.

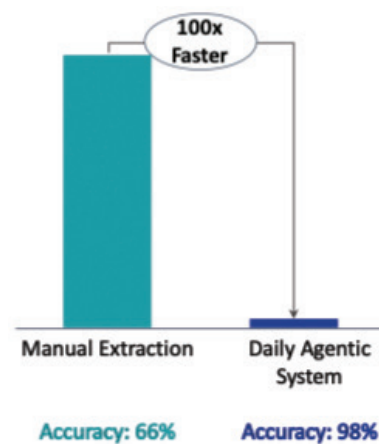


Figure 1: Our agentic AI system for clinical trial adverse event reporting is 100x faster than manual reporting, with improved accuracy.