ABS25-15262. Evaluating a doctor-in-the-loop AI-agent for multiple myeloma chart reviews

Authors: Karla Mariana Castro Bórquez[1], Juan Pablo Capdevila[1], Jennifer Ahlstrom[1], Jay Hydren[1], Alejandro Lopez[1], Jorge Alberto López Carlos[1], Andrea Victoria Pacheco[1], Ana Echenique Alcázar[1], Edgar Ricardo Olivas Domínguez[1], Eduardo Franco Hernandez[1], Miranda De la Luz Franco[1], Martha Paola Anchondo Commesse[1], Magaly Valeria Escobedo Cruz[1], ANA Sahagun[1], Jorge Arturo Hurtado Martinez[1]

Organizations: [1]HealthTree Foundation, Lehi, UT, United States.

BACKGROUND

Real-World Data (RWD) is increasingly vital in healthcare, enabling personalized care and decision-support tools through derived variables like **lines-of-therapy (LOTs) and treatment outcomes.** However, extracting this data is challenging because it is spread across hundreds of unstructured clinical notes. **The traditional manual chart review process is costly and time-consuming**, creating a major bottleneck in research. To address this, we developed a scalable framework: an **AI Agent with a doctor-in-the-loop** proposed to extract multiple myeloma (MM) lines of therapy and outcomes from electronic health records. **This work summarizes the development and evaluates the performance of this AI agent.**

METHODOLOGY

An **AI Agent with a doctor-in-the-loop** was designed to extract MM LOTs and outcomes. A cohort of 94 randomly selected patients diagnosed with MM since 2022 via Healthtree.

Step 1 - Document Filtering: OpenAI's o4-mini model was used to independently evaluate each clinical note to identify documents relevant to LOT and outcomes.

Step 2 - Data Extraction: Gemini 2.5 PRO (selected for its 1 million token context window) processed the aggregated, relevant documents to extract five key LOT elements:

- Medications, Start/End Dates, Procedures, IMWG Outcomes, and Treatment Status.
- The model also returned database IDs of supporting source documents for traceability.

Step 3 - Validation: Clinicians reviewed the model's output in a structured interface, displaying the extracted values and the original source document texts.

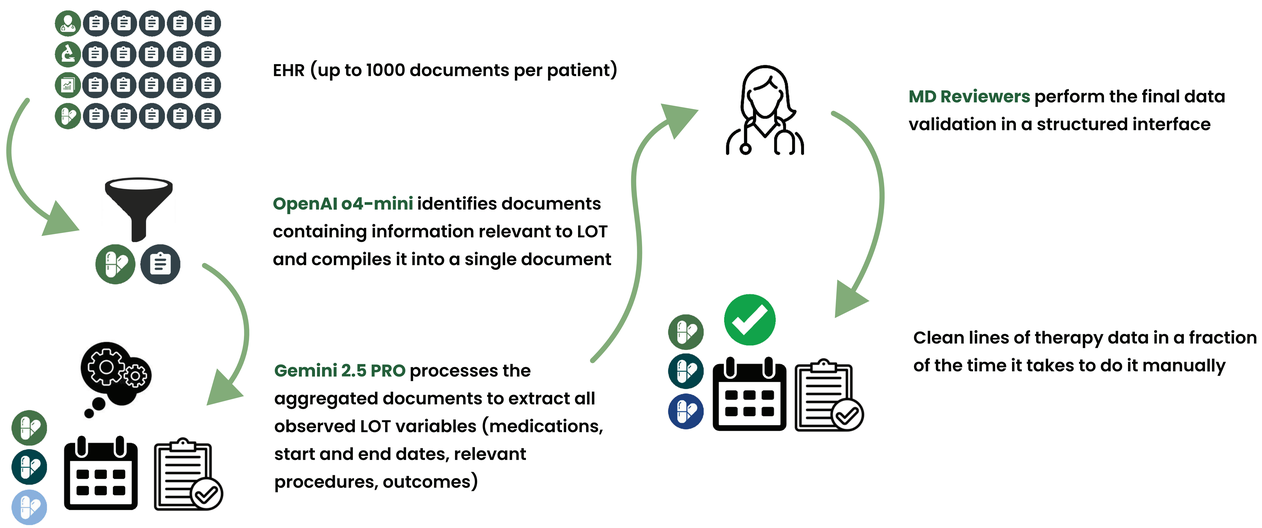
Each element was verified, corrected if needed and model performance was evaluated by calculating element-level accuracy as the percentage of model-generated outputs that matched

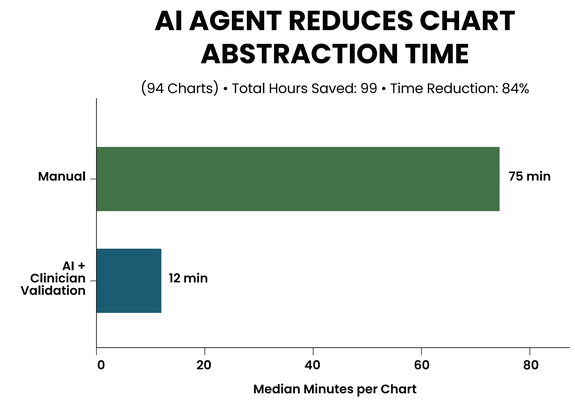
CONCLUSION

This study demonstrates the feasibility of combining frontier LLMs with doctor-in-the-loop validation for the extraction of MM LOTs from unstructured clinical documents.

- The **high fidelity** across the treatment elements analyzed, combined with **rapid clinician validation and document traceability**, positions this approach as a scalable alternative to manual abstraction.
- The **AI Agent addresses long-standing bottlenecks in RWD curation** by offering a reliable, low-burden method for transforming free-text data into structured treatment trajectories.
- This framework has potential applications in trial eligibility matching, retrospective outcomes research, and **real-world clinical decision support** across various EHR systems.

RESULTS





ACKNOWLEDGEMENT

Our deepest gratitude to the patients and caregivers who share their data and experiences through HealthTree. Your contributions make this research possible.

For more information on clinical research collaboration, please contact research@healthtree.org

Jay: Receive research funding from Adaptive Biotechnologies, BioLinRx, Gamida-Cell, Sanofi, GlaxoSmithKline, Regeneron, Pfizer, Johnson and Johnson, and Takeda Oncology. Jenny: Patient advocacy committee: Pfizer, BMS, Janssen, Takeda Oncology, and Sanofi.

Scan to learn more about how HealthTree can accelerate your research