O-12

Assessment of AI-assisted Clinical Trial Matching for Neuroendocrine Neoplasms

Kenneth L Chen[1], Josh A Mailman[2], George A Fisher[3], Shagufta Shaheen[3]

[1]Division of Oncology and Division of Hematology, Department of Medicine, Stanford School of Medicine; [2]NorCal CarciNet; [3]Division of Oncology, Department of Medicine, Stanford School of Medicine

BACKGROUND

For many patients with neuroendocrine neoplasms and their physicians, finding an appropriate clinical trial can be difficult.

METHODS

We characterized the performance of clinical trial matching using common techniques: search engine query (Google), free generalist LLM chatbots (ChatGPT, Gemini 3.5 Flash, Claude Sonnet 5 Max), clinical trial specific AI (ancora.ai), and ClinicalTrials.gov. For search engine query and generalist chatbots, we wrote: “neuroendocrine cancer clinical trials near me.” For clinical trial specific AI and ClinicalTrials.gov, we selected neuroendocrine neoplasms from a dropdown menu. We characterized trial results into possibly relevant “true” matches or irrelevant “false” matches. Trials that were non-therapeutic, clearly labeled non-recruiting, or greater than 100 miles away were excluded. We characterized the overlap and uniqueness of results from each source. We then characterized the specificity of the clinical trial specific AI by searching as a patient in three hypothetical cases, with a moderate level of clinical knowledge (organ of origin, metastatic status, and last therapy). We then broadened our search to the entire United States, comparing generalist against clinical-trial specific AI for our hypothetical pancreatic NET case.

RESULTS

We found from search engine query of local clinical trials returned 8 true matches and 7 false matches. Fewer results from generalist LLM chatbots (1-4 true and 0 false matches), most comprehensive results from clinical trial specific AI (15 true and 0 false matches), and poor signal-to-noise ratio from ClinicalTrials.gov (12 true and 29 false matches). Four trials were reported by all three modalities (search engine, trial specific AI, & ClinicalTrials.gov), one by search engine query alone, seven by clinical trial AI alone, and none by ClnicalTrials.gov alone. When search criteria were narrowed to three specific cases with knowledge of organ of origin, metastatic status, and last treatment, the clinical trial specific AI returned only 0-1 possibly relevant and 1-2 irrelevant trials. On our broadened nationwide AI-driven search, generalist LLMs produced a range of results (no

specific trial listing, 0 true with 2 false matches, 6 true with 0 false matches), and clinical trial specific AI returned 5 true with 4 false matches.

CONCLUSIONS

Search engine query provided admirably robust results though required labor-intensive parsing of false matches. ClinicalTrials.gov provided noisy search results with many irrelevant trials. AI-assisted clinical trial search provided variable results depending on the specifics of the query.

ABSTRACT ID #36660