innovate GenAI Use Case

OM1

Using AI and Unstructured Data to Fill in Gaps in the Patient Journey

Background:

Understanding patients' disease activity trajectories is critical to evaluating safety, efficacy, and unmet need. Medical Affairs professionals need clear visibility into these trajectories as they evolve over time, and within patient cohorts. High-quality, regular data capture is certainly present in clinical trials, and at times in other protocolized cohort studies. Unfortunately, the same can't be said for patients' 'real-world' experiences – even though most of their lives are spent outside formal study contexts. In part, this challenge can be traced to the difficulty and time needed to gather data on outcomes. Assessments of disease activity can interrupt clinical workflows; self-assessments, while potentially less burdensome to complete, are often less reliable and suffer from poor completion rates. These data gaps lead to substantial 'missingness' in understanding what happens to patients in the long stretches they spend between more formal periods of data-gathering.

However, even in these gaps, data are present. They may not be perfect – organization may be challenging, for example, or format may be difficult to parse. Richness can still exist in these data, but historically, the difficulty has been in integrating them, extracting information, and synthesizing outputs into consistent, coherent endpoints that are usable for studies and conclusions. Manual review can accomplish this task, but is obviouosly not scalable. We have been attacking this problem using AI tools, including generative AI most recently. Gen AI is especially adept at 'rapid summarization' – for example, identifying content in a medical note around whether a patient's condition improved after starting a medication. These capabilities hold great promise for improving the quantity of available outcomes data without sacrificing quality, but must be tested carefully to ensure accuracy and generalizability.

Process:

Calibrating a model to extract or estimate disease activity requires a source dataset – in this case, medical notes from providers – and labeled outcomes. For example, in heart failure, cardiologist notes can be matched with available scoring on the New York Heart Association (NYHA)'s Classification scale, describing the severity of a patient's heart failure. With sufficient sample data, we can test different methods of best using the unstructured data available to estimate the relevant disease activity metric, and compare results against non-AI methods for accuracy and efficiency (for example, manual curation or chart review and estimation). Cross-testing among different datasets capturing the same clinical conditions provides validation of robustness. In addition, gen AI can be uniquely proficient at drawing from different data types at the same time for these kinds of extraction or estimation tasks. By providing

models with examples of full patient records containing many data types, for example, we can evaluate how useful additional data are when processed by these tools.

Outcomes:

We have investigated endpoint extraction and estimation across a number of endpoints in a range of diseases and therapeutic areas (e.g., NYHA Classification in heart failure; CDAI in rheumatoid arthritis; PHQ-9 in major depressive disorder; SLEDAI in systemic lupus erethymatosis). In each case, AI-based techniques have provided significant value, in speed, accuracy, or both, relative to baseline methods of understanding disease activity without these tools. Experimentation with gen AI has yielded interesting results in particular. First, as hoped, gen AI is able to process certain kinds of unstructured data and output disease activity measures in high volume, while preserving high precision – an excellent solution that can meaningfully increase structured data availability. Just as importantly, this performance does *not* generalize to all conditions and endpoints. In particular, those involving both concept identification and 'contextual categorization' – not simply understanding whether evidence of disease is present in a patient, for example, but what that means when given that patient's larger context – may be best suited for hybrid approaches where gen AI does some, but not all, of the work. For example, liver function in a patient with advanced nonalcoholic steatohepatitis will very likely appear much worse than that of a healthy comparator, but in the specific patient's context, could actually indicate stable (as opposed to worsening) disease. Finally, gen AI seems capable of picking up signals and evidence that human evaluators either do not understand, or ignore, consciously or not. In cases where this happens frequently, we could actually see improvements relative to the 'gold standard' – while vastly improving processing volume and speed.

Medical Affairs professionals will always benefit from more, and better, data on patients' real-world experiences. While still early, gen AI is already showing great promise in applications extracting and synthesizing meaningful outcomes data from less accessable, unstructured inputs. These inputs are available, but hugely underutilized. We've found that with the right combination of clinical and technical expertise, we can harness gen AI to begin realizing much of this value, with more developments coming soon.

For more information visit OM1.com.