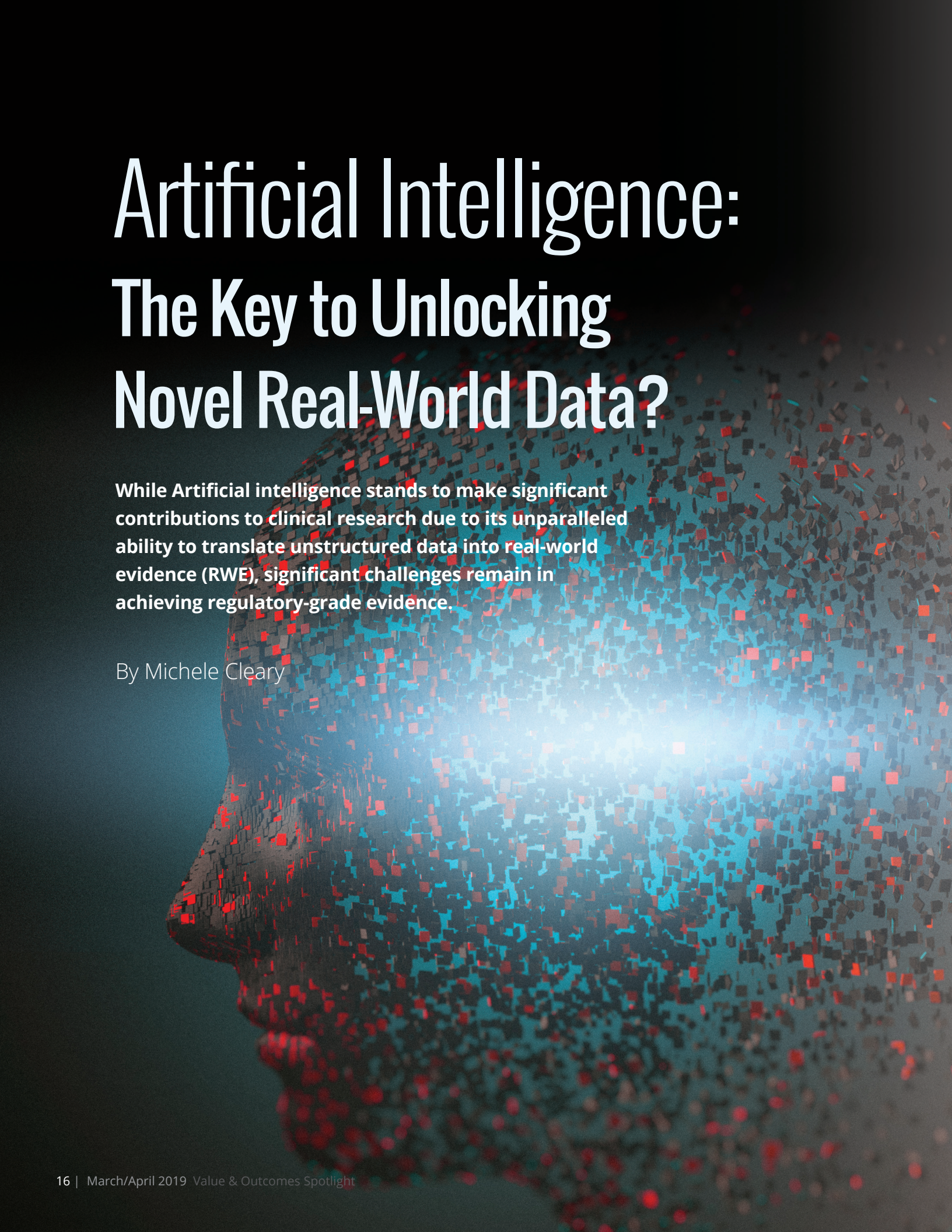
Artificial Intelligence: The Key to Unlocking Novel Real-World Data?

While Artificial intelligence stands to make significant contributions to clinical research due to its unparalleled ability to translate unstructured data into real-world evidence (RWE), significant challenges remain in achieving regulatory-grade evidence.

By Michele Cleary



Artificial intelligence (AI) is revolutionizing healthcare services. From improving disease detection to supporting treatment decision making, AI has become ubiquitous in care delivery.

Now AI is poised to transform the drug and device development process, helping researchers refine the approval process and significantly cutting both the time and the expense needed to bring products to market. While AI has long been used to facilitate recruitment of study subjects, optimize study design, and support patient adherence to study protocols, AI's greatest contribution to clinical research may still be on the horizon—unlocking the data richness that lies within the mountains of novel real-world data (RWD) sources.

This article explores how AI may improve clinical research through its ability to better translate RWD into real-world evidence (RWE), thus providing more valid evidence of clinical benefits and risks. Dan Riskin, MD, of Verantos, Rich Gliklich, MD of OM1, and Sebastian Schneeweiss, MD of Aetion all shared their valuable insights into how AI is transforming clinical research.

THE SEARCH FOR REGULATORY-GRADE DATA

With innovations in digital data, HEOR researchers are facing explosive growth in novel RWD sources. But as researchers move from traditional RWD sources (eg, registries and claims data) to these novel data sources, unstructured data present a significant opportunity and challenge. These novel data sources include doctor notes, discharge summaries, lab or imaging reports, and even social media posts. Some estimate up to 80% of electronic health record (EHR) data may be unstructured.[1,2]

The challenge lies in structuring RWD so valid clinical assertions can be made. AI may provide the key to unlocking these unstructured data, helping researchers identify clinically relevant data points critical to the approval process, which currently are not available in structured data fields. In addition to helping trial operations (eg, recruitment), AI can also help researchers process large volumes of disparate novel RWD to identify critical signals of clinical outcomes, including potentially new biomarkers or postmarketing safety signals.

Regulatory parties appear open to AI's growing role in the development process. The European Medicines Agency (EMA) recently presented its strategic goals regarding how AI may support regulatory decision making, proposing the need to develop AI capabilities to drive "collaborative evidence generation—improving the scientific quality of evaluations."[3] The US Food and Drug Administration (FDA) is also embracing the use of AI to expand the use of novel RWD, with former FDA Commissioner Scott Gottlieb recently stating, "Advancing real-world data into regulatory-quality real-world evidence is a key strategic priority for the FDA." ›

In addition to the FDA, other US agencies are welcoming AI. For instance, the Centers for Medicare and Medicaid Services (CMS) recently launched the CMS Artificial Intelligence Health Outcomes Challenge to support private AI innovation to improve the agency's predictive modeling practices.[4] And the Center for Drug Evaluation and Research's Office of Surveillance and Epidemiology (OSE) has been exploring ways that AI may improve the agency's ability to identify and prioritize drug-related adverse event reports.[5]

THE CHALLENGE OF DATA VALIDITY

Perhaps the greatest challenge in using RWD is in making valid clinical assertions. As these data sources are used more frequently not only to assess comparative effectiveness but also to make access determinations, data validity—data accuracy—becomes critically important. According to Dr Riskin, "If we're changing the standard of care based on clinical assertions, then data validity matters. And in our world, data validity breaks down to data accuracy and data generalizability."

Clinical assertions made from administrative claims data have long been known to carry uncertain validity. Billing codes do not necessarily represent clinical conditions, especially when upcoding occurs. Nor do they differentiate between ruling out a diagnosis or confirming a diagnosis. But per Dr Gliklich, "Some data are better than no data with respect to safety signal."

To improve data validity, data abstracters have been used to evaluate the unstructured data within other RWD sources, such as EHRs or imaging reports. When budgets allowed, multiple abstracters could be deployed to improve accuracy. But data accuracy is still limited by differences across abstracters. And given the time and budget demands, abstraction is rarely an efficient approach to achieving data validity.

AI can also help researchers process large volumes of disparate novel RWD to identify critical signals of clinical outcomes, including potentially new biomarkers or postmarketing safety signals.

MAKING RWD USABLE THROUGH AI

AI can improve the validity of clinical assessments derived from novel RWD through natural language processing (NLP) and machine learning (ML).

NLP is a common first step to AI. NLP involves linking words, phrases, and terms listed within unstructured data (physician notes) to indicate a specific condition or event. For instance, physicians may use a mix of terms to indicate the patient has had a recent myocardial infarction (MI): "heart attack," "MI," "myocardial infarc." NLP must differentiate between "rule out MI" and confirm MI. Common NLP techniques used to abstract these clinical indicators include simple word-based models for text classification, structured models for syntactic parsing (recognizing a sentence and assigning a syntactic structure to it), collocation finding (finding sequence of words or terms which co-occur more often than would be expected by chance), word-sense disambiguation (identifying which sense of a word is used in a sentence), and machine translation (translation of text by computer with no human involvement). However, while an improvement over billing codes, NLP provides only marginal improvement in data accuracy over abstraction.

Between NLP and ML lies inference, whereby computer programs search for patterns across data sources to infer a condition. For instance, searching a patient's EHR, the program may find troponin, EKG changes, chest pain—signs of a probable MI.

THE IMPORTANCE OF DESIGN

While many life science companies are currently using NLP techniques in their drug approval research, and some have introduced inference methods, these AI-aided results may still be insufficient to show an effect size.

Dr Riskin proposes going further by incorporating ML, arguing that NLP could achieve data accuracy levels of approximately 85%. If the effect size in the study is a 10% to 20% difference in groups, that level of data accuracy will be insufficient. ML including pattern recognition increases accuracy levels above 90%, sufficient to make valid clinical assertions.

ML offers the most sophisticated analysis, utilizing algorithms and statistical models to simulate human learning. ML algorithms may include patient demographic data, such as age, gender, and disease history, as well as relevant disease-specific data, such as diagnostic imaging, gene expressions, physical examination results, clinical symptoms, or medications. ML has been used extensively in oncology and immunology, translating imaging and digital pathology into usable clinical data that help clarify treatment choices and transform oncology care. The use of ML in these disease areas is not unexpected given their reliance on imaging and genetic data—deciphering these types of data is an AI strength.

Dr Schneeweiss identified 2 key ML use cases. The first is causal inference, for which ML would help identify additional covariates and new causal inference techniques, such as collaborative targeted maximum-likelihood estimates. The second use case is for predictive analytics, where ML could help target those patients who may best respond to a given treatment.

One key advantage of ML is its ability to operate on numerous predictive features in datasets including outliers, noise, and collinearities, without the stability and reliability concerns of traditional statistical modeling. This enables complex patterns and interactions to be identified. Using pattern similarities between patients with or without a given diagnosis, this approach can confirm a diagnosis in patients for whom the disease is present but is undiagnosed or underdiagnosed. As the volume of RWD continues to grow, so will the demand for sound ML. For as Dr Schneeweiss emphasized, "The less structured the information is, the more helpful machine learning will be."

WHAT IS NEEDED FOR GOOD AI?

To fully employ the benefits of AI, computing prowess alone is not sufficient.

Dr Riskin argued that good AI requires the right technologies, good scientific design, and the right data sources.

By the right technologies, he called for using all 3 AI approaches discussed here—NLP, inference, and ML. Next, he emphasized that good scientific design, namely knowing the expected effect size in advance through proper study design, is critical. This step identifies what the level of required accuracy will be in advance of running the study and then checking accuracy during the study. While rigorous chart abstraction and NLP may improve specificity, it may not improve sensitivity, so both aspects of accuracy need to be protocolled and tested. Finally, he strongly encouraged the use of more-advanced data sets with rich, unstructured data linked at the patient level to extend research capabilities beyond those allowed with traditional RWE (eg, registries) or randomized controlled trials (RCTs).

While many life science companies are currently using NLP techniques in their drug approval research, and some have introduced inference methods, these AI-aided results may still be insufficient to show an effect size.

ACCEPTANCE OF AI DATA

Currently, the FDA is evaluating which models may be appropriate for regulatory safety or approval decision making. Dr Gliklich emphasized that acceptance of AI data will be dependent upon how comfortable end users of AI data are with performance characteristics. He noted that positive predictive value in detecting safety events within unstructured data must pass a certain threshold in order to gain acceptance by regulatory bodies. He stated that "as we move into other areas of how AI might be used with unstructured data to generate a usable signal of safety, effectiveness, or efficacy, will depend on generalized performance metrics that are understood, validated, standardized, and surpass known thresholds."

THE DISCUSSION CONTINUES

ISPOR continues its discussion on AI application within the drug and device development process during its 2019 Conference in New Orleans, Louisiana. ISPOR will be hosting "Global Developments in Artificial Intelligence and Machine Learning in Healthcare." This Spotlight Session will focus on trends in AI and ML from the perspectives of North America, Europe, and the Asia Pacific regions. Presenters will address issues surrounding causal inference, as well as the differences between unsupervised and supervised methods within ML. Presenters will also review how AI and ML methods are currently being used in healthcare delivery, drug discovery, health technology assessment, regulatory approval, and safety surveillance. The session will close with presenters forecasting how AI use may evolve over the next decade.

Recently the European Commission acted to increase the availability of healthcare data sharing through the Digital Single Market.[6] European Commissioner for the Digital Single Market and Vice-President Andrus Ansip said, "The Digital Single Market is rapidly taking shape; but without data, we will not make the most of artificial intelligence, high-performance computing, and other technological advances. These technologies can help us to improve healthcare."

As RWD sources continue to expand, so will the need for sound AI methods. ISPOR looks forward to engaging researchers, regulatory bodies, and other stakeholders during the 2019 ISPOR Conference to advance AI applications in clinical research.

REFERENCES

1. Hernandez-Boussard T, Kourdis PD, Seto T, et al. Mining electronic health records to extract patient-centered outcomes following prostate cancer treatment. AMIA Annu Symp Proc. 2018;2017:876-882.

2. Grimes S. Unstructured data and the 80 percent rule. Breakthrough Analysis. August 1, 2008. http://breakthroughanalysis.com/2008/08/01/unstructured-data-and-the-80-percent-rule/. Accessed March 27, 2019.

3. European Medicines Agency. EMA Regulatory Science to 2025. https://www.ema.europa.eu/en/documents/regulatory-procedural-guideline/ema-regulatory-science-2025-strategic-reflection_en.pdf. Accessed March 27, 2019.

4. Centers for Medicare & Medicaid Services. Artificial Intelligence (AI) Health Outcomes Challenge. https://innovation.cms.gov/initiatives/artificial-intelligence-health-outcomes-challenge/. Accessed March 27, 2019.

5. Han L, Ball R, Pamer CA, Altman RB, Proestel S. Development of an automated assessment tool for MedWatch reports in the FDA adverse event reporting system. J Am Med Inform Assoc. 2017;24(5):913-920.

6. European Commission - Press Release. Data in the EU: Commission steps up efforts to increase availability and boost healthcare data sharing. April 25, 2018. http://europa.eu/rapid/press-release_IP-18-3364_en.htm. Accessed March 27, 2019.

ABOUT THE AUTHOR

Michele Cleary is an HEOR researcher and scientific writer with more than 15 years of experience in the healthcare field.