Reviewing Current Trends of Artificial Intelligence Use in Generating Real-World Evidence from Patient Medical Charts

Sneha S. Kelkar, MSc, MPH. Viktor Chirikov, PhD

OPEN Health, New York, US, Contact: snehakelkar@openhealthgroup.com



INTRODUCTION

- Use of real-world data (RWD) to generate meaningful outcomes has rapidly evolved in the last 2 decades.
- Initial efforts were focused on the application of RWD to generate real-world evidence supporting scientific hypothesis, clinical decisions, post-marketing treatment effectiveness and safety, as well as regulatory submissions.
- The current trends have shifted towards the application of Artificial Intelligence (AI) to RWD to enhance its application for evidence generation.
- Patient medical chart review data are by far the most informative source of RWD. However, significant manual efforts and a tedious abstraction process has limited its utilization to generate rapid, concise and affordable real-world evidence in the past.
- Al tools such as Machine Learning (ML), Natural Language Processing (NLP) models and Deep Learning Algorithms are evolving the ability to use patient medical charts data.

OBJECTIVES

To present an overview of the current use of AI with patient medical charts as a source of RWD and to understand the roles AI and human intelligence need to play in the use of such data for HEOR studies.

METHODS

- A targeted literature review was conducted in Medline and Embase for articles published since 2022.
- The search strategy combined strings for "Artificial Intelligence" and its components such as "Machine Learning", "Natural language processing", "Deep Learning", "Computational Intelligence" with category terms used to identify "medical chart reviews". Screening was conducted independently by two researchers.
- As this was a targeted search, the results were supplemented by searching for independent terms such as "synthetic data" or "data curation" with "medical chart" in search engines such as Google, or policy documents, regulatory/HTA guidelines, and stakeholder's websites.
- Data extraction was qualitative and was interpreted from the perspective of the strategic plans of ISPOR 2030¹ and Medical Affairs Professional Society (MAPS) 2030,² with focus on:
- Use of AI, or related techniques, in increasing the efficiency and speed of medical data abstraction
- Use of AI, or related techniques, in the ability to generate more robust and quality data
- Use of AI, or related techniques, in generating synthetic data or accounting for missing information in medical charts

RESULTS

- The initial search yielded 1,217 articles, of which 30% were conference abstracts and 18% were literature reviews. We found 249 articles relevant to the topic of interest.³
- Articles reviewed AI efforts across 27 countries (>80% focused on AI use in United States).
- Articles showed that AI has been applied to retrospective and prospective medical charts, registry studies, but that efforts were very specific to a research/hospital center or a disease area.
- There was significant variation in the research objectives, intended use of AI, the data sources that fed into the AI models, and the performance of these AI models.
- Very few papers directly compared AI vs human efforts with respect to medical record abstraction.
- Several papers provided information on individual AI model construction, model parameters, challenges and performance matrix that can assist future research to build integrated models (Figure 1).

DISCUSSION/CONCLUSION

- Multiple agencies advocate for the use of real-world data to advance patient-centered care. Future efforts should emphasize the application of AI to chart review, guided by medical affairs and directed by human researchers, ensuring a fast, affordable, reliable, transparent, and unbiased real-world evidence generation (Figure 2).
- The findings of our review align with ISPOR's commitment to longterm sustainability of healthcare systems and MAPS goals to transform traditional evidence generation studies by adopting digital philosophy.

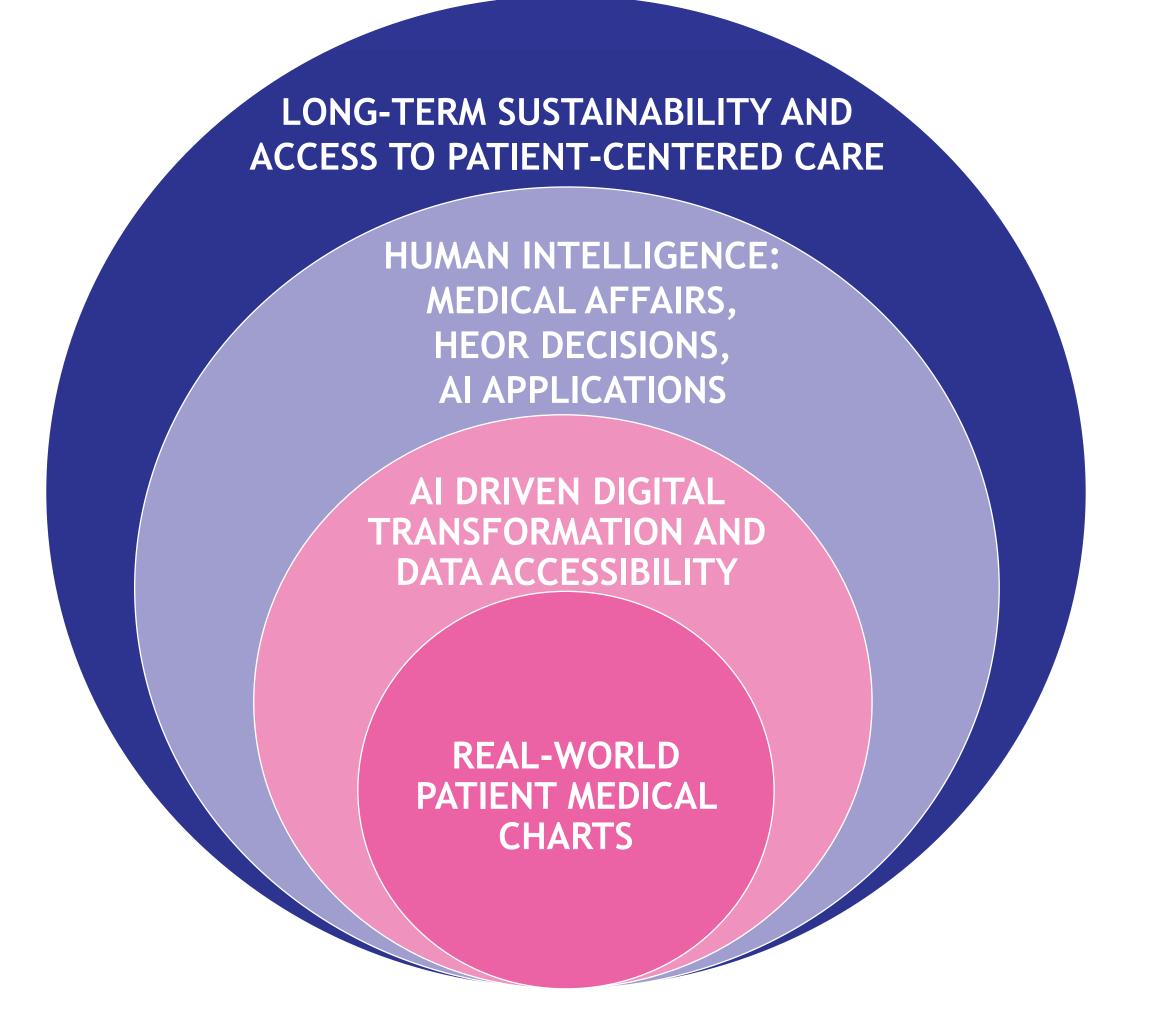


Figure 2. Interpretation of Findings within ISPOR/MAPS Strategy Plans

Figure 1. Summary Of Main Findings And Model For Achieving Synergy Between Al And Human Intelligence For HEOR Studies

Al Assistance in Accelerated Treatment Decisions

Al provides unprecedented speed over human researchers providing quick and concise information from medical charts. However, current models are still isolated, non-generalizable, with high sensitivity but varying specificity.

Al Enabled Data Curation And Extraction

Al use is anticipated to increase for data consolidation, identifying false positives, annotating, coding data, and extracting a pre-specified list of key variables from structured and unstructured data.

Al Models Used In Predictive Outcome Analysis
Increased demand for using Al models to make data-driven predictions
on selected outcomes such as trial eligibility, relapse risk, or need for
interventions based on clinically relevant indicators, which are
identified by humans.

Handling Data Richness and Diverse Data Sources

Digitization of healthcare records is generating substantial volume of individual patient data. All is best fit to process large quantities of longitudinal data from clinical records, pathology, imaging, audio or video reports.

Increasing Record Usability and Synthetic Data Generation

Newer AI models focus on individualized imputations to reduce data missingness. Deep learning models are generating synthetic patient records, alleviating privacy concerns and providing data accessibility.

Commercial Availability of Clinical Data

Several research grants and current efforts are aiming to produce cleaner, usable datasets from patient medical charts. These efforts will likely make clinical data commercially available for research.



- 1. ISPOR—The Professional Society for Health Economics and Outcomes Research. ISPOR Strategic Plan 2030. An ISPOR Strategic Report. July 2024
 - . Medical Affairs Professional Society The Future of Medical Affairs 2030. https://medicalaffairs.org/future-medical-affairs-2030/ [Accessed April 12]
- 3. Targeted literature review data on file.