Is Artificial Intelligence the Next Big Thing in Health Economics and Outcomes Research?

Juan-David Rueda, MD, MS, University of Maryland School of Pharmacy, Baltimore, MD, USA; **Rafael Alfonso Cristancho, MD, MS, PhD**, GlaxoSmithKline, Collegeville, PA, USA; and **Julia F. Slejko, PhD**, University of Maryland School of Pharmacy, Baltimore, MD, USA

To harness the enormous potential of AI in health economics and outcomes research, we need to improve the quality of healthcare information systems and data, train researchers and decision makers on these methods and applications, and define some basic guidelines for any AI-driven research activity.

The human brain has several capabilities that make it unique, including perception, learning, problem solving, decision making, linguistic abstraction and generalization, creativity, pattern recognition, forecasting and more. Intelligence is the ability to understand an issue or problem, and by applying previous knowledge, solve it. Artificial intelligence (AI) is the use of machines to perform processes that mimic this capability.[1] AI integrates multiple cognitive functions to sense, cognize, and perform tasks.

While AI can be classified in multiple ways, the most-used definition divides AI into 2 broad categories: strong AI and weak AI. Strong AI refers to the concept that machines can think and perform tasks on their own, just like a human being, with little to no human interaction. This has been depicted in popular films and television. Weak AI is much more focused and frequently used. Its goal is to solve a specific task, eg, finding the best route on your smartphone or using an application that recommends music or films based on your preferences, like Pandora™ or Netflix™. Nonetheless, other subcategories of AI offer vast potential to explore, such as image recognition, natural language processing, expert systems, speech, planning, and robotics, among many others.[1]

AI research and its applications in data analysis have been adopted rapidly in other fields, particularly in technology and marketing. In healthcare, with the increasing use of information systems, the access to large amounts of data across the healthcare systems and potentially, from other sources that routinely collect health-related data, leverage these applications and optimize many processes and decisions. Specifically, in the field of health economics and outcomes research (HEOR), we rely on healthcare systems data, such as administrative claims or electronic health records, to generate evidence that can help to inform decision for patients, providers, healthcare systems, and policy makers.

We have identified potential opportunities for using AI in HEOR, matching 4 well-established applications of AI:
1) natural language processing;
2) text data analysis; 3) machine learning (ML); and 4) deep learning (Figure 1), to 5 of the most common types of HEOR research activities: 1) burden of illness; 2) drug utilization and patterns of use; 3) patient-reported outcomes (PRO); 4) comparative-effectiveness research (CER); and 5) economic evaluations (Table 1).

NATURAL LANGUAGE PROCESSING

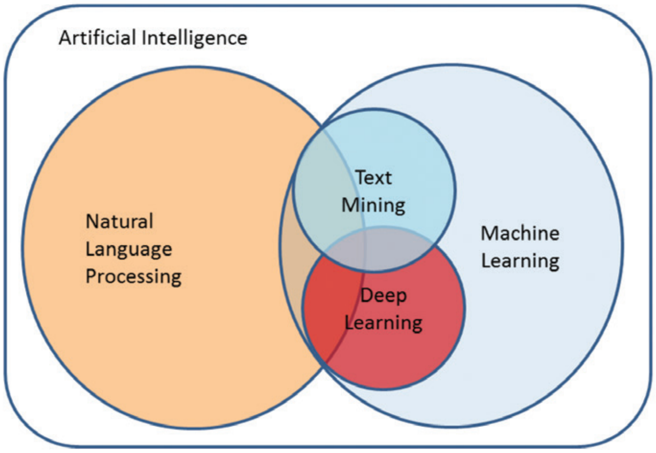
Natural language processing (NLP) is the field that aims to make human language

Table 1. Potential for Use of AI in HEOR by Study Type

	Burden of Illness	Drug Utilization and Patterns of Use	Patient- Reported Outcomes	Comparative Effectiveness Research	Economic Evaluations
Natural Language Processing		+++	+++	+++	
Text Data Analysis	++		+++		
Machine Learning	++	+++		++	+++
Deep Learning	+++			+++	

The rating represents the strength of the application of each method to the HEOR research activities.
+ = less applicability; +++ = high applicability.

Figure 1. Relationship of common applications of AI in research



accessible to machines.[2] This is one of the most prominent and successful fields in AI. We can find examples around us all the time: Siri, Alexa, Google home, etc. The objective of NLP is not only to establish the structure between words in a text (syntax) but also to understand the meaning (semantics) and the context meaning (pragmatics).[3] Algorithms that use NLP have been implemented in chatbots, making them capable of applying deductive coding (supervised ML) and inductive coding (unsupervised ML). In this context, the chatbot is trying to identify common themes from the source file or document, similar to the job performed by a human researcher when coding qualitative data.

In the first approach, supervised ML, a "code book" is used to link each one of the sentences in the interview. The second, unsupervised ML is more exploratory, allowing the chatbot to compile sentences that seem related to a given theme. This application could be used easily in PRO research. There is also an interesting potential application of this technology to the identification of adverse events, or other outcomes not routinely or consistently coded in electronic medical records (EMRs) and frequently used in drug utilization, CER, and PRO studies. Performing systematic scanning of open fields with text in EMRs or physician dictation notes, NLP could capture and analyze additional information to confirm and contrast the findings using only structured data fields or codes.

TEXT DATA ANALYSIS/TEXT MINING

Text data analysis or text mining refers to the conversion of unstructured text data into structured data. The concepts of text data analysis overlap with NLP and data mining, but text mining is limited to written sources. Text mining identifies syntax, semantics, and pragmatics that would otherwise remain hidden in a written document. Rather than a simple keyword search, the machine uses text data analysis to read and analyze documents.[3] EMRs and medical and prescription claims datasets commonly have structured and unstructured data that contain valuable medical information and are used frequently in research. However, discrepancies exist between codes used for billing purposes and the notes from the doctor's office. Combing through medical records' unstructured data is time-consuming and difficult to standardize. Text data analysis can make this process much more efficient to enhance the implementation of CER and to generate real-world evidence. Already some companies are using these technologies to facilitate the development of systematic reviews, for example IBM Watson™ and Doctor Evidence™.

MACHINE—OR STATISTICAL—LEARNING

Machine—or statistical—learning has a great potential for application in HEOR for its ability to learn and perform tasks. ML, named as such because to acquire new knowledge, the machine "learns" from experience and tunes the algorithms over time, requires vast amounts of data.[4] Its goal is to transform data into intelligent action and perform a specific task. Models that use clinical and demographic information for prediction of events, such as severe exacerbations in patients with asthma[5]; or to diagnose a condition using specific patterns applied to image recognition, for instance, diagnosis of a genetic syndrome using face photography[6]; or using voice recognition to detect changes related to dementia[7]; are just a few examples of current applications.

One of the most commonly used ML algorithms is neural networks. A neural network mimics the structure of the cells in the human brain.[8] Neurons are connected through synapses. In a neural network, multiple layers of algorithms (neurons) feed data into other algorithms, creating a very intricate system to perform a specific task. This system comprises 3 layers: the input layer, or original data; the hidden layer, or "black box"; and the output layer, which is the specific task performed.[8] Similar to the brain, the explanation of the interaction of neurons is meaningless; the relevance is focused on the outputs obtained that are tangible and improve over time. If we define clear, specific rules linked to a dynamic dataset with the relevant inputs for updating or for the adaptation of a previously developed economic model, the machine could perform this task using neural networks and update the results in real time for multiple countries or healthcare systems, based on the data available.

DEEP LEARNING

Deep learning refers to the process of understanding large amounts of data with multiple hidden layers in a neural network, increasing the computing power over very large and complex datasets.[9] Burden of illness studies, which are aimed to determine the healthcare resource use, costs, and humanistic impact of a given condition, will require data from multiple sources, including patients, providers, and health care systems. Often, a combination of datasets from epidemiologic surveys or registries, claims datasets, and patient surveys are used in order to achieve this goal. The application of deep-learning techniques could perform these analyses more efficiently.

The use and impact of AI on our daily activities are undeniable. AI helps us connect to each other, decide what to watch or listen and what or when to buy, and often, answers our questions faster than ever before. Nonetheless, in healthcare and research, the adoption of AI is just starting, and many barriers and challenges are emerging. For example, the collection and use of private data is increasing across many different platforms but it still unclear how that data can or will be used in the future by those who already have the information or are collecting it. >

Fresh in our minds is the recent Facebook data breach, exposing more than 50 million users.[10]

Privacy issues can be even more sensitive with medical and other health records, which may be subject to similar security risks. Some ethical concerns have been raised as well, specifically regarding the potential of AI to favor some subgroups simply based on having more or better information, similar to the traditional information bias but at a different level. Along the same lines, access to technology and AI applications and its potential benefits is not the same for everyone, potentially increasing certain disparities. Finally, the quality of the data, as with many other data-driven applications, will determine the quality of the results. In our field, data quality is heterogeneous and can lead to hidden errors that are difficult to identify.

The use and impact of AI on our daily activities are undeniable. AI helps us connect to each other, decide what to watch or listen and what or when to buy, and often, answers our questions faster than ever before. Nonetheless, in healthcare and research, the adoption of AI is just starting, and many barriers and challenges are emerging.

As we described above, many processes inside AI can become too complex or difficult to understand, like a black box, that is difficult to report and in some cases could be proprietary, limiting reproducibility. We advocate for full transparency of methods, data, and algorithms. Currently, there is no guidance in the reporting of models that use AI (specifically ML) in our field. Finally, as many have predicted in the movies, we could encounter a critical issue known in AI as the "control problem." This problem can be summarized as: How can we create machines that help us without harming us? This could be a problem if AI is assigned to maximize goals but finds an undesirable solution, as illustrated by the Greek myth of King Midas or more recently, in the HBO series Westworld.

In order to harness the enormous potential of AI in HEOR, we need to improve the quality of healthcare information systems and data, train researchers and decision makers on these methods and applications, and define some basic guidelines for any AI-driven research activity.

REFERENCES

1. Russell SJ, Norvig P. Artificial Intelligence: A Modern Approach. Malaysia; Pearson Education Limited; 2016.

2. Nilsson NJ. Principles of Artificial Intelligence. Burlington, MA: Morgan Kaufmann; 2014.

3. Zhai C, Massung S. Text Data Management and Analysis: A Practical Introduction to Information Retrieval and Text Mining. Williston, VT: Morgan & Claypool; 2016.

4. Abu-Mostafa YS, Magdon-Ismail M, Lin H. Learning From Data. Vol 4. New York, NY: AML Book; 2012.

5. Finkelstein J. Machine learning approaches to personalize early prediction of asthma exacerbations. Ann N Y Acad Sci. 2017;1387(1):153-165.

6. Kruszka P, Addissie YA, McGinn DE, et al. 22q11. 2 deletion syndrome in diverse populations. Am J Med Genet A. 2017;173(4):879-888.

7. Baldas V, Lampiris C, Capsalis C, Koutsouris D. Early diagnosis of Alzheimer's type dementia using continuous speech recognition. In Lin JC, Nikita KS, eds. Wireless Mobile Communication and Healthcare. MobiHealth. Springer, Berlin, Heidelberg; 2010:105-110.

8. Haykin S. Neural Networks. Vol 2. New York, NY: Prentice Hall; 1994.

9. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436.

10. Cadwalladr C, Graham-Harrison E. Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. Guardian. 2018;17.