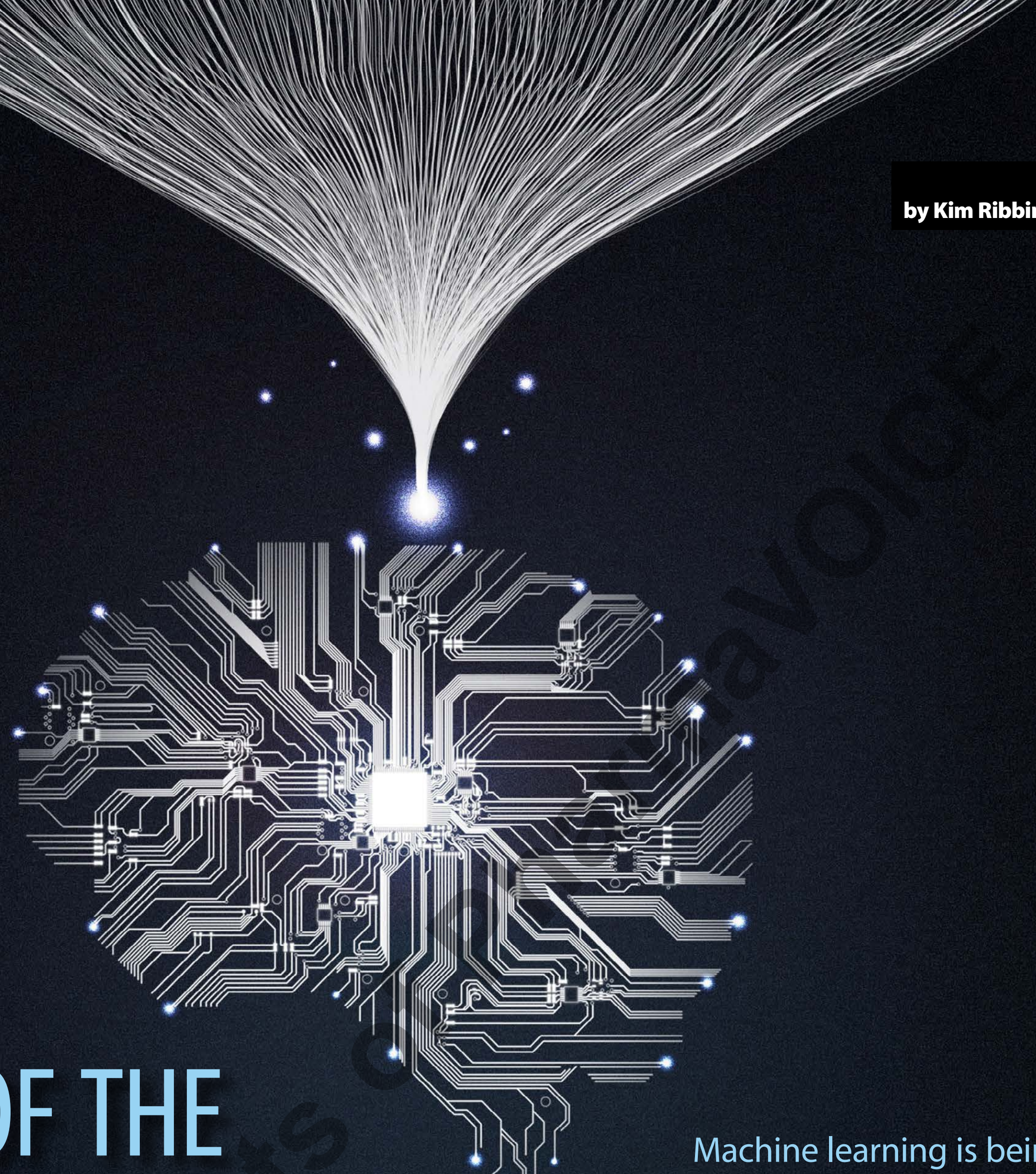


by Kim Ribbink

AGE OF THE MACHINE: Using Algorithms to Advance Outcomes

Machine learning is being applied across the industry to support decision-making in clinical trials, R&D, regulatory, and commercial processes.

What's the difference between machine learning and artificial intelligence? The answer lies in the word intelligence, since AI is about enabling machines to make decisions, whereas machine learning refers to the ability to teach machines to extract information or patterns from data.

In life sciences, machine learning is being more widely applied to discover patterns hidden in the massive amounts of healthcare data being generated today and to support decision-making, says Basheer Hawwash, Ph.D., principal data scientist, Remarque Systems.

In essence, machines are there to carry out laborious, basic work, leaving researchers and other experts to focus on what is relevant, says Marco Anelli, M.D., head of ProductLife Group's Information, Knowledge & Intelligence Group.

Machine learning is being used to increase the efficiency and effectiveness of clinical trial recruitment in terms of patient identification and identifying clinician influencers.

SUSAN ABEDI
81qd

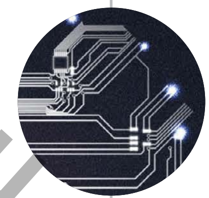
"Take research for COVID-19 for example, machine learning probably won't come up with a big discovery but it will sift through thousands of potential molecules and remove those that aren't relevant so that experts can focus on just those that hold potential," Dr. Anelli says.

According to Vimal Mehta, Ph.D., CEO of BioXcel Therapeutics, machine learning approaches can help unlock potential new uses for drugs in clinical trials or find new therapeutic indications for drugs that previously failed. "Machine learning could also help to speed up processes and provide treatments faster," he says. "Machine learning algorithms could be applied to parse the information about compounds, biochemical pathways, different drug mechanisms, disease pathologies, symptoms, and diseases that can be used to create new relationships by building meta data and then translate that meta data into novel connections, enabling the discovery of therapeutics."

Machine learning can process both labeled and unlabeled data, Dr. Hawwash says, explaining that labeled data is typically when researchers already know the outcome but want to predict how future data will behave. An example is processing MRI images of cancer to build a prediction model to detect the disease.

"Algorithms dealing with unlabeled data — raw data that hasn't been categorized — are incredibly useful for a crisis, such as the COVID-19 pandemic," he says. "We don't know what the outcome will be, but we try to use all this source data to answer questions that nobody has answered before."

Machine Learning and Medical Imaging



Machine learning (ML) and artificial intelligence (AI) technologies are gaining ground in medical imaging. For many health IT leaders, machine learning is a welcome tool to help manage the growing volume of digital images, reduce diagnostic errors, and enhance patient care. Despite its benefits, some radiologists are concerned that this technology will diminish their role, as algorithms start to take a more active part in the image interpretation process while ingesting volumes of data far beyond what any human can do.

ML — and CAD applications in general — show promise, and radiologists have much to gain from incorporating this technology into their operations:

- AI can evaluate an enormous number of imaging variables much faster, and more consistently, than a radiologist.
- Algorithms facilitate decision-making and education for inexperienced radiologists.
- CAD can automate mundane reading and measurement tasks, freeing radiologists to focus on patient interaction, research, and complex higher-order thinking.
- ML can automate radiologist workflow, placing more time-sensitive cases higher on the radiologist's worklist.
- Machines have the potential to improve diagnostic accuracy dramatically, prevent medical errors, and reduce the overuse of testing.
- ML can act as a next-generation clinical decision support tool for radiologists, offering segmentation, classification, and pattern recognition that can be used to propose statistically significant guidance for image analysis.
- Analyzing images can be highly subjective; machines replace subjectivity and reader variability with quantitative measurements that can improve patient outcomes.

There are, nevertheless, many challenges, such as concerns among radiologists that ML will lead to fewer jobs and a diminished role, possible legal challenges if an algorithm leads to errors or misdiagnosis, and the sheer complexities involved in building algorithms that apply to a broad set of scenarios. To mitigate some of these issues, ML implementations should look at several key considerations:

- Engage all stakeholders in the planning process. ML can help radiologists improve patient outcomes, but ML initiatives can fail if healthcare organizations do not address fears.
- Be mindful of the application scope. Many ML algorithms are narrow in their application, working across select modalities to inform decisions on specific diseases.
- Incorporate ML as a complement to the radiology staff. Even when algorithms are accurate, radiologists still need to apply their judgment, using the algorithm as a secondary support system to optimize care.

Source: Extracts from article by Partha S. Anbil & Michael T. Ricci, IBM Healthcare & Life Sciences Practice

Machine Learning and R&D

Dr. Hawwash says machine learning can be, and is being, used in every stage of drug development, starting with the ability to analyze the DNA of a disease much faster, as well as predict the pharmaceutical properties of a broad range of molecular compounds.

"During the planning and start-up phases, existing literature about similar clinical trials can be fed into a text-mining algorithm to help design a study, data can also be used to facilitate site selection and to quickly and accurately identify patients for trials," he says. "Once the study is in progress, machine learning can support remote monitoring that identifies data risk and patient safety concerns."

Loubna Bouarfa, Ph.D., founder and CEO at OKRA Technologies, says machine learning algorithms have already been used in the design of automated drug development pipelines, guiding, and speeding up drug discovery, preclinical, and clinical studies.

For its acute agitation drug, BioXcel is developing an algorithm to be used with the Apple Watch that potentially predicts when a patient will have an agitation episode, enabling the administration of BXCL501 before the onset of agitation. Agitation can result from multiple disorders, like schizophrenia, bipolar, Alzheimer's, dementia, opioid withdrawal symptoms, and delirium. BioXcel is currently conducting Phase III trials of BXCL501 for the treatment of acute agitation in patients with schizophrenia and bipolar disorder.

"The drug was identified using machine learning and AI," Dr. Mehta says. "The whole process is driven by a machine-learning approach. We hope we can validate the model by filing our first NDA in 2021, and upon approval, go to market in early 2022."

Susan Abedi, executive VP, commercial solutions at 81qd, says machine learning is being used to improve clinical trial recruitment in different ways. One way is by identifying healthcare professionals who are highly likely to encounter target patients with rare diseases, enabling companies to more effectively carry out trial recruitment.

"Machine learning can identify influential clinicians who are connected to other clinicians managing relevant patients," Ms. Abedi says. "Biopharmaceutical companies can partner with clinical leaders to serve as investigators for trials and clinical leaders can leverage their networks to drive clinical trial recruitment of eligible patients through referrals."

"Machine learning replaces the mundane tasks of scrolling through literature and clinical trial results," Dr. Bouarfa says. "In terms of the conduct of clinical trials, machine learning can be deployed in every single step of protocol design, such as background information, objectives and evaluation criteria, subject selection, study procedures, or power calculation, among others."

Researchers can also apply machine learning to determine whether a drug is carcinogenic or not, Dr. Anelli says.

"Through machine learning we can get an accurate idea as to whether a drug is dangerous — this is a valid and validated approach," he says. "This is a use of machine learning that is being applied quite extensively. It's important since it allows researchers to filter out many drugs, saving years in time and millions of dollars. This approach also prevents the need to kill vast numbers of lab animals in toxicology studies."

Other important uses of machine learning include drug optimization and drug repurposing, Dr. Anelli adds.

"COVID-19 has brought drug repurposing to the fore, with significant research into different marketed drugs as potential treatments for the virus," he says. "Existing drugs are being examined by means of machine learning algorithms to determine if they can be successfully repurposed to prevent the lung and organ damage caused by the virus."

Machine learning is used to take care of as much of the basic work as possible, allowing experts to concentrate on the most relevant information.

DR. MARCO ANELLI
ProductLife Group

Machine learning also potentially enables studies to enroll fewer patients, which could speed up the enrollment process, be less costly, and reduce the risk to patients, Dr. Anelli says. The process has been used to improve the efficiency of the statistical analysis of clinical trials, and it has huge potential to aid precision medicine, he notes.

"When collecting thousands of variables for thousands of patients, the result is billions of potential correlations," Dr. Anelli says. "If the goal is to correlate each of these parameters with each of the others to determine the effectiveness of a drug on a group of people based on their genes, sex, and so on, then it's something that can effectively be done using a machine learning-based approach."

Dr. Bouarfa says by combining data from demographics, pre-existing conditions, lab tests, imaging, -omics, and wearables, machine learning could identify characteristics of patients most likely to benefit from treatment. Furthermore, the historical data could be used to predict the sample size, recruitment times, and dropout rates for future trials.

"For clinical trials, we need a change of mindset, moving away from the classical clinical trial framework and accepting a more adaptive approach," she says.

One area where machine learning is invaluable is in genomics, says Mark Kiel, M.D., Ph.D., co-founder and chief science officer at Genomenon.

"One significant application involves examining data from genome sequencing results on large cohorts of patients annotated with empirical evidence in published literature, and associated databases relevant to the disease of interest," he says.

Machine learning offers tremendous advantages when it comes to complex decision-making.

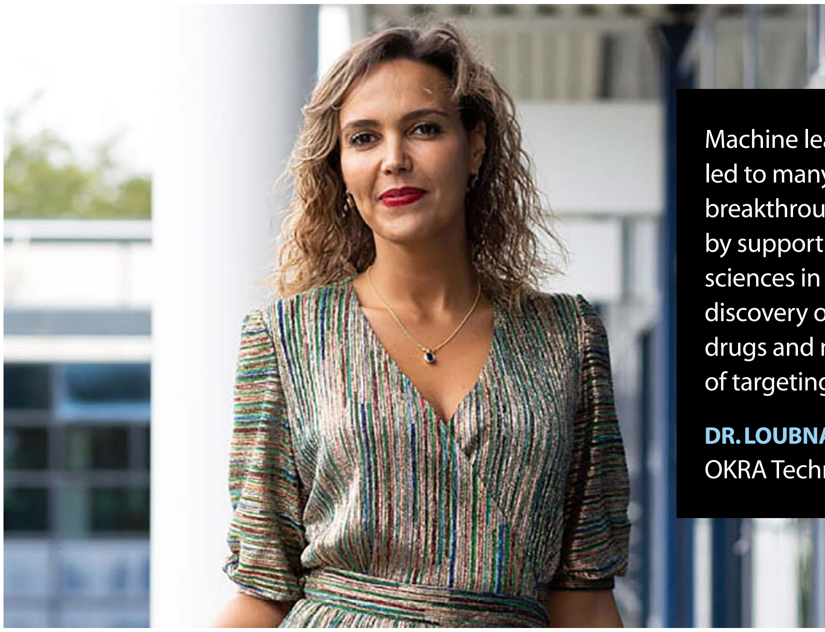
DR. BASHEER HAWWASH
Remarque Systems

"In the case of genomics, machine learning techniques are being used to stratify patients into responders and non-responders based on a variety of clinical parameters in an effort to understand the genetic underpinnings of differential drug responses or clinical phenotypes," Dr. Kiel says. "At Genomenon, we aggregate the information for a given disease, phenotype, gene, pathway, or treatment to assemble de novo a comprehensive genomic landscape of all genetic associations, focusing on causative genetic variants. We then prioritize these results by the strength of this associative evidence."

Dr. Kiel says there have been many breakthroughs with the company's Mastermind Genomic Landscapes to better understand the genetic drivers of disease. These include identifying vastly more pathogenic and likely pathogenic variants for each disease indication than previously known, significantly expanding clinical trial patient population, and gaining novel insight into the mechanism of action for each therapeutic target.

In a first for machine learning, a long-acting potent serotonin 5-HT1A receptor agonist was identified; the compound is currently in Phase I clinical trials for the treatment of patients with obsessive-compulsive disorder.

"By using machine learning, we were able to rapidly sift through algorithms to pinpoint the most effective molecules to engineer the drug," Dr. Hawwash says. "This process accomplished in just a year what typically takes five years. This demonstrates the power of machine learning to help drug development become faster and more efficient — within the



Machine learning has led to many scientific breakthroughs by supporting life sciences in the discovery of new drugs and methods of targeting.

DR. LOUBNA BOUARFA
OKRA Technologies

decade maybe all new drugs will be created by machine learning."

The Machine Potential

According to Ms. Abedi, machine learning can use a broad range of claims data to fundamentally change how the industry approaches both patients and physicians in a more nuanced and impactful way.

"With patients, machine learning can help accelerate the process of getting the right therapies to the right patients," she says. "For example, patient-finding solutions use predictive analytics and machine learning–based algorithms to examine real-world data to identify patients with difficult-to-diagnose diseases, which can drive earlier treatment."

With regard to physicians, Ms. Abedi says machine learning can help identify clinicians who drive behavior change and who can therefore impact the care of patients well beyond their own practices. She notes that collectively, machine learning-based network mapping and natural language processing can optimize patient care by helping to ensure that therapeutic interventions get to the right patients.

The opportunities to apply machine learning also extend into other parts of the life-sciences business, including regulatory processes.

Machine learning-based approaches are being used to help with regulatory compliance, Dr. Anelli says, citing the IBM Watson initiative that is being used to help organizations set up their compliance in all areas — from the dossier submission to legal regulatory requirements to regulations around production and beyond.

"A good example of where machine learning is highly effective is signal detection, because there is a massive amount of data, and pharmacovigilance departments need to extract the signal from the noise," he says.

Managing the Challenges

Dr. Kiel says with machine learning techniques in general, one of the main challenges is overfitting the data — when a model doesn't generalize well from training data to subsequent datasets. Problems with overfitting are amplified when the analysis is meant to inform clinical decision-making or the conduct of a clinical trial.

"When applying black-box machine learning techniques to clinical trial design and analysis, the goal is to ensure there is a solid evidence predicate for any decisions that are made as a result of the analysis," he says. "In effect, the path to the result must be clearly defined and understood at the human level by clinicians and investigators, as well as corroborated statistically. Ideally, the results will also be corroborated by some orthogonal analytic technique as well."

Dr. Mehta says with machine learning, access to and use of relevant data are key, as is building the right model for the machine.

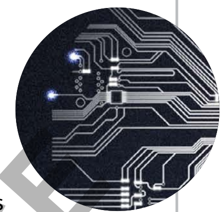
Still regulatory concerns can hold companies back from applying machine learning, Dr. Hawwash says. Reservations are understandable, he says, given that machine learning algorithms aren't static.

"For example, if a clustering algorithm is used to group clinical trial sites, and that grouping is used to determine the frequency of visits, it might later be hard to explain to regulators why a site belonged to a certain cluster," he explains. "However, the advantages machine learning offers are tremendous, and we should be pushing to use it at least as a guideline for decision-making. Fortunately, the industry's resistance shows signs of loosening given the new medical breakthroughs using machine learning."

The fast-moving nature of machine learning means legislators and regulators are challenged to keep up, Dr. Anelli says.

"Dr. Phil Tregunno, who oversees vigilance, intelligence, and research at the MHRA, has spoken about the legislative challenges

Machine Learning: Preventing Theft



A potential for machine learning is training software to recognize known patterns of drug thefts in healthcare facilities, commonly called drug diversion.

"When we began leveraging machine learning, we compiled a large database of known drug diversion cases where individuals had been stealing drugs, including data from various health IT systems," says Tom Knight, CEO of Invistics. "We programmed our software to recognize these events as potential indicators of drug diversion. The more information we provided, the easier it became for our software to comb through similar data from other facilities and detect incidents that exhibit the same patterns that have already been linked with drug diversion."

When developing a solution to aid in the detection of drug diversions, Mr. Knight says one challenge is to join all of the databases from a variety of different health IT systems — not just EMRs or ADCs — but also purchasing records, payroll, and attendance systems, and a variety of other inventory management systems.

"Another issue that is top-of-mind is improving accuracy so the system alerts for as few false positives and false negatives as possible," he says. "Machine learning has significant benefits here as it learns from mistakes, as well as proven incidents of diversion. Every false positive and false negative provides data and improves the accuracy of the machine-learning algorithms."

We use machine learning as a tool to discover patterns for down-stream investigation.

DR. MARK KIEL
Genomenon

Given the current COVID-19 situation, imagine having the ability to feed pandemic information into machines so if we are hit with another pandemic we could quickly process that information and transform R&D.

DR. VIMAL MEHTA
BioXcel Therapeutics

as well as the supervision of machines," Dr. Anelli says. "Dr. Tregunno points out that machines are inherently 'conservative,' because in order for them to 'react' there has to be a strong signal. He warns if we give too much power to the machine we may end up with a bunch of traditionalists in charge instead of innovation-oriented entities."

Ms. Abedi points out that machine learning must be structured to guide action and not just provide information or there is a risk of analysis paralysis.

"Overthinking in the decision-making process can result in no action at all or delayed action," she says. "In the dynamic world of healthcare, agile marketing demands speedy decision-making based on the best data available at the time."

Another pitfall to avoid is becoming too siloed, Ms. Abedi says, adding that biopharma companies must ensure critical work is carried out across medical, commercial, creative, and data science teams to bridge data insights, define the right questions, structure assessment, and build models that translate outputs into actionable insights.

"Integrating machine learning with clinical expertise and strategy is the linchpin of leveraging data to maximize impact," she says. "Successful analytics teams or partners should be cross-functional translators of data insights."

To mitigate some of the challenges, Dr. Bouarfa says pharmaceutical companies should keep several key considerations in mind. First, have the data environment ready since to train machine learning systems, the data needs to be consistent and aligned to industry standard references. Second, machine learning algorithms should be designed with a focus on the problem to be solved. Third, all internal barriers to adoption, such as the integration of third-party apps, need to be removed. And fourth, machine learning should be used to empower people, not replace them.

Implementing Machine Learning — Cautiously

"Machines don't get tired, they can read millions of publications and make connections based on the direction of the data," Dr. Mehta says. "By combining machines and drug developers, unique insights emerge that can increase the efficiency of the overall drug development process."

Dr. Mehta says with machine learning, access to and use of relevant data are key, as is building the right model for the machine.

Interest in machine learning and other forms of AI is increasing, Dr. Mehta says, and larger pharma companies have started embracing these platforms. Machine learning requires a different mindset. While drug developers typically have a hypothesis that they want to test, machine learning looks at all of the experiments and integrates that knowledge before determining what's best from a drug development perspective.

"These approaches are complementary to one another and will help to drive innovation," he says.

Dr. Hawwash says it's important that attention is paid to vetting and verifying the quality of the input data to make certain that the output is reliable. It's also important to recognize that most machine-learning algorithms have an inherent randomness, so their outcomes should be considered as guidelines for better decision-making — not as the final arbiter.

Furthermore, machine learning is simply a tool.

"The key consideration in my mind is the need to not rely exclusively on the results of machine learning approaches to drive research or clinical decisions, but rather to have the results of machine learning lead to appropriate confirmatory and empirical down-stream studies," Dr. Kiel says. PV

SCAN HERE FOR BONUS CONTENT
OR GO TO
bit.ly/PV0620-Machine

New eBook

Telehealth: The Past, Present and Future

Coming in June 2020

Telehealth: The Past, Present, and Future is an electronic publication which provides an overview of:

- Relevant terminology
- Major players
- Categories of telemedicine
- HCP and patient perspectives
- Key factors in the evolution of telemedicine
- Impact of COVID-19
- Predictions for the future

Telehealth is entering a new age.

This eBook will be released on June 10, 2020, in advance of the **Virtual DIA Annual Meeting**.

Contact a sales representative for advertising information.

▶▶ Added Value

Added Value Includes
Distributed to PharmaVOICE list of 60,000 industry executives
Appears within the Featured Contributed Content section on the PharmaVOICE.com homepage for 30 days
Hosted in the eBooks section of PharmaVOICE.com for 12 months
Featured in a Content Box on a PharmaVOICE eNewsletter
Promoted through Social Media including posts on Twitter, LinkedIn, and Facebook

For more details or to sponsor a content page, contact Lisa Banket, Publisher, at 609-730-0196 or lbanket@pharmavoice.com; Cathy Tracy at 203-778-1463 or ctracy@pharmavoice.com; or Suzanne Besse at 561-465-5102 or sbesse@pharmavoice.com.

ENHANCED PATIENT-CENTRICITY: Optimizing Patient Care Through AI/ML/DL

A Sponsor Perspective

DEFINING THE FIELD

ARTIFICIAL INTELLIGENCE (AI) is the science and engineering of making intelligent machines, especially intelligent computer programs.

MACHINE LEARNING (ML) is a system that has the capacity to learn based on training on a specific task by tracking performance measure(s).

DEEP LEARNING (DL), a subset of machine learning, applies algorithms using an artificial neural network comprising layered connections. These connections evaluate and process input data to yield a desired output classification.

The biopharma industry is increasingly realizing the potential of artificial intelligence (AI), machine learning (ML), and deep learning (DL) based technologies to improve the patient outcomes by deriving insights from real-world data generated during medical care. Such approaches are yielding insights into how to more effectively accomplish tasks that experienced professionals can sometimes perform intuitively based on experience, but that are hard to describe in a formal manner. While traditional statistical tests still have an important place in healthcare analytics, by helping infer the relationships between variables, ML models are increasingly valued for their ability to work with very large data sets and predictive accuracy.

Recent studies illustrate compelling applications of AI/ML/DL for diagnosis, treatment, disease management, and patient journey mapping in several non-communicable diseases, which are generally chronic diseases. As the COVID-19 pandemic takes hold in the United States, there are signs that these technologies may help in infectious diseases too.

Patient Journey Mapping via AI

Diagnosis

Challenges in diagnosis can be a major driver of the high cost of care and poor patient outcomes. For conditions that are difficult to diagnose and lacking in pathognomonic signs and symptoms, this can be amplified. For example, fibromyalgia, characterized by widespread musculoskeletal pain and fatigue, where patients can take as long as five years to receive a final diagnosis, presents a case in point and an important opportunity for improvement. Recent studies show the ability for AI/ML/DL to advance fibromyalgia diagnosis through use of imaging, biomarkers and behavioral/emotional indicators.

One of the more promising areas for AI-based computer-aided diagnosis is in the processing of medical images — a vast and rapidly growing source of big data that accounts for at least 90% of all medical data, according to IBM estimates. For example, a recent study by Li et al in Radiology journal (March 19, 2020) showed that a deep learning model can accurately distinguish COVID-19 from community-acquired pneumonia and other lung diseases on chest CT scans.

A wide-ranging analysis by Liu et al in The Lancet Digital Health (October 2019) found the diagnostic performance of DL models to be equivalent to that of healthcare professionals in classifying diseases using medical imaging, yet with increased speed and capacity. The systematic review comparing deep learning performance with healthcare professionals in detecting diseases from medical imaging identified 31,587 studies published since 2012, of which 82 were included and 69 provided sufficient data to construct contingency tables, enabling calculation of test accuracy. Sensitivity was found to range from 9.7% to 100%, (mean 79·1%, SD 0·2) and specificity ranged from 38.9% to 100% (mean 88·3%, SD 0·1).

In fibromyalgia, specifically, ML and medical imaging have been used to distinguish the brain scans of those with this condition from those without. One study, published by Lopez-Sola et al in Pain (2017), identified a brain signature that characterizes fibromyalgia central pathophysiology at the neural systems level and used ML techniques to identify a brain-based fibromyalgia signature. Combined activity in various patterns "classified patients vs. controls with 92% sensitivity and 94% specificity in out-of-sample individuals," according to the authors. The study provides initial characterization of individuals with fibromyalgia based on pathophysiological, symptom-related brain features, and establishes "a framework for assessing therapeutic mechanisms and predicting treatment response at the individual level." These may constitute objective targets for therapeutic interventions.

Biomarkers also present fertile opportunity for the use of AI/ML/DL. A recent study focused on microbiomes found that ML could diagnose fibromyalgia with 87% accuracy, based only on the composition of the microbiome, according to a second article in Pain (2019). Researchers at McGill University in Canada identified 19 bacterial species that were either increased or decreased in individuals with fibromyalgia. The authors conclude, "To the best of our knowledge, this is the first demonstration of gut microbiome alteration in nonvisceral pain. This observation paves the way for further studies, elucidating the pathophysiology of fibromyalgia, developing diagnostic aids and possibly allowing for new treatment modalities to be explored."

Another ML study involving neural networks, by Andres-Rodriguez et al in the International Journal of Medical Sciences (2019), indicated that IL-10 is the best immune biomarker predicting fibromyalgia diagnosis. This study suggested that the severity of widespread pain was best predicted by quality of sleep, perceived stress, anxiety, and three cytokines (IL-6, IL-10, and CXCL-8 {IL-8}), while severity of fibromyalgia is best predicted by stress, anxiety, and IL-10.

Other studies have leveraged behavioral and emotional characteristics as indicators of fibromyalgia. One study, by Orru et al in Frontiers in Medicine (2020), found that ML models achieved an overall accuracy higher than 80% in detecting both patients with fibromyalgia and healthy controls. This study analyzed the accuracy of the Toronto Alexithymia Scale, a measure to assess alexithymia, which is the inability to recognize emotions and their subtleties and textures in fibromyalgia patients.

These examples demonstrate the ways AI/ML/DL can support more accurate and faster diagnosis of conditions that can severely impair patients' quality of life. In the future, AI-based diagnostic approaches could complement physicians' efforts, creating macro efficiencies in the healthcare system and significant quality-of-life benefits for patients.

Treatment

AI/ML/DL is also opening the door to more effective treatment options and better outcomes by predicting which treatment protocols are likely to succeed based on patient characteristics, comorbid conditions, and treatment rationales. Recent studies show that different approaches to cluster and subgroup analysis can support more effective treatment choices in difficult to treat conditions, as illustrated by pain-related conditions, overactive bladder or erectile dysfunction.

One example is the use of ML to predict pain reduction outcomes for patients with painful diabetic peripheral neuropathy. Methods used in one study by Alexander et al in PLoS One (2018) included hierarchical cluster analysis, coarsened exact matching, regressions optimized using shrinking and penalty search algorithms, and microsimulation. According to the authors, "These analyses reinforced the predictive value of utilizing patient subgroups that reflect more complex patterns of fixed patient characteristics and 'on-treatment' variables that change over time." Similarly, in another study, published by Rahman et al in Obstetrics & Gynecology (2020), an ML model was validated in predicting the likelihood of anticholinergic treatment failure in patients with overactive bladder syndrome. Using an external validation data set, the model's sensitivity was 80.4% and specificity was 77.4%. The model is available at https://oabweb.herokuapp.com/app/pre/.

Use of neural networks, which learn tasks largely on their own by analyzing huge quantities of data, can provide additional avenues for predicting treatment success. One analysis focused on erectile dysfunction (ED) and associated risk factors in males aged 40 to-70 years across the U.S., Italy, Brazil, and China. This study (by Goldstein et al, published in the International Journal of Clinical Practice, 2019) identified natural clusters of male characteristics per country, quantified ED dynamics in these profiles, and compared profiles. Clusters were mainly predicted by unhealthy behaviors, ED risk factors, and ED regardless of positive health characteristics/behaviors. The analysis identified meaningful subgroups of men with heightened ED risk factors, which could help healthcare providers target interventions. These examples show the possibilities for making effective treatment decisions and better manage patient treatment.

Disease Understanding and Management

Digital health management has offered long-held hope for extending clinical resources in understanding and managing NCDs by virtually connecting patients and healthcare providers through digital technology, such as mobile phone apps. Data from personal devices can be gathered to support optimization of just-in-time adaptive interventions to support positive health behaviors such as smoking cessation. This use of technology can create a positive feedback loop: patients receiving timely, personalized support, engage more frequently with clinicians, thus generating more data, which in turn allows providers to employ the most appropriate intervention.

For example, a subject who is trying to quit smoking might access coping strategies when experiencing strong cravings. While helpful, such pull interventions rely on a person to be aware and motivated. To address this challenge, a push approach can be helpful, using sensors and algorithms to decide when an intervention is needed, and what intervention might be most useful. In addition to distraction, for example, stress reduction might also be useful for quitters. One study, by Klasnja et al in Health Psychology (2015), used a micro-randomized trial (MRT) to determine the effects of different interventions delivered at different times. For example, it would be possible to measure whether quitters smoke less after delivery of a distraction exercise than otherwise, and how contextual and psychosocial factors (such as location or level of busyness) moderate the observed changes in the efficacy of the intervention. For a multi-component intervention, the components can be randomized concurrently, making MRTs a form of a sequential factorial design. In a study lasting several weeks or months each person might be

The Authors

Technologies using AI, machine learning, and deep learning present a unique and powerful opportunity to enhance patient outcomes by deriving insights from real-world data generated during medical care. There are signs that these technologies may have a role in flattening the curve in the COVID-19 pandemic by helping to protect healthcare workers, prioritize care for vulnerable populations, and predict trends. Potential applications across therapeutic areas include diagnosis (for example, AI-based analysis of medical images), treatment (where novel approaches to cluster and subgroup analysis can support more choices based on patient characteristics, and neural networks can help predict treatment success), disease management (with digital technology to connect patients and healthcare providers, and help manage various conditions), and patient journey mapping. Use of smaller data sets for machine learning models is seeing growth, with potential to help additional enterprises develop AI strategies. But pharma companies often face barriers in terms of accessing the right data in a compliant manner. Specific challenges include accessing the right volume of data, improved data interoperability, and complying with data governance and privacy requirements. There remains a need for agreed regulatory approaches, operating models, and governance to enable further developments and additional research. The authors of this paper explore a selection of recent studies and examine the hurdles that researchers in industry and academia may need to overcome to fully realize the promise of AI/ML/DL for patients.


KELLY ZOU, PH.D.
PStat, VP, Head of MAI, R&D and Medical, Pfizer Upjohn


JOSEPH IMPERATO, ME, MS
Dir., BT, Pfizer Upjohn


JUAN OVALLE, M.D., MBA
CMO, Global Medical and Clinical, R&D and Medical, Pfizer Upjohn


JIM Z. LI, M.D., PH.D.
Sr. Dir., HEOR Lead, MAI, R&D and Medical, Pfizer Upjohn


NIKUJ SETHI, MS
Sr. Dir., BT, Pfizer Upjohn


AMRIT RAY, M.D., MBA
Global President, R&D and Medical, Pfizer Upjohn

randomized hundreds or thousands of times. The authors conclude that micro-randomized studies have the potential to optimize just-in-time adaptive interventions, "enabling the creation of interventions that can effectively support individuals whenever and wherever they most need the support."

In the infectious diseases arena, AI may also have a role in curbing the COVID-19 pandemic, by potentially helping to protect healthcare workers, as noted in an article in The Lancet Digital Health (April 1,2020); prioritize care for vulnerable populations, as mentioned in a report by Health IT Analytics (March 24, 2020); and predict trends.

Capability Requirements

AI/ML/DL approaches are rich with promises to help derive insights from real-world data such as administrative databases and registries. The healthcare industry is increasingly pursuing efforts to leverage AI/ML/DL to enhance disease understanding and the effectiveness of their therapies. However, pharma often faces significant barriers in terms of accessing the right data in a compliant manner. Specific challenges include accessing the right volume of data, improved data interoperability, and complying with data governance and privacy requirements.

Adequate Big Data

Early machine learning tools required huge data sets to generate useful results, limiting the kinds of models that could be devised. Use of smaller data sets for machine learning models is seeing growth. While massive data sets allow for easy training, developers and researchers are using new techniques to mine and transfer data that allows for training on limited labeled information. Models for machine learning can be trained with small data sets using "few-shot" and "n-shot" approaches, as described by George Lawton in a Tech Target article (November 13, 2019) — both of which require expertise in DL architecture and mathematical formulations. Few-shot learning has potential to help clean and label data sets for ML models, and to grow more data. This ability to learn with limited labeled data could allow companies to make innovative use of large pools of otherwise unusable data. Few-shot approaches could reduce the need to amass large sets of the right data and to invest in the compute to train a model on those datasets — efforts that are challenging and costly.

"Zero-shot" techniques are also being explored, with the ability to learn from related data or from descriptions of what to look for in the data – without the need for any designated data sets. These leverage DL networks that have already been trained by supervised learning in other ways, without the need for additional supervised learning. These types of training models, requiring limited data, have the potential to help additional enterprises develop AI strategies.

Common Data Model

Even with smaller data sets becoming more useful, the range of sources of data besides randomized controlled trials (RCTs), such as RWD as in the imaging and biomarker examples discussed earlier in this paper, still require a standard format across multiple data sets. Many networks are emerging using common data models (CDM) to solve for this need.

One such network is the Observational Health Data Sciences and Informatics (OHDSI) program, a public initiative representing a network of research centers for standard analytical methods and tools. It leverages the Observational Medical Outcomes Partnership (OMOP) common data model, enabling systematic analysis of its disparate observational databases, many of which have been created for different purposes and in different formats. The OHDSI network encompasses about 2 billion de-identified patient records. Benefits of this initiative include the ability to: conduct faster and more reliable studies across datasets and data types; lower the costs of ownership, including understanding coding schemes, writing statistical programs across databases or developing software; and broaden data access via the OHDSI network and remote multi-center database studies.

Data Privacy

Concerns remain over the privacy of protected health information (PHI). This can be achieved by de-identifying personal information using a risk-based approach to develop a high quality dataset for secondary use, such as analytics. These approaches use machine learning to determine the likelihood of patient re-identification, thus preserving as many critical data elements as possible to support rich insight, while still ensuring compliance. Examples are widespread, including conditions ranging from diabetes to gout to HIV.

These capabilities are not easy to build. As a result, many biopharma companies are partnering with external experts to accelerate their AI/ML/DL efforts. Some of the most prominent recent examples include:

- Novartis has partnered with MIT, IBM Watson, Quantumblack, and Intel to advances its efforts in AI, including in clinical trials, drug discovery and patient analytics.
- Takeda also has partnered with MIT to develop healthcare AI, with the goal of driving development and application of AI. The program aims to fund six to 10 ML research projects each year, with a focus on diagnosis, prediction of therapeutic response, biomarkers, process improvements, drug discovery, and trial optimization.
- Roche has partnered with Owkin, a French data science and AI firm, to accelerate drug discovery, drug development, and clinical trials, and with GNS Healthcare, a big data analytics company. Roche also bought Flatiron with the aim of using AI to advance oncology research and optimize patient care.
- GSK has partnerships with multiple AI firms, including Exscientia, BERG, Cloud Pharmaceuticals, and Insilico Medicine.
- Ten big pharma companies formed an AI partnership in Europe to share data to aid in drug discovery. The Machine Learning Ledger Orchestration for Drug Discovery (Melloddy) project uses AI, including a blockchain-based system, developed by Owkin.

Conclusions

When it comes to AI/ML/DL, some biopharma companies may still be asking, "Why build this set of data-driven capabilities?" The answer is that this is no longer optional. In fact Deloitte's Second Annual Real-World Evidence (RWE) Benchmarking survey of biopharma companies demonstrated significant commitment to leveraging ML with RWE. The survey indicated that 60% of companies currently use ML to analyze RWD with 95% saying they plan to implement a capability in the near future. The need to address the current COVID-19 pandemic — and work towards the goal of "flattening the curve" in the spread of infection — adds further urgency to these efforts. However, there remains a need for agreed regulatory approaches, operating models and governance to enable further developments and additional research.

Whether biopharma partners build, acquire, or outsource these activities, it is still a significant effort requiring rigor and careful planning. Recent additions to the literature show that the effort can make a meaningful difference to stakeholders across the healthcare system and that the intersection between AI/ML/DL and digital health management presents a unique and powerful opportunity to enhance patient care. PV

Note: The authors are employees of the Upjohn Division of Pfizer Inc. The views expressed are their own and do not necessarily represent those of their employers. The authors thank Jill Dawson, Ph.D., and Elizabeth Powers, VP and Category Lead, Safety Evidence for Regulators, Real World Solutions, IQVIA, for writing and content development support.