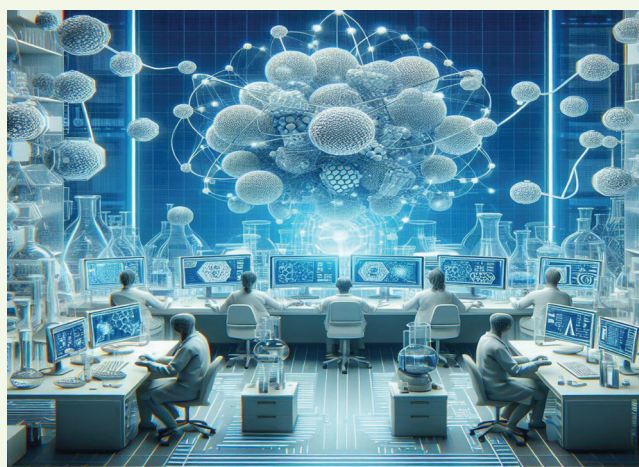


AI REVOLUTIONISING DRUG DISCOVERY

A SCIENTIFIC PERSPECTIVE

Himanshu Jain



DRUG discovery is a multi-stage process filled with financial as well as logistical challenges. According to some estimates, the cost to bring a single drug to market is around 2-3 billion dollars. This cost is mainly inflated by the high rejection rate of candidate molecules during clinical trials and the inefficient discovery phase. However, we can make this process much faster and cheaper by identifying strong candidates for trials and by developing a less labour-intensive discovery pipeline. The solutions are very clear, but according to the American Society for Biochemistry and Molecular Biology, 90% of the drug candidates fail in clinical trials even today.

Attempts have been made to solve these challenges, but nothing much has changed, and the industry still mostly relies on traditional methodologies. This is where Artificial Intelligence (AI) comes into play and promises to drive significant advancements in solving financial and logistical challenges. With the launch of powerful AI tools in the last few years, the world has realised the vast capabilities of AI. Since then, AI tools have revolutionised various fields, from ad-making to financial management. AI tools have the capability to analyse vast datasets, which can help make predictions and bring improvements in the drug discovery field as well. This includes identifying new drug targets, predicting the effectiveness of potential pharmaceuticals, and optimising dosage strategies. Moreover, AI can advance drug repurposing and help tailor pharmaceutical agent design. These applications underscore AI's pivotal role in enhancing the efficiency and effectiveness of drug development processes.

Diverse Domains of AI's Impact on the Industry

The pharmaceutical industry is fundamentally dependent upon efficacious experimentation and robust data analysis.

Consequently, it is characterised by an extensive data repository that can be used to devise innovative solutions. AI's capability to use novel datasets for establishing correlations empowers researchers to integrate them across various segments of the drug discovery pipeline. Notable areas where AI manifests its efficacy include the following:

- **Target Identification**

Drug discovery revolves around identifying and validating potential targets, which could be biological entities or processes targeted by drugs to treat diseases and are modifiable through pharmacological intervention. Historically, this process has been hindered by lengthy and unpredictable timelines due to its reliance on empirical observations and chance discoveries.

The extensive biological data available, coupled with AI's adeptness in processing it, support the utilisation of multidimensional datasets for pinpointing potential therapeutic targets with exceptional precision. By employing predictive techniques, we can evaluate and validate these targets in the early phases of drug discovery, thereby reducing costs and timelines associated with the process. This approach leverages AI's capability to comprehensively analyse complex biological information, leading to cost-effective but more successful drug discovery endeavours.

- **Drug Design**

Another crucial aspect of drug discovery involves identifying potential drug candidates, typically through labour-intensive screening of compound libraries. Minimising side effects and potential toxicological risks is important during this process. AI presents a solution by automating this step, effectively exploring the vast chemical space, and precisely predicting the properties of new molecules. This automation accelerates screening while improving candidate selection accuracy. Researchers can further utilise predictive models to identify drug candidates for target proteins, facilitating the design of novel therapeutics. This takes care of one of the biggest problems in drug discovery, the high rejection rate of drug candidates during the trials.

- **Drug Repurposing**

When the COVID-19 pandemic emerged, numerous companies rushed to develop treatments for the virus. During this period, many existing drugs were also prescribed for alleviating COVID-19 symptoms or

treating the infection. That entire process is essentially called drug repurposing. AI can speed up this process through its ability to predict hidden relationships between existing drugs and diseases, thereby presenting innovative therapeutic solutions for unresolved medical conditions. This would save a huge amount of money and time by reducing the need for new drugs and solving several diseases.

- **Combination Therapies**

Combination therapy or polytherapy is any therapy that uses more than one medication or modality. It is an area of exploration as it is believed that it can treat complicated diseases like cancer. The basic reasoning behind this approach is that we expect a combination of drugs to positively affect each other's potency and, hence, have a more significant effect on the target protein.

Also known as multimodality therapy, this approach can be quite complex. However, AI, equipped with extensive biomedical data, possesses the capability to identify synergistic combinations of drugs. They can predict the binding capabilities of different drug candidates and their overall impact on a target protein. Researchers may uncover new therapeutic interventions tailored to address emerging medical challenges by leveraging AI to develop combination therapies.



- **Predictive Modelling and Simulation**

This stage employs predictive modelling and simulations to provide researchers with a better understanding of complex chemical compounds. While observations and mathematical models have traditionally been relied upon, capturing the complexity of biological systems through these techniques is challenging and unreliable. Incorporating diverse data like ethnicity, age, and gender is difficult for these models, so they only simulate the interactions with some limited accuracy.

AI comes into play by simulating drug-protein interactions better and accurately predicting the behaviour of a given molecule. Machine learning-based models also

enable researchers to integrate diverse data sources and simulate interactions between drugs, biological systems, and disease pathways. This empowers AI solutions to predict drug responses in diverse patient populations and assist with dosing adjustments, automating a significant drug discovery pipeline component without compromising efficiency.

Collaborations between Pharma and AI Professionals

While AI plays an important role in this entire field of drug discovery, collaboration with pharmaceutical scientists truly unlocks remarkable benefits. The crucial role of AI researchers lies in empowering pharmaceutical scientists with data intelligence. By harnessing past trial data and outcomes, AI acts as a catalyst and enables pharmaceutical researchers to arrive at more promising conclusions regarding effective and cost-efficient drugs. This catalytic effect accelerates their ability to reach their set targets and work goals by efficiently sifting through historical data to eliminate unfruitful compounds. It also provides valuable guidance for more focused exploration. Consequently, this streamlines tasks, reduces time and effort per task, and enhances the overall efficiency and productivity of pharmaceutical scientists, enabling them to tackle a broader range of challenges, thereby significantly benefiting their organisations.

While AI holds immense promise in drug candidate identification, like human-driven processes, there is no guarantee that AI-discovered compounds will behave as envisioned and pass clinical trials. Limitations arise from the data used and the intelligence created, which govern the AI's understanding. These limitations include the reliance on past data, the inability to innovate in drug design, and the potential oversight of certain possibilities or inaccuracies in interpretation. Moreover, the training speed on large datasets presents a significant bottleneck, making repeat training and upgrades costly and cumbersome. In some cases, AI may even stifle the creativity of scientists. This underscores the importance of leveraging the domain intelligence of pharmaceutical experts to train and utilise AI-derived insights more effectively.

Pharmaceutical researchers must collaborate closely with AI experts to achieve optimal operational efficiency. These researchers interpret results, shape selection criteria, and validate AI outputs. Once trained on the required data, AI becomes an invaluable ally, swiftly verifying numerous permutations and combinations for drug selection and shortlisting.

When scientists and AI researchers deeply understand each other's fields and work seamlessly together, they can exponentially accelerate the drug discovery process. Compound identification with a higher likelihood of success can be achieved by these scientists in as little as one year, compared to the conventional 4-6 years. Hence, this seamless collaboration provides tremendous benefits in terms of time to reach outcomes and develop effective drugs, streamlining

the research process by rapidly predicting and analysing vast amounts of data to guide selection processes.

The collaboration between AI scientists and pharmaceutical experts isn't just a future prospect; it's already underway. Since 2016, trials for drugs discovered using AI-driven processes have been conducted worldwide, dramatically shortening the discovery phase. Across various companies, scientists have utilised AI predictive modelling and other advanced techniques in designing drugs for conditions like Ebola, Amyotrophic Lateral Sclerosis (ALS), and idiopathic pulmonary fibrosis, some of which are already at Stage 2 trials. The collaborative nature of this field is exemplified by the composition of AI startups in the realm of drug discovery, which often comprise a blend of industry AI experts and seasoned pharmaceutical researchers.

Despite these collaborations and remarkable progress, we still have a long way to go. Even after almost a decade since the work on AI in drug discovery started, we still don't have a single publicly available drug made entirely by AI. The widespread implementation of AI in drug discovery has not yet become commonplace.

Challenges Faced by AI in Drug Discovery

The efficacy of AI hinges on the accessibility of extensive multidimensional datasets, which presents a notable challenge for the technology. Issues such as data availability and quality persist across various facets of drug discovery. Moreover, the pharmaceutical industry operates within a highly regulated and scrutinised environment, posing an additional challenge in providing transparency regarding AI-driven solutions. Every decision within this domain relies on robust scientific evidence, and relying solely on predictions from deep learning models may not suffice to convince scientific bodies.

Another challenge lies in the resources required to support AI tools. AI models are complex and computationally heavy, making the software slow and very expensive to run. According to a report in the Yale Environment Journal, even basic interactions with the present AI tools consume around 0.5L of water, and companies have been consuming millions of litres to train their AI models. The International Energy Agency (IEA) also projects that by 2026, data centres will consume almost 1000 terawatts of electricity, comparable to Japan's total consumption. The environmental burden imposed by AI has led to calls for the development of sustainable AI and several protests in different countries to protect natural resources as well. The cost of these resources, combined with the added need to protect other depleting resources presents another challenge in developing complex AI tools and their expansion.

Furthermore, the field itself is new and upcoming, which makes developing AI models that account for the different variables in the pharma industry quite challenging. Working with millions of parameters and fine-tuning them to achieve high levels of accuracy is a very complex task for even the best of AI engineers. Overall, the lack of knowledge, resources, and data, combined with the regulatory backlash, keeps impeding

the implementation of AI in the drug discovery field. The historical data on technology adoption in the pharmaceutical sector further underscores this observation.

Traditionally, the pharmaceutical industry has tended to trail behind other sectors in embracing new technologies, primarily due to several key factors. Among these factors are the complexities and stringent regulations inherent in the industry, which pose significant barriers to swift technological integration. Additionally, the substantial financial investments required to implement advanced technologies serve as a deterrent for many pharmaceutical companies. Despite the potential benefits these technologies may offer in optimising business processes, the perceived return on investment often falls short of justifying the considerable expenditure of millions, if not billions. Given the lucrative pharmaceutical sector, many companies may be disinclined to allocate substantial resources towards adopting emerging technologies that may not yield immediate, tangible advantages.

Nevertheless, the promise of a revolutionising change still exists, and not just in drug discovery but in allied fields such as medicinal chemistry and personalised medicine as well through prediction for absorption, distribution, metabolism, and excretion properties of various materials, virtual screening of prospects to prioritise them for experimental testing, multiple parameter optimisation for balancing multiple design properties such as potency and absorption, and various design methods. In structure-based design, we can predict the binding affinity and interaction between small molecules with target proteins and simulate the drug-protein interactions. In ligand-based design, we can use generative models to map correlation (potential causation) chemical and biological activity to structure and vice-versa.

In De-Novo Design, we can use the models to generate new chemicals based on the required properties. According to a study by the National Cancer Institute, AI can also be used to develop personalised treatments for complex diseases like cancer. It can help precisely match cancer drugs to patients by predicting whether their cells would respond to the given drug. This shows that the potential is limitless, and the initial steps have already been taken to redefine the process to achieve AI-driven efficiency.

Prioritising effective data organisation is paramount to accelerate the broader acceptance of predictive, personalised, and preventive medicine. This lays the groundwork for domain research scientists to harness this data towards desired outcomes efficiently. Central to this advancement is fostering seamless collaboration across diverse domains, made possible by integrating AI into existing processes and addressing bottlenecks. Such concerted efforts are essential for ushering in a transformative era of medicine. Over time, I envision universal adoption and a profound revolution in the drug discovery field, significantly improving efficiency and accelerating progress for the better.

Mr Himanshu Jain is currently a software engineer at Meta Inc. based out of San Francisco, California. Email: himanshukj17122000@gmail.com