AI in Drug Discovery and Development

Prof. Rajnish Kumar
Dept. of Pharmaceutical Engineering and Technology
IIT-(BHU), Varanasi

Week-11
Lecture-52

Welcome to the course "AI in Drug Discovery and Development." In today's session, we will talk about successful case studies of AI in drug discovery. So, as we have seen, people have been using AI for discovering new targets, new lead molecules, and even new drugs. So, in this session, we will talk about some of the successful case studies where AI has been used for discovering either new targets or even new drugs as well. So, by the end of this lecture, you will be able to understand the impact of AI on accelerating drug discovery and development. Analyze key case studies such as the discovery of halicin, baricitinib, and in silico drugs that have been discovered using de novo drug design.

And then recognize the opportunities and limitations of AI-driven drug discovery in real-world settings and reflect on lessons learned and future directions for AI applications in pharma R&D. So, this is a recent review where these authors have beautifully summarized some of these successful case studies of, you know, the AI-enabled drug discovery. So, where you can see, for example, this is the number of AI-discovered molecules in the clinical trial, and it starts with 2015. By the year 2023, you see at least 67 assets that have been evaluated in clinical trials or which are still undergoing a clinical trial and where AI is playing a major role in their discovery.

So, they have even classified it, you know, which assets are in phase 1, phase 2, and phase 3, and even one of them is being launched in the clinic as well. So, the AI discovered drugs and vaccines that are rising rapidly with exponential growth in small molecule biological candidates. So, AI is showing potential to accelerate the overall drug discovery timeline, and leading pharmaceutical companies are increasingly adopting AI; often we are partnering with AI-native biotech firms. However, a key concern remains regarding the clinical success, safety, and efficacy of AI-discovered drugs, and preliminary industry-wide analysis has begun on clinical success rates, but further validation with more data is needed.

And on the right-hand side, you can see that these are, you know, in by the year 2023. So, if we talk about the number of AI-discovered molecules in the clinical trial again. So, out of those 67 molecules, they found that those are in clinical trials in 2023. So, many of them are the AI-discovered targets, where the target was discovered using AI, and the light green color shows where the target was discovered with AI. The green color shows where the

small molecules were designed using AI, and then in the dark, you know that a green color indicates the AI-discovered antibodies.

And then, in the orange color, you have the AI-discovered vaccines. And then in the red color, you have AI-repurposed molecules, and there are some other clinical assets that have been discovered using AI as well. So, this is a summary table from that publication only, where we kept only those that are in phase two clinical trials. So, you can see here that there is a molecule named BMF 219. So, which was discovered by the A2A pharmaceutical or Biomea company, which is a small molecule where AI was used for small molecule design and optimization, is currently in phase 2 clinical trial during the year 2023.

Likewise, you have a lot of molecules; many of them are, you know, small molecules, and many of them are like vaccines. Many of them are, you know, biologicals as well, like the antibody, and we're almost at each step, starting from the target discovery to the small molecule design and optimization. So, AI has been used and it has been successful until now. Okay, so now we will talk about some of those individual, you know, success stories, like the COVID Moonshot project. So, COVID Moonshot is an open science collaborative initiative launched in 2020 to develop affordable, patent-free antiviral drugs against SARS-CoV-2 by targeting the main protease, Mpro.

In 2019, the whole world was affected by this deadly disease, COVID-19, and there were efforts to discover new drugs; the COVID Moonshot project is one of them. So, where AI was used for, you know, compound prioritization, ML models rank thousands of crowdsourced compounds based on predictive ability, drug likeness, and synthetic feasibility. And then AI was used in structure-based design as well, where the integrated AI modalities were combined with crystallographic data of MPro for efficient hit-to-lead optimization. It was used for generative modeling, you know, as well as fragment merging. So, A was employed to design new scaffolds through fragment merging and de novo compound generation to make the workflows automated by combining AI predictions with high-throughput screening and real-time structural feedback.

The successful outcome resulted in the DNDI6510, which is a promising oral antiviral highlighting AI's role in accelerating real-world drug discovery. So, it was a multi-institute collaboration, and it was also crowd-sourced, where researchers from around the globe were contributing to the design and synthesis of these molecules, especially in the design. So, they designed at least 18,000 molecules, and then those were synthesized with the help of several companies, one of which was Inamine. And then many of those synthesized structures were co-crystallized with the COVID-19 M Pro, and more than 10,000 measurements were taken, resulting in several potent lead molecules, one of which is in advanced phase. Okay, coming to another example where baricitinib was repurposed for

the treatment of COVID-19.

So, where the several AI methodologies were used, like the knowledge graph, which analyzed over 500 million biomedical relationships to identify drugs with antiviral and anti-inflammatory properties. And then, GNN was also used, which predicted baricitinib's dual action: inhibition of viral entry via AAK1 and suppression of the cytokine storm via the JAK-STAT pathway. And then NLP was also used. We scanned over 24,000 publications and clinical trial data for hypothesis validation and candidate prioritization. And some of those tools, AI tools that were used for AWS cloud, processed over 18,000 drug candidates in just 90 minutes.

You can see the computational power. And then a proprietary AI platform, which integrated cheminformatics, transcriptomics, and clinical outcomes to enhance candidate discovery. And then open datasets utilizing COVID-19 patient records, viral genomics, and cytokine profiles to inform the model's prediction. Another example is an open-source anti-malarial discovery named AI for Malaria. So, it is a collaborative AI effort to predict anti-malarial activity against Plasmodium falciparum using open-source data sets and community-built models.

So in this MAIP platform, it is trained on 6.5 million bioactivity points from 11 data sets, which achieved a 12 times hit rate enrichment over random screening. And then DeepChem is also being used for molecular featurization, like converting them into these fingerprints to predict toxicity and activity. And then AutoML from H2O.ai is being used for automated model selection and SAR optimization for compounds with IC50 less than 50 nanomolar and clean ADMET profiles.

Ok then, another AI-discovered drug, or at least the drug candidate that was very popular, is Helicine. So, it was the AI-discovered antibiotic that was discovered at MIT in the Broad Institute. So, it is an AI-discovered broad-spectrum antibiotic that was identified via virtual screening of repurposing libraries effective against multidrug-resistant pathogens. So, in this case, deep learning was used, where the model was trained on over 2,500 antibacterial compounds to predict growth inhibition across the species. and then message passing net neural network which analyze molecular structures to identify candidates with novel mechanisms.

And finally, virtual screening was done, which evaluated over a hundred million compounds in days, identifying helicin. Which was originally a diabetic drug candidate, actually emerged as a top hit, and then it was found that helicine was very potent in inhibiting the pathogen Acinetobacter baumannii. And then other similar compounds were also identified, which had a broad spectrum of antibiotic activity. Okay, coming to another

example where the DSP 1181 was discovered. So, the DSP 1181 is a 5-HT1A full agonist developed for obsessive-compulsive disorder.

So, the whole discovery was done in less than 12 months, which is a major acceleration compared to the typical time taken for the discovery of 4 to 6 years. So, it was developed by Sumitomo Dainappon Pharma and Exscientia Limited leveraging Exscientia's AI platform. And then the contribution of AI was the de novo molecular generation with synthetic feasibility and predictive modeling of pharmacological activity, ADMET, and brain penetration. An iterative design feedback cycle every two weeks where they synthesize only 350 molecules in less than one year. Out of those 350 molecules you know, they were able to identify this DSP 1181 as a potent compound.

So, the DSP1181 showed potent activity, an early onset of therapeutic effect, and entered phase 1 trials in Japan in 2020. So, the broader impact is that it demonstrated AI's capability to streamline drug discovery in psychiatry by combining in silico design, rapid synthesis, and translational biomarkers. Okay, coming to another example, which is Dexmedetomidine, named BXCL501. So, it was developed by Bioexcel Therapeutics using an AI-driven drug re-innovation platform. So, this is actually a repurposing again.

So, where the sublingual film formulation of BXCL501, which is a selective alpha-2 adrenergic agonist, was developed. It was indicated for acute agitation in schizophrenia, bipolar disorders, and Alzheimer's disease and is under investigation for MDD, opioid withdrawal, and PTSD as well. So the role that AI played in this case was that it mined vast biomedical datasets to identify new indications for approved drugs, and it enabled faster hypothesis generation, formulation design, and patient stratification. It also helped accelerate development by focusing on compounds with established safety profiles. So, the clinical regulatory milestones included the FDA breakthrough therapy designation for Alzheimer's-related agitation given to this drug, and then it had a positive phase three tranquility result in Alzheimer's disease in 2023.

FDA approved it as Egalmi for agitation in schizophrenia and bipolar disorder, so this is, you know, one of the drugs that is in, you know, clinics now for the use of patients in the treatment of acute agitation in schizophrenia, bipolar disorder, and Alzheimer’s disease. Okay then, another very popular drug, which is still in phase 2 trials, is the discovery of rentosertib. So, it is not only the discovery of a small molecule, but the target itself was discovered by using, you know, AI in silico medicine. So, this is from, you know, in silico medicine. So, it is a first-in-class anti-fibrotic small molecule for idiopathic pulmonary fibrosis (IPF) and the first AI-discovered drug target, as well as the molecule to enter clinical trials.

So, where you know a lot of data in the form of GWAS clinical data, it was used to discover a target, and once the target was discovered, which is TNIK. Once the target was discovered, using generative modeling and de novo drug design, several molecules were designed for that target, then optimized, synthesized, and tested. It came to the discovery of rentosertib, which is currently in phase 2a clinical trial and has successfully completed that. And that is not only a drug that is coming from the in silico medicine, but you have a number of, you know, drugs that are in different stages of, you know, clinical development. So you can see on the website of the in silico medicine which has been developed using you know.

All those sorts of AI tools, like Panda Omics or the generative chemistry tools like Chemistry 42, or the clinical trial design tool like Medicine 42. So all these assets have been discovered with the help of AI-based tools in silico medicine, and this is one of the companies that is doing very well at it, actually. Ok, and then benevolent AI discovered novel kinase inhibitors. So, benevolent AI uses its AI platform to discover novel kinase inhibitors targeting undrugged kinases in complex diseases such as cancer and autoimmune disorders. So AI was used for, you know, knowledge graphs which integrate over 500 million biomedical relationships to prioritize kinase targets based on disease mechanisms, genetic links, and safety profiles.

And the graph neural network, which predicted kinase drug interactions by analyzing graph topologies, identified baricitinib as a JAK-1/2 inhibitor for Covid-19. And then the tools that are being used were like proprietary KG, which combines data from PubMed, ChemBL, and proprietary sources. and AWS cloud like scaled computations for screening 18,000 plus track candidates in 90 minutes and GAN and VAE which generated novel kinase inhibitors, scaffolds and optimize binding and admet properties. And then another example is ABSCI and Owkin. So, they have partnered to co-develop antibody-based therapeutics targeting novel pathways in immuno-oncology, immunology, and inflammation.

So, they have this Owkin predictive AI that analyzes biomedical and organoid data to identify and validate disease-relevant targets with strong translational potential. And then they have Absci's generative AI, which designs de novo antibodies optimized for binding safety and manufacturability without needing prior binding data. And then the tools are Owkin's platform, which is a multi-modal AI integrating clinical omics and imaging data. And then he has a AbSci's platform which combines generative AI with rapid wet lab validation, so the outcome was that they reduced target to candidate development to 18 to 24 months. And then they have multiple antibody programs progressing to the preclinical stages now.

So, another example is the collaboration between QURIS AI and Merck. So, Merck partnered with QURIS AI to integrate its bio AI clinical prediction platform into drug development, improving the preclinical safety evaluation of small molecule candidates, especially for drug-induced liver injury. After a two-year validation study showed superior accuracy over the traditional method, the AI, which was used in machine learning, was trained on QURIS AI's patient-on-a-chip data to predict daily nephrotoxicity with high precision. and the generative AI that simulates clinical trial outcomes using biomarkers and genomic diversity across interconnected organoid systems And then Exscientia is another key player that is developing personalized oncology treatments. So Exscientia uses AI to design personalized oncology treatments by analyzing patient tumor samples and optimizing compounds for specific tumor profiles.

So the AI it is using is like multi-objective optimization to select compounds targeting specific tumor characteristics and an AI-driven precision medicine approach tailored to individual patients. So you can see here, for example, you have that initial data integration, and then the AI is learning the features and designing the compounds. Compounds are being synthesized, tested, and then developed into clinical drug candidates. And then all the data generated during this process is fed back into the model to make it more learned. And then granite 001 is another example, which is an individualized neoantigen cancer vaccine designed to stimulate a patient's immune system to recognize and kill cancer cells based on their unique tumor mutations called neoantigens.

So the AI algorithm analyzes each patient's tumor DNA and RNA to predict the most immunogenic neoantigen mutations likely to trigger a strong immune response. So the tools they are using are Gritstone's Edge TCGA and COSMIC, so you can see here, for example, the patient data is obtained. So the standard tumor genome panel test plus HLA test, and then it has the shared neoantigen, yes, so then they have developed this slate 001. And if there is no, you know, neo and shared neo antigen, then you have the gritstone; using gritstone edge, it has, you know, private neo antigens.

Yes, then the granite 001 is being used. And then you have the in-situ, which has discovered a target for NASH known as alcoholic steatohepatitis. So, in vitro leveraged AI to identify novel targets for NASH using human genetics and cellular models, addressing unmet needs in a complex disease area. So, here the AI was used, you know, integrated with functional genomics to uncover causal pathways and therapeutic targets. And also the AI-driven analysis of genetic and phenotypic data to prioritize the targets. Okay, coming to the summary, so until now we have seen how AI has been used for discovering new drugs, new targets, hit optimization, and lead optimization.

So starting from drug target identification to its validation, hit identification, and lead

optimization. In the clinical trials as well, we have seen that there are several success stories that have shown the role of AI in accelerating drug discovery and development. So, there are multiple AI-discovered assets that have entered, you know, phase 2 clinical trials, and they are in an advanced stage. So, there are AI technologies such as knowledge graphs, generative models, and GNNs that are driving innovation in target discovery, molecule design, and repurposing. The clinical validation of AI-designed drugs shows promise in diverse areas like fibrosis, oncology, autoimmune disorders, and CNS disorders.

Partnerships and investments are growing, with AI-native biotechs collaborating with the pharma giant, and that is, you know, the need of the hour because. If the pharma companies that have a lot of data lying with them actually share that data with the companies that have that kind of expertise, AI expertise is essential, and only if they come together will it be successful in discovering drugs or new targets for diseases. So, we can say that AI is no longer experimental; it is becoming translational, proving capable of accelerating the pipeline from concept to clinical impact. Okay, in the end, I have an activity for you: choose a drug candidate that was discovered using AI and create a timeline tracking its progression through preclinical and clinical development stages. So you can pick, you know, any molecule like rentosertib; you can pick any other molecule as well and just track its progress, such as how the molecule was discovered using AI; either AI was used in the optimization of that molecule, or it was used in target discovery for that molecule.

So just track down its progress until it's progressed up to now, actually. So where has it reached, and what is the fate of that molecule? And then I have suggested some of the popular, you know, some of the important articles; for example, this is one of the seminal articles where you can see all that data. So, all the data about which of those drugs and molecules are in which stage of clinical trials. So, all those clinical assets you can obtain from this paper. And then there are some of these reports from the BCG, Boston Consulting Group, which has recently published a report on unlocking the potential of AI in drug discovery.

So, all of these resources are quite useful to know more about how successful AI has been in discovering and developing new drugs. And with that, thank you.