AI in Drug Discovery and Development

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

Week-07
Lecture-32

Welcome to the course "AI in Drug Discovery and Development." In today's session, we will talk about the importance of ADMET and its predictions in drug discovery. So, by the end of this lecture, you will be able to understand the importance of ADMET. Why early phase prediction is necessary in drug discovery and the various failures and success stories of some drugs due to suboptimal ADMET parameters. and how companies are utilizing early-stage ADMET prediction with various examples. As you have already seen, ADMET parameters or properties are very important because more than 50 percent of clinical trials fail due to safety.

And efficacy issues, and in all those failures, these ADMET parameters play a very important role. So, let us go through this important topic, you know. We have already discussed this. However, the admit is a critical part of drug development, helping to predict the drug's performance in the body and its safety profile before clinical trials.

So it helps streamline the process and minimize risk by identifying potential issues early in the development process. And this we have seen is that absorption is the process by which a drug enters the bloodstream after administration. and distribution where the drug is, you know, spreading throughout the body and reaching its target tissue. In metabolism, the process of chemical alteration of the drug in the body is usually done mainly by liver enzymes. and excretion, where the elimination of the drug or its metabolite from the body occurs mainly via urine, feces, or through the and the toxicity profiles, the harmful effects of a drug can often be related to a dose or the prolonged exposure to the drug.

So let us see why admitting is a major factor in drug failures. So, these admission properties play an important role as they account for the failure of around 60% of drug molecules during the drug development process. An early prediction of these properties would lead to a significant cost reduction in the field of drug research. So, if we can predict whether a drug will be toxic or not, a drug candidate will be toxic or not, and if it is toxic, So the early decision can be made, and you can save a lot of time as well as money. So, the effect of drug binding on the ADME properties of pharmaceutical proteins depends on factors like target quantity, turnover rate, location, and drug binding affinity.

So you can have a look at some examples like Pfizer's Exubera, which is an inhaled insulin.

It was discontinued due to poor absorption and safety concerns, emphasizing the critical role of admixture properties in drug development. There was also a drug called Vioxx, which is Rofecoxib, a selective COX-2 inhibitor. It was withdrawn from the market during phase 4 due to its unforeseen toxicity. So, there are actually some challenges in predicting admit properties during lead optimization.

During the lead optimization, what you have to do is improve both of these things. For example, the poor systemic exposure can come either from a distribution problem, meaning the drug is not being well distributed, or from clearance issues. So, it is cleared from the body really quickly. There are renal plasma and hepatic clearance; all of these play a role in it. And then in the distribution, in the case of distribution, these factors we already studied, like the volume of distribution, the blood-brain barrier, plasma protein binding, and those transporters which actively, you know, for example, if it is a CNS drug, there are these PZP and P-glycoproteins that actively expel those molecules from the brain.

so that they do not make it available for the CNS tissue regarding that drug actually. So likewise, another issue is the poor oral bioavailability, which can result from first-pass clearance if a drug is being cleared very quickly through hepatic metabolism. Or it is not stable in the gut, actually, so that can lead to poor bioavailability, or it can be a problem associated with absorption. So where are the physicochemical properties like pKa, solubility, or log P rendering the drug? Less absorbed in the gut, as well as, you know, reducing plasma membrane permeation, like the paracellular or transcellular permeation. So, because of these factors, the drug can have poor oral bioavailability as well.

So, these are two major factors that need to be optimized during the lead optimization stage because a lead is, you know, a molecule that should have optimal pharmacokinetic properties. and these are the two the major areas where we work on those you know those molecules to improve their bioavailability as well as the systemic exposure. Okay, so by using ADMET predictions, one can balance efficacy and safety. ADMET predictions help balance drug efficacy and safety by evaluating these properties early in development to optimize therapeutic effectiveness and minimize risk. So, we need to think about why highly potent molecules may fail due to poor admittance properties.

So despite their strong therapeutic effects, issues like inadequate absorption, improper distribution, harmful metabolism, or slow excretion. Or even high toxicity can limit their effectiveness, causing safety concerns or leading to drug accumulation, preventing successful clinical use. So, we will discuss one case study of terfenidine here. So terfenidine is an antihistaminic drug used to treat allergy symptoms, but it was withdrawn from the market due to cardiac toxicity because of its channel inhibition. So, it has been reported to cause cardiac death in at least 125 cases in the United States and 14 cases in the

UK, as reported by these authors.

So, the FDA recommended removing it from the market in 1997 due to its pro-arrhythmic risk for long QT-related torsades de pointes. And then it became, you know, not available in the clinic because of its associated toxicity. So, then there are several approaches by which to determine all those admitted properties of those molecules during the lead optimization. So, if we talk about physicochemical properties like solubility, it's a kinetic solubility at different pH levels because, in our body, different tissues have different pH levels; for example, the gut has a slightly acidic pH. And then the intestine has a slightly basic pH, and then blood has a neutral pH.

So, at different pH levels, we need to determine the solubility of molecules, and that determines the bioavailability of those molecules. So, there are physicochemical methods by which we can determine this kinetic solubility, and then there is another term called thermodynamic solubility. Wherein the thermodynamic solubility, you dissolve the solid compound at different pH levels, either in buffer or in the solvent of your interest. And then there is another thing called chemical stability at different pH and log D. So, all these properties need to be optimized, and that is done during the lead optimization stage.

And then with the membrane permeability and transporters. So, properties related to membrane permeability and transporters are used; usually, there is the PAMPA kit, the Parallel Artificial Membrane Permeability Assay. So, you either use the GIT permeability or the BBB permeability models. So, there are these commercial kits available where you can experimentally determine the permeability of those molecules, which is equivalent to or closely resembles the membrane permeability or the BBB permeability. And then with the CACO2 cell permeability, that is another permeability study assay, which can be used.

And then you have the PGP substrate identification, and then you can do the PGP inhibition as well, so all these assays—these are PGP, as I said, related to the CNS. where it renders molecules not available in the CNS by exporting them out of the brain actually. And then you have assays associated with determining protein binding, such as plasma protein binding, microsomal binding, or the blood-to-plasma ratio. And then there are assays that can be used to determine the metabolism as well, like microsomal stability or S9 stability. Which is determining the phase 1 and phase 2 metabolism, and then you can use the simulated gastric and intestinal fluid stability, plasma stability, CYP450 phenotyping, or metabolite prediction.

Or identification of CYP450 inhibition IC_{50} values, or phenotyping, or you can do, you know, CYP450 time-dependent inhibition as well, where you can determine the IC_{50} values. And then there are toxicity assays as well, where you can do the cytotoxicity in

various cell lines, use the MS tests, or use the HERGE assay, which actually indicates cardiotoxicity. So, these are some of the experimental methods that are being used for optimizing the leads in drug discovery. Okay, talking about those ADMET prediction approaches. There are basically two approaches: one is called in silico approaches based on molecular modeling, and the other is based on data modeling.

So, then you have, for example, molecular modeling; mainly, it is based on the metabolism. And then you have data modeling; you can either use the QSAR (Quantitative Structure-Activity Relationship) models or you can use the PBPK modeling. This is called physiologically based pharmacokinetic modeling. Then, by using those, you can obtain the comprehensive properties. And then by using all those ADMET tools, you can model the multiple pharmacokinetic parameters at once.

And if you look at these kinds of developments, basically molecular modeling tools, they are much less capable of predicting those properties. However, talking about the latest tools, admit software programs are much more capable of predicting multiple properties at once. Okay, so let us talk about the computational approaches. As I said, there are basically two data modeling approaches: KSAR and PBPK modeling. So ADMET analysis and predictions in KSAR mainly depend on many molecular descriptors, including topological, geometrical, physicochemical, or electronic descriptors.

And the PBPK modeling always predicts parameters concerning the dose size and dose frequency, such as the volume of distribution at steady state. Total drug clearance and the fraction of the dose that reaches the portal vein are important because most drugs are taken orally. So, you can see that in QSAR-based modeling, we use multiple descriptors or features to create a predictive model to predict all those ADMET properties. So, talking about ADMET absorption, the bioavailability depends on, as we discuss, the superposition of two processes. One is the absorption and another is the first-pass metabolism.

So, absorption in turn depends on the solubility and permeability of the compound, as well as interactions with transporters and metabolizing enzymes in the gut wall. So, some of the important properties for determining permeability seem to be the size of the molecule as well as its capacity to make hydrogen bonds, its overall lipophilicity, and possibly its shape and flexibility. So, these are some of the properties that are responsible for, you know, absorption. So, you see an example here, like there is this compound. So, it had a very low solubility of 0.03 micromolar and a logD of 3.2. However, when this methyl sulfonyl group was removed and the benzene ring was replaced with a pyridine. So, you could see that solubility improved a lot. You saw a steep rise in the solubility. However, the log D 7.4 was similar. So here, you know, the lipophilicity is the key admit property, which is an interrelationship between log D, log P, and ionization. So, for the monobasic compounds,

$$\log D_{pH} = \log P - \log(1 + 10^{(pKa - pH)})$$

you can say that the log D pH is equal to log P minus log(1 plus 10 to the power of pKa minus pH). For example, if the log P is 4 and PK is 9.4, then log D will be 7.4, which will be 2. And if the logP is 4 and PK is 7.4, then the logD 7.4 will be 3.7. And for monoacidic compounds, it will actually be like this.

$$\log D_{ph} = \log P - \log(1 + 10^{(pH - pKa)})$$

So, looking at this example called Entresto, which is a combination of sacubitril and valsartan, is an example of improved bioavailability. So, Entresto is a combination of sacubitril and neprilysin inhibitor, and valsartan and angiotensin II receptor blocker. So, both of these components exhibit pH-dependent solubility, being freely soluble at neutral pH but much less soluble at lower pH values. Additionally, WellSartan has low permeability, classifying Entresso as a BCS Class IV product, which typically presents challenges in achieving adequate bioavailability. So, WellSartan's poor bioavailability in standalone formulation posed a significant hurdle, ensuring that both SacoBitril and WellSartan were effectively absorbed when co-administrators required an innovative approach.

So Novartis developed a co-crystal complex of the sodium salt of saccharobitryl and valsartan in a hydrated form. So, this co-crystal formulation enhances the bioavailability of valsartan compared to its standalone version. So then, this co-crystal complex dissociates upon oral administration, allowing the active moieties to act independently in vivo, thereby improving the drug's overall efficacy. So, then there is another example of this BMS 986205, which is an IDO1 inhibitor. The reformulation of this compound improved its oral bioavailability.

So this compound inhibits the IDO1 enzyme to enhance the immune response against cancer cells. And the challenge was that the original formulation had very low oral bioavailability, limiting its effectiveness. So, reformulation improved the absorption and bioavailability for better systemic exposure. And the outcome was enhanced oral bioavailability, which makes it more effective and convenient for cancer immunotherapy. Talking about the admission in distribution means we need to target the right tissue.

So the plasma protein binding is a major issue that we have discussed earlier, and most drugs bind to the plasma protein very strongly. So that a very small amount of the drug is freely available in the blood, which is actually available for the target to bind to. So, the drugs bind to the plasma protein, affecting their free concentration, and high PPB limits

drug availability at the target site, while low PPB increases the free drug fraction and potential efficacy. Another thing is the BBB permeability, which determines if a drug can cross the BBB to reach the central nervous system or not. A good BBB permeability is essential for CNS-targeted therapies, such as those for neurodegenerative or psychiatric disorders.

Talking about loratadine versus terfenadine, you know terfenadine is an antihistamine that was withdrawn due to cardiotoxicity, which leads to QT prolongation. BBB permeability caused central nervous system side effects as well. Therefore, loratadine is a derivative of terfenadine with a structural modification that reduces BBB penetration. So, it improved the safety profile by minimizing CNS side effects. Greater BBB selectivity allows for effective treatment without crossing into the brain.

So then terfenidine, which was, you know, CNS toxic, was converted into loratadine, which is a derivative of terfenidine. So now this loratadine was not able to cross the BBB and was not entering the brain, thus avoiding the side effects shown by terfenadine. So, talking about the admittance metabolism, the rapid clearance reduces drug efficacy by eliminating it too quickly, while toxic metabolites can cause harmful side effects. So, optimizing metabolic stability ensures a longer drug half-life, and avoiding toxic metabolites enhances safety. So, the CYP enzymes metabolize lipophilic drugs, making them water-soluble for excretion and influencing drug-drug interactions.

One example is amoebifradil, which was withdrawn from the market due to its CYP3A4 inhibition, which led to a fatal drug interaction with other medications metabolized by this enzyme, causing serious side effects and safety concerns. So, let us discuss another example, the same as terfenadine. It is an antihistamine. So, it was sold by, you know, the company Hoist; later, it was converted to Sanofi with the brand name Saldane. So, the issue with terfenadine was that it was metabolized by CYP3A4 to the active compound fexofenadine.

So, when it was co-administered with CYP3A4 inhibitors, like ketoconazole or erythromycin. So, there was a lot of unmetabolized terfenidine accumulating. It was leading to QT prolongation and cardiac arrhythmias. So, the solution was that terfenidine was withdrawn due to the risk of cardiac toxicity and replaced with fexofenadine, which is the safe active metabolite.

So talking about those admissions and excretions. So, drug excretion is the final step in the ADME process and consists of a series of pathways that remove an administered drug or its metabolites from the body. Renal and hepatic clearance are crucial for preventing drug accumulation and toxicity. And we have seen that effective drug excretion ensures

that drugs are properly metabolized and eliminated, minimizing the risk of accumulation in the body. An example is Sidofovir, where the nephrotoxicity was due to renal accumulation. So, this OAT1, organic anion transporter 1, mediated the uptake of Sidofovir, which led to selective accumulation and toxicity in renal proximal tubular cells.

And then another example is Oraflex, which was developed by the Eli Lilly company. So, Benoxaprofen, which is a non-steroidal anti-inflammatory drug, is used to treat conditions like osteoarthritis and rheumatoid arthritis. So, there was a challenge with the excretion because Benoxaprofen had a prolonged half-life of approximately 20 to 30 hours. Primarily undergoing biliary excretion and enterohepatic circulation. However, in elderly patients and those with renal impairment, the half-life could extend up to 148 hours, leading to significant drug accumulation and side effects.

So, the accumulation of binoxaprofen in patients with compromised renal function led to severe hepatotoxicity and phototoxicity. So, the outcome of this was that due to this safety concern, binoxaprofen was withdrawn from the market in 1982. And this case underscores the critical importance of considering excretion pathways and patient-specific factors during drug development as well. Talking about the ADMET in toxicity, ADMET predicts drug toxicity by evaluating potential risks like organ genotoxicity and cardiotoxicity early in development. So, there are those in silico tools which can help identify safe drug candidates, reducing late-stage failures and clinical complications.

So why does the toxicity account for over 30% of the clinical failures? Because many of the drugs, despite showing efficacy in the early stage, may cause harmful side effects or organ damage that are not detected until the later phase of testing. So, these unforeseen toxicities lead to the withdrawal of the drug or halting development, making it a major factor in clinical trial failures. An example is Trovafloxacin, which was hepatotoxic, and its hepatotoxicity led to its market withdrawal. Talking about another drug, Troglitazone, which was sold under the name Resulin by Pfizer. So, it is an anti-diabetic drug; it caused idiosyncratic hepatotoxicity and liver failure.

Despite the preclinical safety data, human metabolism formed toxic quinone metabolites. So, there are over 60 confirmed liver failure deaths reported as being associated with the use of this drug. And why it failed toxicity prediction is because the human-specific metabolisms are not detected in animals, as most of the time those toxicity studies are performed in animal models. And then those animals sometimes do not recapitulate human metabolism, actually. So, some of those can be missed during these studies, and then the reactive metabolites are not flagged early because of the limitations of the assays.

And then the insufficient long-term toxicity monitoring was also not performed during the

development stage. Okay, so how does admit optimization avoid late-stage drug failure? If we can detect the toxicity early, it identifies risks like cardiotoxicity. or hepatotoxicity using in vitro and in silico tools like Protox-II or Direct Nexus, or there are multiple tools; actually, we will talk about them in later sessions. And then we can optimize the pharmacokinetics, which ensures proper absorption, distribution, metabolism, and excretion to improve drug effectiveness. And we can improve resource efficiency as well, focusing on the most promising drug candidates and reducing wasted resources.

At times, it can meet the regulatory compliance, which includes the safety standards, preventing delays. And rejections in the clinical trials, I would like to discuss this drug here, atorvastatin, which was, you know, optimized, admitting to a blockbuster success. So, you know the absorption-related optimization. High bioavailability for effective cholesterol reduction was needed, and then for metabolism. So, minimize drug-drug interactions, as it is metabolized by the CYP3A4 enzyme, and regarding excretion.

So, it has a low renal excretion of less than 1 percent, thereby reducing the kidney-related risks. So, atorvastatin is extensively metabolized in both the gut and liver by oxidation, lactonization, and glucuronidation, and the metabolites are eliminated by pyloric secretion and direct secretion from the blood to the intestine. So, that is how you can see that it has very low renal toxicity issues, and it has very good absorption as well. If we talk about the lessons from historical drug failures related to admissions, we can emphasize the importance of early toxicity testing and comprehensive screening to prevent late-stage failure. And if we can reduce late-stage failure, then we are saving a lot of time as well as the company's investment.

and then utilizing in silico tools for early risk prediction can help identify potential issues before the clinical trials like rofecoxib Which is a coaxed inhibitor was withdrawn from the market due to an increased risk of cardiovascular prothrombotic events that are being observed in the vigor. and approved trials. And then the withdrawal of cerevastatin involved an uncommon but serious adverse reaction that rhabdomyolysis. So, how you know those pharma companies they are using admit optimization. So, pharma companies like Pfizer and Novartis use optimization to screen and refine the drug candidates early, ensuring better absorption, metabolism, and lower toxicity.

And by understanding and predicting how a drug will be absorbed, distributed, metabolized, and excreted. By understanding its potential toxicity, companies can improve the chances of developing a successful therapeutic. And the process is heavily supported by computational tools, preclinical studies, and iterative design to balance efficacy and safety, which is crucial for getting new drugs to the market efficiently. So, we talk about the future of AI in admit modeling. So, the future of AI will enhance drug discovery by,

you know, improving the prediction accuracy, which will further lead to speeding up the development process and will also enable personalized medicine.

So AI will integrate omics data for better toxicity and pharmacokinetic predictions, streamlining regulatory processes, and a better explainability and model validation in AI-driven truck development ensure transparency, trust, and regulatory acceptance. And they are crucial for accurate predictions, clinical applicability, and meeting the safety standards. So, coming to the summary, these ADMET properties determine lead compounds' viability for further development, and early ADMET prediction helps prioritize leads with favorable pharmacokinetic and safety profiles. Inaccurate ADMET profiling is a major cause of late-stage drug failure, highlighting its role in optimization, and predictive ADMET modeling enables efficient refinement of chemical structures during lead optimization. Integrating AI-based ADMET tools accelerates the selection of optimized leads with higher success potential.

So in this session, we have discussed why ADMET optimization is important during the lead optimization stage and what some of the failures and success stories are in the pharma market. And in the coming sessions, we will discuss various AI-based tools that are being used for ADMET prediction during the drug discovery stage. So, in the end, I have some suggestions for you: you can go through this literature to get more info about this topic. And then I have an open question for you as well: Can a molecule truly be a lead if its pharmacokinetics are unknown? So just think it over. And with that, thank you.