

Predictive Physicochemical Property Modelling and Process Parameter Optimisation: AI-Driven Platforms for Pharmaceutical Formulation Development

Dr. Soo-Yeon Oh, Professor of Computer Science, Yonsei University, South Korea

1. Introduction to AI in Drug Formulation

Artificial intelligence (AI) is an advanced computerized tool that makes it possible to understand and process information. It is widely used in pharmaceutical sciences, especially in reviewing and analyzing extensive databases in a reasonable time. Although there is a reliable system for monitoring information via these stored databases, the AI-based methodologies for advanced drug formulation applications are limited. Incorporating AI systems in pharmaceutical manufacturing can smoothly and effectively reduce the number of drug formulation trials. Thus, AI-based applications in the development of pharmaceutical products promise flexible drug delivery system optimization based on both in vitro and in vivo evaluations with considerable cost and time savings.

The optimization of drug delivery systems is obligatory to reduce the shortage of the inadequate in vitro and in vivo evaluation of a drug formulation process before the animal or human stages. Integrating AI-based technologies with experimental protocols will help to optimize the formulations more rapidly. The use of AI in drug delivery systems changed the traditional trend, where researchers need to prepare and validate each formulation separately. Drug delivery systems prepared to mediate AI can predict activities and animal toxicity that facilitate a joint approach. AI finds special significance in the fast optimization of nanoparticles in the formulation of protein drugs. Fabrication of pharmaceutical metals can be monitored under the accelerated AI process. During conventional drug formulation development, almost 13–15 years and millions of dollars were used primarily for drug in vitro and in vivo evaluations.

1.1. Overview of AI-Driven Platforms

According to data provided on the life sciences market, the pharmaceutical industry has shown the fastest growth in the market for artificial intelligence (AI). Thus, it is not a surprise that several AI-driven platforms are now available for various aspects of drug formulation. Such platforms are mainly based on data analytics and machine learning capabilities. Based on their functionality, various platforms can be used to relate various factors to oral absorption and biopharmaceutical properties, to relate formulation building blocks to CMC properties, to relate excipient selection to drug-drug interaction suppression or enhanced dissolution, and to predict dosage form disintegration and/or dissolution profiles. Two AI technologies are currently available in the market: mechanistic modeling by several platforms and machine learning. As of the day of the coining meeting, platforms based on atomistic simulations are not yet in the market.

Data analytics, commercialized by several platforms, is an AI technology that may assist with oral exposure risk assessment and biopharmaceutical decision support by relating various physicochemical and pharmacokinetic properties to oral absorption. These platforms can be integrated into the overall new molecular entity selection and optimization process or combined with other predictive tools. A comparison of various AI technologies is provided in the following sections, and four illustrative case studies are provided. Finally, a discussion on AI-assisted formulation is provided in terms of potential uses, data issues, validation approaches, and user interface.

2. Challenges in Traditional Drug Formulation

The conventional process of developing new drugs is both time-consuming and highly inefficient. Technologies used to provide formulation for small beneficial molecules frequently lack the fine granularity needed to identify optimized candidate drugs. Time consumption and massive financial input make this a challenging project. Effects can normally be accomplished utilizing methods and empirical assessments that can also be carried out utilizing robotics and high-throughput screenings over a short period. Two tactics are commonly used in drug development. For optimization, strategies intended at enhancing stability were utilized, while strategies for enhancing bioavailability were utilized for verifying that the required dose might be released more effectively.

Several complexities exist in building fresh biomedical polymers that have new drug release applications and new mechanisms, despite the numerous years of formulation

technology. Formulation aims to maintain chemical stability, active substances, and biomechanical properties. The use of drugs is restricted, since polymer solvent properties are normally unsuitable when creating solid-state medicine for the pharmaceutical sector. Conventional methods for experiments and mistakes in pharmaceuticals are prevalent. Administering a hit or a packet three times does not imply that a valid medication can be developed to provide necessary treatment, as demonstrated in the formulations and experiments at any time. In each failure, the cause of the error is shown in the cost range only with component input. AI is predicted to have a position in a dynamically changing technological market, but not for the first occasion.

2.1. Stability, Bioavailability, and Delivery Issues

The stability of drug molecules represents a critical issue in drug development, with pharmaceutical companies investing significant resources to control degradation processes and extend the shelf life of their drugs. Consequently, state-of-the-art technologies are needed for assessing and optimizing not only formulation processes but also product stability. A decrease in bioavailability, or the rate and extent to which the active pharmaceutical ingredient (API) within a pharmaceutical product becomes available at the site of drug action, results in lower therapeutic effectiveness. In particular, the solubility and permeability of the drug molecule in the human body significantly influence bioavailability. Finally, effective delivery of the drug to the site of action—or more generically, to its intended target cells—is not always guaranteed. This is true not only at the formulation level but also on a micro- and nano-particle scale, for which nuclear dye-labeled nanosystems can be employed to show that some particles and nanodrops formulated with standard methods are not reaching the envisaged location.

The discussed issues undoubtedly deserve attention in the pharmaceutical field, which is why bioavailability and delivery processes are accurately studied through a series of scientifically sound approaches and methodologies. For example, a detailed protocol covering different stages of in vivo studies for improving the solubility, bioavailability, and stability of poorly water-soluble APIs has been reported. Similarly, some classic methods are well known for evaluating API release and dissolution from solid and semi-solid vehicles. Despite remarkable advances in this field, case studies are provided

herein to highlight the potential of novel advanced tools in identifying the aforementioned issues in API delivery.

3. Role of Machine Learning in Drug Formulation

Drug formulation can be optimized more effectively and efficiently by using AI-based modeling. Nowadays, many machine learning algorithms exist, making it possible to predict clinical outcomes or stability based on formulation variables such as excipients' concentrations or types. There is also an increasing ability to handle very challenging and computationally demanding stability data sets, demonstrating how enhanced directed stability can be achieved when a greater degree of complexity is taken into account. Excipients' effect scoring and interaction scoring can be performed. However, the final transformative power of an AI-driven modeling technique in the world of pharmaceutical formulation is in the area of predicting formulations' ability to remain unchanged under volatile conditions, be they temperature, storage conditions, or other changed environmental conditions.

Pharmaceutical development is often hindered by low bioavailability and solubility, which need precise excipient dosing to form an ultimately stable clinical formulation capable of treating patients. These excipients are expensive in terms of costs, and stability results are generally not available to help guide decision making. Recent formulation examples will be used to illustrate the power of a machine learning model in predicting the formulations likely to be stable in new systems and to help the process develop faster. That same AI model can also be used to evaluate the potential success of a stability enhancement strategy, predicting the most likely formulations to be successful.

3.1. Predictive Modeling for Stability Enhancement

One way of applying AI techniques in the field of drug formulation is to predict the stability of complex multi-component medications using predictive modeling. This type of model can forecast the stability of the formulation based on input variables that are easily changed, circumventing an expensive and inefficient trial-and-error development process. More and more computational and predictive models are being deployed in the pharmaceutical field, which speeds up drug formulation. When done properly, building a predictive model is relatively quick, significantly reduces the resources needed, and is environmentally friendly. When the formulator wants to select or design a stable

formulation, the algorithm will predict the probability of the product's practical shelf life. The developed models were able to handle complex datasets and were applied in a case study, which involved the selection of the best performing formulation. In some cases, it may be more efficient and faster to build a predictive model to simulate both the degradation profile and the effect of excipients on that profile. Nevertheless, reduced by the number of data points, excipient influences may not demonstrate this. Model building and down modeling are activities requiring a tremendous amount of reliability on input data. Information requirements for simple physicochemical models are, however, very different from life science models.

4. Applications of AI in Pharmaceutical Product Development

Pharmaceutical product development involves several technical and business complexities that have historically extended development timelines and increased drug product development costs. AI-focused technologies have been increasingly used in different stages of the development lifecycle. AI can be employed to accelerate the process of drug formulation efficiently, overcoming the challenges of traditional formulation development. From drug discovery to API and dosage form design, as well as clinical trials and manufacturing, multiple applications exist for AI integration. The incorporation of AI would allow easy maintenance of the high-quality standards defined by regulatory bodies and enhance decision-making within an organization. AI incorporation is not limited to the initial stages of drug development but continues until commercial manufacturing. Using the AI integrated formulation software, the formulation data of immediate-release tablets and topical creams, ointments, and gels were entered, resulting in the generation of TCFs with 0% of FIDs. This, in turn, led to an 87.5% overall formulation efficiency for software operational direct-by-design formulations. These studies have indicated that AI could be a useful tool in drug product development because it can provide immediate results. As AI is evolving, regulatory bodies even have to adapt to these fast-growing technologies, obligating them to enhance their infrastructure periodically. They are required to work on the guidelines to be followed during the utilization of these AI algorithms and also provide training to the reviewers, helping to assess these incorporated AI models for approval of the products. The challenges associated with the adoption of AI can be navigated by using alternative strategies to enhance explainability. AI agents can now provide analyses of their decision-making progress when integrating decisions from other

domain experts. Furthermore, this progress algorithm provides an explanation based on a decision tree for a decision that has been made. The decision is the basis of the real-world trial associated with the selection of the polymers within the software training sequence. It is reasonable to believe that the adoption of AI technologies in the exploration phase could lead to a higher degree of efficiency in the formulation of cannabis therapeutic products.

4.1. Enhancing Formulation Efficiency

AI platforms specializing in the formulation of small molecules and biological drugs in the industry aim to reduce the timeline taken to proceed from drug candidate to patient. Recently, significant steps have been made in contrast to traditional trial-and-error approaches, significantly reducing the time and resources spent on optimizing biological drug formulations. Such examples include small-molecule platforms that accelerate the optimization of medium and screening of crystallization methods, and engage in the optimization of lyophilization conditions for a range of formulations. It is worth noting that significant human expertise is involved not only in experimental execution but also in interpreting the AI-designed experiments' results. Although it appears to signpost the direction in which drug formulation development is heading, it is essential to recognize that data underpinning some of the claims made by such companies are still unpublished and unverified. As such, drug development managers should not be swayed solely by companies simply riding the AI-for-formulation development bandwagon without having an in-depth review of the data.

Adopting AI methodologies can help automate, thus speed up, the development pipeline all the way from formulation design to stability testing. For example, given the speed with which advanced analytics can provide insight, the platform can propose a new formulation within 45 minutes of training a data-analytic model with the solubility of a set of assay formulations. Paired with simulation, AI becomes more than a form of high-throughput experimentation and can help systems achieve a better understanding of the formulation components and their interactions, thus exercising more insight over experimental design. The most crucial component of developing AI models is human input. As a result, by not working with people who are developing or directly connected to the experimental formulation of interest under investigation, many models might be flawed. Critics argue that the only means of validating an AI platform is through its

actual performance relative to a traditional formulation approach. Making claims led by expert human intuition is at the very least biased and at most intensely confounding. In addition, drug development executives have concerns about AI platforms' purported risk of overstating the benefits of an approach that is decades away from being proven.

5. Future Directions and Potential Impact of AI in Drug Formulation

What might AI in formulation look like a decade from now? AI-driven formulation will become even more deeply integrated with all aspects of the formulation process. For example, AI models might automatically generate new types of drug delivery systems for clinical trials based on a wide variety of patient needs and convenience criteria. In particular, it is expected that various techniques will be pushed to a higher readiness level, all converging around the challenge of being able to design optimal drug formulation under all possible aspects, both in terms of API quality and molecular details for designing rational MFA, and in terms of the most effective and scalable formulation for targeting cost and time constraints.

Significant enhancements are expected for all AI components, from the developable drug and process models, MFA workflows, and the advanced logic AI engines. The full-scale integration of various expert systems into an intelligent platform will provide unique capabilities to the pharmaceutical scientists involved in drug development and production. A further development can lead to full integration of the AI-generated formulation outcomes with the pre-clinical, clinical, and commercial models and data. All these novel capabilities will require collaboration between pharmaceutical scientists and multidisciplinary experts in several fields, such as materials, food, and chemical engineering, as well as professionals in computational biology, bioinformatics, and pharmacometrics. Such a large range of AI applications and calculations will need to face significant regulatory, safety, and ethical issues, and require the highest level of AI verification and validation on top of the conventional qualitative and quantitative models, and new metrics for assessing the wide array of experimental validation challenges and data.

6. Conclusion

In summary, AI-driven platforms for the formulation of drugs offer innovative ways to solve classic industrial challenges that, so far, have not been answered using other analysis methods. The hope for these platforms is to offer more effective optimization

solutions. To do so, they rely on interactions between AI-based predictions and human decision-making, which aims, through AI and increasingly batch experiments, to explore a wider formulation design space and generate new, effective solutions and knowledge. This does not mean that AI has all the answers and needs to replace human judgments—it resonates with the message of applying AI for the enhancement of human characteristics to handle complex problems. Overall, AI technologies represent an innovative way of helping with the optimization of formulations, aimed at improving efficiency, accelerating analyses, and the best selection of formulations, and, crucially, predicting subsequent behavior and leveraging that prediction.

We believe that response surface characterization and model development/factoring strategies will lead to enhanced predictive power, the ability to use the results for more complex system optimization, and broaden the utility of these platforms in supporting the overall objective in drug product development. While the initial examples of these platforms are not currently capable of predicting crystallization, they have clearly demonstrated the impact and potential high value of AI-driven approaches in transforming formulation strategies. We contend that combining human judgment with current and AI-based analytical predictions can deliver platforms that enable the optimization of the drug product as a whole, considering complex compounded effects based on chemical and physical interactions across the system. Finally, examples of AI platforms using more efficient surrogate or faster predictive model development would significantly reduce experimental resource requirements. Further AI-driven development will enable the definition of current and future drug product function. In summary, AI-based platforms are already beginning to impact formulation strategies, with the potential for widespread impact in the near future. However, the approaches, training, and methods started here need to adapt and expand to deal with new analytical techniques and further change in the future drug formulation landscape. In general, we expect that AI will increasingly form the cornerstone for the optimization decision-making in formulation practices.