

Predictive Excipient Compatibility and Dissolution Modelling: AI-Driven Optimisation of Solid Dosage Form Formulation

Dr. Hatice Özbay, Associate Professor of Electrical and Electronics Engineering, Istanbul Technical University, Turkey

1. Introduction

It is widely recognized that drug development, the aim of which is to formulate new, quality preparations in user-friendly dosage forms, is becoming more and more complicated today. Therefore, new solutions based on modern technologies must be sought in order to maintain effectiveness, reduce costs, and increase the efficiency of the entire process—artificial intelligence (AI) is one of them. AI is especially important in the process of estimating the solubility of active substances in one-step matrix dosage forms. The use of AI at this stage of the formulation process will significantly improve the accuracy of the results, as well as allow for measurement to be carried out at much higher speed.

Drug development involves researching innovative new active ingredients as well as already known ingredients for which the patent has expired, that is, for which non-original preparations are formulated. Before a preparation is formulated in a user-friendly dosage form, the type of formulation should be determined, as well as the oxide forms. These tests are both time- and cost-consuming. The search for new solutions based on the AI model is supported by AI inhibitors, which enhance the accuracy of results. There is thus a need for global, multidimensional optimization, taking into account both the theoretical and practical aspects of the problem. We thus intend to formulate subject applications for a drug formulation process. Knowledge of the quality of the interactions between knowledge about wear and active disintegrants contributes to improved accuracy and efficiency, as investigators believe, is a transformation of the efficiency of their sciences.

1.1. Background and Significance

For many years, trial and error methods were employed for the discovery, invention, and optimization of drug formulations. These traditional methods include wet granulation and direct compression techniques and have significantly contributed to the advancement of pharmaceutical sciences. However, they are not free from limitations due to a lack of mechanistic conceptualization. For instance, wet granulation is slow and does not ensure good mixing, while direct compression leads to intermixing. Moreover, the old paradigm demands labor-intensive materials and time. In response, we are in the era of the so-called fourth industrial revolution. This era is characterized by rapid advancements in artificial intelligence, big data analytics, high-throughput methods, and systems biology approaches that allow researchers, scientists, pharmacists, pharmaceutical engineers, biologists, and business entities to process, analyze, chemometrically deduce, correlate, and predict new product development insights.

The inclusion of AI technologies in the drug development process has several potential benefits for pharmaceutical scientists and manufacturers. AI has been growing rapidly, mainly due to its performance in data-driven tasks. In fact, AI applications have featured as primary parts of groundbreaking technologies. A review of the literature has found that AI is used in precision medicine, predictive analysis, genomics, metabolomics, personalized medicine, diagnostics, drug discovery, protein structure prediction, de novo molecular design, computational biology, ligand-based drug design, image analysis, modeling, molecular docking, biophysics, biomass, pharmacology, biotechnological processes, and analytical chemistry, among others. Aside from identifying novel uses for AI in the pharmaceutical industry, researchers are also exploring predictive modeling capabilities of artificial neural networks. This and several other studies focus on a variety of tools to optimize drug processing. To illustrate, commercial firms are developing AI-based tools for optimizing pharmaceutical unit operations to compete in a rapidly evolving yet economically challenging business environment. Integration of AI in the pharmaceutical industry value chain will ensure patients benefit from modern cutting-edge drugs. Moreover, this will facilitate a competitive and sustainable supply chain.

2. Fundamentals of Drug Formulation

The development of new pharmaceutical products typically consists of multiple stages, including drug formulation, pre-clinical and clinical development, and final drug release. Formulation fundamentals were examined and predicted by drug-discovery platforms that could be considered a key to the beginning of drug design. This review presents the fundamentals of drug formulation as an essential part of pharmaceutical product development. It highlights the basic requirements for the drug formulation process that most drugs pass through and also identifies, on a molecular basis, several areas where artificial intelligence can support this traditional, almost manual process to increase drug efficiency.

The most commonly prescribed drugs are presented in the form of solid and liquid formulations; semi-solid forms, such as creams or ointments, are used for skin-related problems. In addition to serving the purpose of drug administration, drug formulations can also be readily available, have good stability, long shelf life, and good appearance. The drug formulation processes involve two important steps: the first step is to prepare a drug product, and the second step is to test the drug. A major part is the preparation of drug products, which involves several processes such as powder mixing, granulation, compression, tablet coating, or stabilization in capsule formulation followed by formulation. This overview provides aspects that are fundamental for drug formulation strategies. Each form of drug has its own characteristic properties depending on the forming process.

It is influenced by critical factors such as the mixing type, used excipients, flow behavior, and the amount of lubrication required. Due to the function of the drug and its dosage form, the critical properties of the drug must meet certain quality attributes to ensure drug efficacy. Safety and quality are the three most critical aspects of pharmaceutical products that must be taken into account throughout the formulation process. Bioavailability, among others, is one of the most crucial aspects of drug formulation. In addition, the interaction between different excipients within the formulation must have a synergistic effect. Therefore, the identical components of all newly discovered medications, the excipients, should be selected. Success in drug formulation is evaluated based on the possible obstacles. Dosage formulas demonstrate unpredictable deviations rather than generally utilized compound modeling systems. The physical and chemical

characteristics of a compound can inhibit its effectiveness and create obstacles if not correctly adapted to various requirements such as acceptable bitterness, bad aroma, a delicate texture, and a variety of other requirements tempting to customers. A combination of a few of these issues may lead to suboptimal compound dosages.

2.1. Key Concepts and Processes

Creating pharmaceutical products involves a broad array of interconnected concepts that revolve around formulating drug delivery systems, including the mechanism of drug release at the sites of action and absorption. The production includes the design of formulation and process, and all necessary batch trials and tests to prove theoretical assumptions in practice. An important part is also technological validation, which includes optimization research and stability testing. In contrast to process optimization in other modern industries, in pharmaceutical production, optimization is an iterative process, especially when it is oriented to the minimization of production costs. To optimize the strategies, formulation characterization and optimization methods have been developed, and the improvement of those is still an active field of research. This chapter includes the most important aspects of the formulation process, from initial steps through evaluation method development and characterization. Formulation strategies and evaluation approaches presented in this chapter are based on a combination of science and regulatory requirements.

Formulation optimization is an integral part of any drug development process. Formulation, as a separate area, is related to the analysis of formulations' properties and helps to develop better formulations based on the results of the analysis. Formulation and in-process controls are an integral part of the pharmaceutical quality system. Developing new technologies and data management approaches permits further opportunities for evaluating excipient attributes that are now considered critical and evaluating the influence of manufacturing parameters on product properties. There is a distinction between scientific and regulatory points of view, which are exposed in different strategies of optimization. Generally, in order to propose a specific optimization strategy for the formulation of a specific active substance, knowledge-based methods are applied. There are a number of parameters that may be considered for formulation characterization and optimization processes. Most of them can be used as barriers for drug quality, stability, and efficacy, and include pharmacokinetic

parameters, bioavailability, patient compliance, toxicity, and therapeutic equivalence or interchangeability, API release mechanism, and kinetics from the solid delivery system. Key characteristics of the delivery system that can be optimized are as follows: stability, drug release mechanism, API loading in the delivery system, and API release and kinetic rate from the delivery system. The properties of the optimal formulation must limit the possibilities of misuse and abuse. Control of stability and in vitro drug release in correlation with in vivo studies is necessary. We need to recognize to whom the results of formulation design can be applicable. We must determine the target product profile before we can establish what we are trying to test. Concentration in the body, route of delivery, and desired pharmacological effect are the main areas to focus on.

3. Role of Artificial Intelligence in Drug Formulation

With the evolution of technology, artificial intelligence (AI), especially machine learning and neural network-based approaches, has started showing its significant impacts in various sectors, including the development of pharmaceutical sciences. In particular, the drug formulation area is a subsector that can achieve more benefits compared to other pharmaceutical subsectors, as most problems in pharmaceutical development start at the formulation level. Formulations help in the diagnosis, management, and delivery of therapeutic agents in the human body. However, the traditional route of formulation development and optimization is flawed; it is based on the formulation scientist's experiences rather than satisfactory data. Moreover, most of the formulation optimization processes are challenged by the generation of big data and the associated lengthy work processes. AI-based tools can streamline the interest region and reduce both economic and time costs. One major application of AI is in the quality by design concept, enhancing the risk assessment of undesired parameters.

The ultimate aim of the formulation scientist is to maximize the performance of the drug product by considering different drug substances, excipients, processing conditions, patient compliance, and stability while minimizing the overall risks and costs. However, the development time of a formulation depends on the trial-and-error-based product assessment. In all these steps, scientists have to optimize the right composition of the drug substance and excipients to avoid imperfections. Formulation scientists also have to create forms that humans can easily ingest while preserving the necessary properties. The biopharmaceutical classification system class also has to be identified for the right

design of dosage systems. In addition to the aforementioned points, the possibilities of a high-quality drug product should be successfully scaled up at the relevant level to achieve the desired performance in dosage forms. AI can optimize formulations and predict new drug actions. The AI method aids formulation development and troubleshooting processes and selects an optimized formula by providing potential data.

3.1. Overview of AI Technologies

With the increasing demand for pharmaceutical products, drug formulation has become one of the most critical steps in the healthcare industry. Formulations assist in converting active pharmaceutical ingredients into effective and stable drug products. Development of a formulation involves a variety of stages including preformulation studies, formulation design, optimization of formulation methods, and eventually the scaling up of the formulation. Since these steps contribute to the cost and time required during pharmaceutical drug development, there is significant interest in developing technologies that can facilitate the formulation process. The emergence of artificial intelligence and associated technologies has the potential to transform the conventional methods in this domain.

The fields of artificial intelligence have found various applications in pharmaceutical development. These include, but are not limited to, machine learning, data mining, natural language processing, and predictive modeling. Data mining techniques have been used to systematically learn patterns from bioinformatics data in rational drug design. Natural language processing has been used in text mining to unveil the hidden semantics from structured scientific articles and patents, and it can automatically extract useful information from an abundance of text data. Artificial intelligence technologies can potentially be applied to address these challenges and provide sophisticated, efficient, and effective pharmaceutical formulation insights.

Artificial intelligence technologies possess a plethora of advantages for the pharma sector, empowering the industry to optimize its formulations. Artificial intelligence techniques improve the speed and accuracy of the formulation process; help in the fast and early identification of critical quality attributes and critical process parameters, controlling effects, and establishing process design space; quantifying the potential dependence of quality on parameters within a design space; and maximizing accuracy and consistency in predicting the risks associated with variations in the formulation

inputs and manufacturing processes. The current reality also shows multiple examples of successful artificial intelligence applications in the formulation design and development process. Research has identified that artificial intelligence can be leveraged to predict the critical quality attributes of a submicron product among the various variables considered by formulators and researchers in pharma companies. Besides, it is evident that artificial intelligence techniques have the ability to quickly capture the design space based on programmed constraints and requirements, thus eliminating trial-and-error experimentation.

Despite their large list of success stories, the usage of artificial intelligence in drug formulation is not free from challenges and drawbacks. Some of these challenges include the involvement of high-level thinking leading to the need for interdisciplinary expertise in formulation, software programming, and regulatory affairs, the necessity of large, clean, and contextual datasets for the artificial intelligence solutions to be effective, and the requirement for integration with internal data sources, such as pharmaceutical internal databases and IT systems. Other challenges include the existence of potential bias and overfitting, predominated by data rather than domain knowledge, and the potential incorrect use of early-stage formulation data due to the high contribution of inherent data variability. Since available datasets are mostly scattered across various systems and environments, the collection of data into artificial intelligence-conducive datasets is a major stumbling block. However, the increasing availability of data in the future, together with the rapid maturation of artificial intelligence and IoT technologies, could play a crucial role and provide a turning point for artificial intelligence-driven drug formulation advancements.

4. Applications of Machine Learning in Pharmaceutical Formulations

Pharmaceutical formulations, due to their complex composition, often result in very large datasets. Machine learning enables the mining of these data to uncover new insights, trends, and patterns. Predictive models are particularly useful, as they enable the forecasting of formulation outcomes prior to actually completing the experiment. This aids in guiding the process and making decisions, using more than intuition to predict the success of a particular formulation. A range of machine learning techniques are available, which can all be employed for a successful formulation outcome. For

example, regression analysis applies mathematical equations to estimate the relationship between formulation input and resulting outputs.

However, the use of machine learning comes with its own pitfalls. Foremost is ensuring the quality, accuracy, and comprehensiveness of the data being fed into the model, to ensure the output is meaningful. A lack of data can further hinder the performance, interpretability, and robustness of the machine learning model. Furthermore, as the pharmaceutical industry is highly regulated, any results obtained are carefully examined and validated before being implemented. Finally, all successful formulations by machine learning approaches need to meet the standard regulatory requirements. Despite these challenges, machine learning has been successfully applied for the predictable optimization of formulation compositions in several case studies.

4.1. Predictive Modeling and Optimization

4.1. Predictive Modeling and Optimization

In any drug formulation application, it is always advisable to move away from reactive mode in product development to a more proactive mode. The formulation is influenced by a myriad of characteristics that exist or can be introduced in the process. Rightsizing these sources to desired outputs before an undertaking to build to meet the desired quality and efficacy is always advantageous. The capability to instantly let us know of the desirability of any initial decisions is at the heart of predictive modeling, with capability limits imposed. Such a capability unlocks the potential of several formulation scenarios and provisional decisions that can be evaluated prior to implementation. Additionally, one can realize the benefit that can accrue if such decisions are implemented in the formulation process. Furthermore, it is always possible to refine the model in light of historical data as well. Any prediction can be further validated with ongoing experimentation in the laboratory, and history can be accordingly recorded.

In a competitive environment such as pharmaceutical product development, the time and cost of product development can be significantly reduced if the vision into the future of the formulation development can be obtained, which thus helps in minimizing adjustments to meet performance criteria. The optimization aspects of the predictive modeling technique discussed herein are well documented with several examples in pharmaceutical formulations and processes. This approach in the aforementioned

context reduces time and resources in drug product development. It envisions the potential outcome of different scenario formulations ahead of the lab work. It reduces the number of laboratory-scale formulations needed. The statistically designed development trials avoid basing the development on guesswork. Additionally, it aids in step selection that simultaneously balances the goals of economics while protecting outcomes from probabilistic variation. Conventional pharmaceutical development relies on explorative models that describe the response surface of a pharmaceutical product. Later optimization tries to use these linkages to explore some specific formulation space and apply the know-how for the best formulation. With the model-building tools of today's science, the methodology finds the approach too much a posteriori rather than a priori. Conventional development, pre-engineered and pre-qualified validations, and verification studies are carried out in predictive modeling. A hybrid approach that encapsulates machine learning with predictive modeling is also available. The examples in the following illustrations adopt predictive modeling using computational approaches.

5. Challenges and Future Directions

Challenges and Future Directions. A number of aspects related to the integrative role of AI strategies in drug formulation may present challenges. The integration of AI in the drug development lifecycle should meet a number of guidelines, indicating a high level of transparency regarding the adopted AI algorithm, input features, and the utility and importance of the predictive model for carrying out dose optimization in drug formulation. Data quality, reliability, and liabilities will require attention and certification/disclosure. Automating the metabolomics data interpretation and risk mitigation processes is an ethical concern that would require extended consideration from the ethical and legal community. For example, the use of AI-based algorithms allows access to highly confidential patient data, which raises privacy concerns. In addition, algorithmic bias can sometimes occur, as AI systems suffer from pre-existing biases in the underlying data.

On the technological level, scalability is a major challenge, particularly in the context of developing an informatics platform for drug combinations with high throughput, high-content assays. A number of limitations regarding data availability are also present for the AI integration of metabolomics data with systems biology models, e.g., large-scale

models. Furthermore, AI for physics presents another challenge, as AI models trained on relatively small experimental data would need to learn highly complicated molecular theories and equations, which are normally obtained from decades of theoretical developments. The application of AI strategies on informative metadata promises to guide the acquisition and curation of high-quality multi-omic datasets, though this would require extensive regulations to oversee the quality assurance of this data and meta-analysis properties. Regulatory compliance is a potential challenge, especially given the highly experimental nature of the AI-based integrative framework.

Future Directions: AI-based techniques can dramatically enhance the ability to take an integrative and global view of the relationships between drug characteristics and patient health. The formulation process may be further empowered by multiscale and multiphysics models coupling drugs. The application of these considerations in the context of the drug formulation process is also a subject for future work. Regulatory bodies need to engage in continued dialogue with technologists, scientists, and drug developers to facilitate and oversee the development of AI technologies for drug development that maintain workflow autonomy and discuss and resolve associated ethical, regulatory, and safety requirements.

5.1. Ethical Considerations and Regulatory Compliance

A number of ethical issues related to technological optimization in the area of drug formulation processes require further examination. The potential of AI in this space might lend itself to putting innovation first and may not, by default, highlight or prioritize the ethical considerations required in this type of biomedical research. The possible implementation of AI-optimized drug formulation processes may influence issues of informed consent and data processing or data privacy, especially when patient data is processed. The use of shallow or cryptic models that do not easily disentangle variables taken into account to make a decision is a major concern in the ethical scrutiny of algorithmic decision-making.

This category of AI methods bases its decision on a myriad of variables, all of which may be linked to diverse axiological implications, like socioeconomic status or victimhood status. An ethical question that often arises is how, and if, AI can ever be made unbiased or neutral. It is important to establish regulatory guidelines that can operate robustly across different sectors or that can overcome concerns arising from data protection, for

instance. Those interested groups and government-backed regulatory bodies must be able to interpret and understand new guidelines and ensure that those guidelines adequately cover and regulate AI-based applications. A dataset can only be as good and as ethical as it is representative of a diverse range of stakeholder views and as biases—or indeed the lack thereof—are aligned with transparency norms and stakeholder consultation throughout development. Therefore, it is recommended that any AI optimization activities be embedded in a wider ethical framework, where all stakeholders have an equal say in what is perceived as ethical or trustworthy.

6. Conclusion

In this paper, we demonstrate that AI can greatly enhance the drug formulation process, mainly by speeding up the process and optimizing the results in a broader scope. We discussed the current and foreseeable future perspectives of this technology, outlining both its limitations and the critical points that still need to be faced. Further, we analyzed the drug formulation process in all its aspects, pinpointing the current bottlenecks and problems that lay ahead, as well as the new challenges that need to be faced. In conclusion, AI-approaches have the potential to revolutionize the present approach to drug formulations technologically, chemically and biologically. Thanks to the wider applicability of the methods proposed and the possibility of taking an integrative view ruling the constraints deriving from a time-efficient approach, they represent a cornerstone for future pharmaceutical development. The activity of the pharmaceutical formulator is now aimed at discovering a correct sequence formulation method that could combine drug release rate and kinetic modality in the single product, preferably starting from a conventional oral formulation. Given the future possible change in the pharmaceutical approach over the incoming decade, this new area is capable of attracting the interest of a significant and still increasing number of professionals. Further possible regulatory pathways for such complex products are open to debate within the agencies and could benefit from the view of the pharmaceutical industry. Prior to any real and safe development of such products with AI, the cautions for micro- and nanosystems development should be properly considered. The present manuscript has gathered and discussed in a common framework the present and foreseeable future methodologies capable of speeding up and ameliorating the pharmaceutical development. Invalidating compounds at an early step of the pre-formulation process and not on the basis of long and expensive developing pathways

could also lead to a wider knowledge of the biomolecule itself that attributes a better safety-related profile to new patients. We believe that the efficient cooperation of the academic community, regulatory agencies, and the pharmaceutical and industrial sectors could further promote the correct development of these new systems on fields related to the system analyses.