Drug Design, Development and Therapy

Dovepress
open access to scientific and medical research

Open Access Full Text Article

PERSPECTIVES

The Coming of Age of AI/ML in Drug Discovery, Development, Clinical Testing, and Manufacturing: The FDA Perspectives

Sarfaraz K Niazi

College of Pharmacy, University of Illinois, Chicago, IL, USA

Correspondence: Sarfaraz K Niazi, Tel +1-312-297-0000, Email niazi@niazi.com

Abstract: Artificial intelligence (AI) and machine learning (ML) represent significant advancements in computing, building on technologies that humanity has developed over millions of years—from the abacus to quantum computers. These tools have reached a pivotal moment in their development. In 2021 alone, the U.S. Food and Drug Administration (FDA) received over 100 product registration submissions that heavily relied on AI/ML for applications such as monitoring and improving human performance in compiling dossiers. To ensure the safe and effective use of AI/ML in drug discovery and manufacturing, the FDA and numerous other U.S. federal agencies have issued continuously updated, stringent guidelines. Intriguingly, these guidelines are often generated or updated with the aid of AI/ML tools themselves. The overarching goal is to expedite drug discovery, enhance the safety profiles of existing drugs, introduce novel treatment modalities, and improve manufacturing compliance and robustness. Recent FDA publications offer an encouraging outlook on the potential of these tools, emphasizing the need for their careful deployment. This has expanded market opportunities for retraining personnel handling these technologies and enabled innovative applications in emerging therapies such as gene editing, CRISPR-Cas9, CAR-T cells, mRNA-based treatments, and personalized medicine. In summary, the maturation of AI/ML technologies is a testament to human ingenuity. Far from being autonomous entities, these are tools created by and for humans designed to solve complex problems now and in the future. This paper aims to present the status of these technologies, along with examples of their present and future applications.

Keywords: FDA, artificial intelligence, machine learning, drug discovery, drug development, advanced manufacturing

Introduction

The learning of facts by humans comes with their intellectual bias, based on prior learning, beliefs, and genetically driven survival instincts. None of these factors enter when a machine is involved since it cannot think beyond what it is taught. Artificial intelligence (AI) and Machine Learning (ML) definitions vary.[1] ML is a subset of AI that enables ML models to be developed by ML training algorithms through data analysis without being explicitly programmed. AI is a branch of computer science, statistics, and engineering that uses algorithms or models to perform tasks and exhibit behaviors like learning, making decisions, and making predictions. Notably, this computer software can only report or analyze the available data as input or based on reported simulations. This paper presents a historical perspective of the evolution identifying significant events that have led to the evolution of scientific and computational events that now constitute AI/ML, as shown in Table 1.

Optimizing the Utility of AI/ML

If AI/ML were limited in use by corporate and industrial laboratories, there would not have much debate about the risks of their service until the arrival of the public-use ChatGPT4,[30] which stirred up controversies, suggesting that the technology must be regulated,[31–33] as if this were an alien intervention. But this is nothing new; we continue to live with myths and perpetuate them, as the nature of humans and an AI-driven tool can only replicate what we think, as shown in Table 2.

Received: 28 June 2023
Accepted: 24 August 2023
Published: 6 September 2023

Table 1 A brief history of AI/ML development interventions

Year	Intervention
1950	The paper by Alan Turing, "Computing Machinery and Intelligence", proposed the concept of a "universal machine" capable of exhibiting intelligent behavior. [2]
1956	The Dartmouth Conference marked the birth of AI as a field, bringing together leading scientists to discuss AI research and define its goals.[3]
1956	The Logic Theorist program, developed by Allen Newell and Herbert A. Simon, was the first AI program to prove mathematical theorems.[4]
1957	Frank Rosenblatt's Perceptron introduced the concept of a single-layer neural network capable of learning through a process called "perceptron learning."[5]
1957	The General Problem Solver (GPS) program, developed by Newell and Simon, demonstrated an approach to problem-solving using symbolic reasoning.[6]
1972	The development of the PROLOG programming language by Alain Colmerauer and Philippe Roussel brought about significant advancements in logic programming and knowledge representation.[7]
1976	The MYCIN system, developed by Edward Shortliffe, demonstrated expert systems in medical diagnosis, utilizing a rule-based approach. [8]
1982	The R1/XCON system by Douglas Lenat and Randal Davis was a notable expert system used in configuring computer systems, showcasing the power of rule-based reasoning.[9]
1983	John McCarthy coined the term "artificial intelligence" and developed the Lisp programming language, which became a significant language for AI research.[10]
1983	The CYC project, led by Douglas Lenat, aimed to create a vast knowledge base encompassing common-sense reasoning and understanding.[11]
1988	The Backpropagation algorithm, proposed by Paul Werbos, enabled efficient training of multi-layer neural networks.[12,13]
1997	Deep Blue, developed by IBM, defeated world chess champion Garry Kasparov in a six-game match, showcasing the potential of AI in complex strategy games. [14]
1998	The LeNet-5 architecture by Yann LeCun et al revolutionized the field of computer vision and became a fundamental model for image recognition tasks.[15]
2009	The ImageNet project, led by Fei-Fei Li, introduced a large-scale dataset and benchmarks for training deep convolutional neural networks for image classification.[16]
2011	Watson, an AI system created by IBM, won the Jeopardy! Game show against human champions in 2011, demonstrating natural language processing and knowledge retrieval abilities.[17]
2012	The AlexNet architecture by Krizhevsky, Sutskever, and Hinton revolutionized image classification tasks and demonstrated the power of deep learning on GPUs.[18]
2016	AlphaGo, developed by DeepMind, defeated the world champion Go player Lee Sedol, highlighting the advancements in machine learning and reinforcement learning.[19]
2017	DeepMind's AlphaGo Zero surpassed the performance of the original AlphaGo by mastering the game Go through reinforcement learning without any human expert knowledge.[20]
2017	The Transformer model, introduced by Vaswani et al, revolutionized natural language processing tasks by enabling efficient attention-based sequence-to-sequence modeling.[21]
2018	Devin et al created the BERT (Bidirectional Encoder Representations from Transformers), achieving phenomenal results in processing in multiple natural languages that included tasks, including question-answering and sentiment analysis.[22]
2019	OpenAI's AlphaStar defeated professional human players in StarCraft II, showcasing the potential of reinforcement learning in complex real-time strategy games.[23]

(*Continued*)

Table 1 (Continued).

Year	Intervention
2021	AlphaFold wins the CASP competition for predicting the 3D protein structure.[24]
2022	Organizations and initiatives, such as the Partnership on AI,[25] OpenAI's Charter,[26] and the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems,[27] have emerged to address these challenges.
2023	ChatGPT arrives, bringing much controversy.[28]
2023	The FDA issues two papers disclosing how the FDA is encouraging the adoption and applying AI/ML in improving drug discovery and that more than 100 regulatory filings to date have employed AI/ML approaches.[29] Also released is the FDA Guidance

Table 2 Common myths of daily life

Field	Misconception
Computers	ChatGPT Can Destroy Humanity[34]
Biology	Bats are Blind[35]
Astronomy	Black Holes are Cosmic Vacuum Cleaners[36]
Physics	Centrifugal Force[37]
Psychology	Cognitive Biases and Heuristics[38]
Physics	Glass is a Slow-Moving Liquid[39]
Physics	Heavier Objects Fall Faster[40]
Biology	Humans Evolved from Monkeys[41]
Biology	Humans Only Use 10% of Their Brains[42]
Physics	Lightning Never Strikes the Same Place Twice[43]
Climate Science	Misconceptions about Climate Change[44]
Biology	Misconceptions about Evolution[45]
Animal Behavior	Ostriches Bury Their Heads in Sand[46]
Ancient History	Roman Vomitoriums[47]
Astronomy	The Dark Side of the Moon[48]
Chemistry	Water Conducts Electricity[49]
Medicine	We Should Drink 8 Glasses of Water a Day[50]
History	Witch Burnings in Salem[51]
Medicine	You Can Sweat Out Toxins[52]

However, these algorithms, acting much faster than humans can, also make and multiply mistakes. It is this concern that the FDA and other US Federal Agencies have established stringent controls on using AI/ML, where applicable, in developing new drugs.

- Lack of Accountability: One argument against the widespread use of AI/ML is that it lacks accountability since it can generate false or misleading information. However, it is crucial to recognize that AI/ML are tools, and computer software, developed by humans and trained on vast amounts of real-world data. The responsibility for using the

content generated lies with the users, planning to utilize and deploy the technology. They are responsible for ensuring the accuracy and reliability of the information shared. This expectation is no different from the responsibility of the developers in validating the data reported to the FDA.

- Ethical Concerns: Concerns are frequently expressed regarding the ethical implications of AI, particularly in areas such as privacy, bias, and fairness. While these concerns are valid, it is essential to note that the ethical issues do not stem from the AI system itself but rather the data it is trained on; AI cannot make a judgment; it is a regurgitation tool. Regulating AI for drug development has the greatest potential and risk if the AI-generated data are not validated.[53]
- Job Displacement:[54] Critics argue that AI technologies will lead to massive job displacement, leaving many unemployed. However, historical evidence suggests that technological advancements have often created new jobs and transformed existing ones. While AI may automate specific tasks, it can also augment human capabilities, leading to increased productivity and new employment opportunities. Regulations should focus on ensuring a smooth transition by fostering education and upskilling programs rather than stifling innovation that will significantly increase new job opportunities.
- Safety and Security:[55] Concerns are raised in the public media about the potential misuse of AI systems, such as spreading misinformation or generating harmful content. However, there should not be any ethical concerns about using AI/ML for drug development; for example, the AI-driven protocols will not bring alternates to animal testing to prevent currently against animals; how it is possible that if literature discusses such possibility, the AI may propose, entirely on the perspective of reporting not judging. A good example is how one can easily get biased and racist responses from ChatGPT based on how the question is designed. The focus of the search of ChatGPT shifts based on the nature of inquiry with any bias or the ability to judge.
- Unintended Consequences:[56] Critics worry that AI systems may have unintended consequences, such as amplifying data biases or generating unexpected analyses or results. While these concerns are valid, they can be addressed through ongoing research, development, and responsible deployment. Building AI systems that are transparent, explainable, and subject to continuous evaluation and improvement can help identify and rectify any unintended consequences.

One prominent use of AI/ML has revolutionized drug discovery, development, and manufacturing. However, to minimize the risk of faulty decision-making, the FDA and dozens of other agencies have taken significant steps as guidelines that present the control requirements while suggesting more engagement of these new technologies. AI/ML is now widely used to discover its applications, which have led to the entry of many novel tools and approaches that can help create new entities and reduce the cost of development—a dire need to make drugs affordable.

The development and usage of pharmaceuticals will accelerate because of recent technological advancements in data gathering and generation tools, reliable information management and exchange platforms, and enhanced computational capabilities. According to the FDA's latest studies, this dynamic ecosystem offers distinct opportunities, difficulties, and ways to increase their use.[57]

Various drug development efforts have used AI/ML to hasten the drug development process and improve the efficiency and safety of clinical trials. However, it's critical to determine whether AI/ML poses risks or harms.

When results are extrapolated outside of the testing environment, ethical and generalizability issues arise since the AI/ML algorithms have the potential to exaggerate flaws and preexisting biases found in the underlying data sources. As a result of its inherent complexity, an AI/ML system may also have limited explainability or may not be entirely transparent for proprietary reasons. The FDA is creating guidelines for trustworthy AI to address specific qualities, including explainability, reliability, privacy, safety, security, and bias reduction.

The fact that the FDA has raised these concerns points to better use of AI/ML. Some of these concerns can be better understood from examples drawn from current research efforts, as presented in Table 3.

Drug Discovery

One of the areas where AI/ML is being used extensively is early drug development, which includes compound screening, drug design, target identification, target selection, and target prioritization.

Table 3 Examples of Recent Applications of AI/ML in Drug Discovery Applications

Adverse event detection Identify adverse events associated with drugs by analyzing electronic health records, social media, and other real-world data sources. This can assist in early detection, reporting, and mitigating potential drug risks.[58]
Antibiotic discovery Antibiotic resistance is a major public health concern, and there is a need for new antibiotics. Help discover new antibiotics by analyzing vast amounts of chemical and biological data to identify potential antimicrobial compounds.[59]
Autoimmune diseases Aid in developing drugs for autoimmune diseases by predicting disease mechanisms, identifying potential therapeutic targets, and optimizing treatment strategies.[60]
Biomarker discovery Analyze large-scale omics data, such as genomics, proteomics, and metabolomics, to identify potential biomarkers for disease diagnosis, prognosis, and treatment response. This can aid in personalized medicine and drug development.[61]
Biomedical images Analyze biomedical images, such as images from microscopy, to identify potential targets for drug development. This can provide valuable information about the effects of potential drugs on cells and tissues.[62]
Cancer immunotherapy Development of cancer immunotherapies, such as immune checkpoint inhibitors and CAR-T cell therapies. Machine learning models can predict patient responses, identify potential biomarkers, and optimize treatment protocols.[63]
Cell therapy Developing cell therapies involves introducing new cells into a patient to treat a disease. Machine learning models can help optimize the production of therapeutic cells and ir effects on patients.[64]
Clinical trial design optimization Optimize clinical trial design by considering patient characteristics, treatment protocols, and trial outcomes. This can help improve the efficiency and success rate of clinical trials.[65]
Clinical trials Help improve the efficiency of clinical trials, which are critical in drug development. This can involve predicting the outcomes of trials, optimizing the design of trials, or identifying suitable patients for trials.[66]
Clinical trials Optimize the planning and execution of clinical trials, the research that examines the security and efficiency of novel drugs in people. The best trial candidates, trial results, and trial data analysis can all be assisted by machine learning algorithms.[67]
Combination optimization Optimize the selection and dosing of drug combinations to maximize efficacy and minimize side effects. Machine learning models can analyze data on drug interactions, synergistic effects, and patient characteristics to identify the most effective combinations.[68]
Combination sensitivity prediction Sensitivity of cancer cells to specific drug combinations, aiding in identifying effective treatment regimens for personalized cancer therapy.[69]
Combination side effect prediction Side effects of drug combinations by analyzing the known side effect profiles of individual drugs and their interactions. This can help identify potential safety concerns and optimize drug combination therapies.[70]
Continuous manufacturing Optimize continuous manufacturing processes in the pharmaceutical industry by monitoring and controlling key process parameters in real time. This can enhance quality control, reduce waste, and increase efficiency.[71]
De novo design Creating new and effective drug molecules from scratch is a complex task given the many possible molecules. Help with this by generating and optimizing potential drug molecules for specific targets or diseases.[72]

(*Continued*)

Table 3 (Continued).

Delivery optimization Optimize drug delivery systems by predicting drug release profiles, designing targeted delivery vehicles, and enhancing efficiency.[73]
Delivery systems Design of drug delivery systems, which control the release and distribution of drugs in the body. Machine learning algorithms can the effectiveness of different drug delivery systems, leading to better treatments with fewer side effects.[74]
Dermatology Analyze skin images to diagnose skin diseases and predict responses to treatment. This could aid in the development of drugs for dermatological conditions.[75]
Disease mechanisms Understand the mechanisms of diseases, aiding in the development of new treatments. Analyzing large amounts of biological and medical data reveals novel insights into how diseases develop and progress.[76]
Disposition modeling Predict drug disposition, including absorption, distribution, metabolism, and excretion (ADME), to understand drug pharmacokinetics and optimize dosing regimens.[77]
Drug-drug interaction prediction Help discover potential drug-drug interactions and their implications on efficacy and safety, and predict probable drug interactions.[78]
Environmental Impact of Pharmaceuticals Help assess the environmental impact of pharmaceuticals, a factor that is increasingly considered in drug development. Machine learning models can reduce drugs' environmental persistence, bioaccumulation, and toxicity, leading to safer and more sustainable treatments.[79]
Epigenetic discovery Analyze epigenetic data and identify potential epigenetic drug targets that modify gene expression patterns. This can aid in developing therapies for diseases influenced by epigenetic modifications.[80]
Formulation and Dosage Optimization Assist in optimizing drug formulation and dosage to enhance drug efficacy, minimize side effects, and improve patient compliance. Machine learning models can optimize formulation and dosage based on patient characteristics and drug properties.[81]
Formulation prediction Based on molecular descriptors and physicochemical characteristics, predict drug formulation properties, such as solubility and stability. This can aid in the development of optimized drug formulations.[82]
Gene editing Assist in gene editing technologies such as CRISPR-cas9 by predicting off-target effects, guiding the design of gene editing tools, and optimizing editing efficiency.[83]
Gene therapy Aid in developing gene therapies involves altering the genes within a patient's cells. Machine learning models can find these alterations' effects and help design more effective therapies.[84]
Genome-wide Association Studies (GWAS) Help analyze the results of genome-wide association studies, which involve scanning the genomes of many people to find genetic variations associated with a particular disease. This can uncover new drug targets.[85]
High-throughput Screening Improve high-throughput screening, a common method for drug discovery, by predicting the properties and potential therapeutic uses of numerous compounds quickly and accurately. Machine learning models can process and learn from vast amounts of chemical and biological data, making them effective tools for drug screening.[86]
Imaging agents Imaging agents are substances used in medical imaging to highlight certain structures or processes. Design of new imaging agents, which can be used for diagnosis or for tracking the progress of disease or the effect of treatment.[87]

(*Continued*)

Table 3 (Continued).

Immunotherapy Develop new immunotherapies, treatments that use the body's immune system to fight diseases like cancer. Machine learning models can respond to immunotherapy, identify potential targets, and help design more effective treatments.[88]
Manufacturing and Quality Control Optimize drug manufacturing processes, predict potential issues, and improve quality control measures. Machine learning models can analyze production data, detect anomalies, and optimize manufacturing parameters to ensure consistent drug quality.[89]
Metabolism prediction Metabolic fate of drugs in the body, aiding in identifying potential metabolites and facilitating drug development and optimization.[90]
Metabolomics Analyze metabolomic data, which provides information about the metabolites in a biological sample. This can identify biomarkers for disease, understand the mechanism of action of drugs, and identify potential drug targets.[91]
Microbiome analysis The microbiome, the collection of microorganisms living in the human body, can affect the effectiveness and toxicity of drugs. Help analyze the complex data from microbiome studies, potentially leading to new strategies for personalized drug therapy.[92]
Nanomedicine Assist in designing nanoscale drug delivery systems. Machine learning can properties of nanoparticles and their interactions with biological systems, helping to design more effective nano-drugs.[93]
Natural language processing Analyze scientific literature, patents, and other text sources to extract knowledge and insights for drug discovery and development. Natural language processing techniques can help in data mining, knowledge extraction, and text summarization.[94]
Natural product discovery Analyze large-scale data on natural products and their biological activities to identify potential drug candidates. This can expedite the discovery of novel compounds with therapeutic potential.[95]
Neglected diseases Accelerate drug discovery efforts for neglected diseases by predicting potential drug candidates and repurposing existing drugs to treat these conditions.[96]
Neurodegenerative diseases Discovery of new drugs to prevent diseases like Alzheimer's and Parkinson's. Machine learning models can analyze multi-omics data and predict potential therapeutic targets for intervention.[97]
Neuropharmacology Develop new treatments for neurological and psychiatric disorders. For example, machine learning models can analyze brain images and other data to respond to treatment or to identify new drug targets.[98]
Ophthalmology AI has also found applications in the development of drugs for eye diseases. For instance, deep learning has been applied to analyze retinal scans, which can provide insights to understand and treat conditions such as age-related macular degeneration.[99]
Orphan diseases Facilitate drug discovery efforts for orphan diseases by predicting drug-target interactions and repurposing existing drugs for these rare conditions.[100]
Pain management Aid in the discovery of new drugs for pain management by analyzing large-scale genomic, transcriptomic, and proteomic data to identify potential targets and pathways involved in pain signaling.[101]
Patient adherence prediction Predict patient adherence to drug regimens by analyzing patient characteristics, social determinants, and past adherence patterns. This can help identify patients at risk of non-adherence and develop interventions to improve drug adherence.[102]

(*Continued*)

Table 3 (Continued).

Patient stratification Assist in patient stratification, which involves grouping patients based on their predicted response to treatment. Identifying patients who can most benefit is a main goal, facilitating the move towards personalized medicine.[103]
Peptide design Design of peptide drugs, which are a unique class of therapeutic agents. Machine learning algorithms can stabilize potential peptide drugs' toxicity and activity, facilitating their design.[104]
Personalized combination therapy Effectiveness of different drug combinations for individual patients based on their genetic and clinical characteristics. This can help in the development of personalized combination therapy for complex diseases.[105]
Personalized dosing Optimize drug dosing for individual patients by considering age, weight, genetics, and biomarker data. This can help achieve optimal therapeutic outcomes while minimizing adverse effects.[106]
Personalized medicine By analyzing genetic data, predict individual drug responses and help develop personalized treatment plans. This is particularly relevant in oncology, where genetic variations can significantly affect treatment outcomes.[107]
Pharmacoepidemiology Be used in pharmacoepidemiology, the study of the uses and effects of drugs in large numbers of people. For example, machine learning models can analyze large health record databases to identify drug use patterns, drug effectiveness in real-world conditions, and the factors influencing these outcomes.[108]
Pharmacogenomics Analyze genetic variations and drug response data to predict individual drug responses and identify potential adverse reactions, enabling personalized medicine approaches.[109]
Pharmacokinetic modeling Enhance pharmacokinetic modeling, which involves studying how drugs are absorbed, distributed, metabolized, and excreted by the body. Machine learning models can predict drug concentrations in different tissues and optimize dosing regimens.[110]
Pharmacokinetics and Pharmacodynamics (PK/PD) Modeling Help develop PK/PD models, which predict how the body will affect the drug (pharmacokinetics) and how the drug will affect the body (pharmacodynamics). Optimal dosing and frequency are one main element of understanding for individual patients.[111]
Pharmacovigilance signal detection Analyze large-scale pharmacovigilance databases and real-world data to detect potential safety signals and adverse drug reactions. This can enhance the early detection and monitoring of drug safety concerns.[112]
Interaction of drugs with multiple targets, a field known as polypharmacology. Understanding these interactions can help design drugs to achieve the desired effects and minimize undesired side effects.[113]
Precision medicine Enable precision medicine by integrating diverse patient data, including genomic information, clinical records, and lifestyle factors. Machine learning models can analyze these data to predict individual treatment responses and tailor therapies for optimal outcomes.[114] Best drug combinations for individual cancer patients based on their genetic and clinical data, helping to move toward personalized and precision medicine in oncology.[115]
Preclinical safety assessment Aid in the preclinical safety assessment of drugs by predicting their potential toxicities and identifying safety risks, leading to more efficient and reliable safety evaluation.[116]
Predicting interactions Predict potential drug-drug interactions, which are situations where one drug affects the activity of another. This is critical for ensuring patient safety, as drug-drug interactions can lead to adverse effects or reduced treatment efficacy.[117]

(*Continued*)

Table 3 (Continued).

Predicting side effects Potential side effects of drugs contribute to better patient safety and care. Machine learning models can utilize chemical properties, known side effects, and biological data to make these predictions.[118]
Predictive toxicology Toxicity of chemical compounds and drugs, aiding in the early identification of potential safety issues and reducing the need for animal testing.
Pricing and Market Access Analyze healthcare data, market dynamics, and pricing information to optimize drug pricing strategies and improve market access for pharmaceutical companies. This can help ensure the affordability and availability of essential drugs.
Proteomics Analyze proteomic data, which provides information about the proteins in a biological sample. This can help to understand the mechanism of action of drugs, identify potential drug targets, and identify biomarkers for disease.[119]
Quantitative structure-activity relationship (QSAR) models Help develop QSAR models, which show the biological activity of compounds based on their chemical structure. This can accelerate the identification of potential new drugs.[120]
Rare disease discovery Aid in discovering drugs for rare diseases by analyzing various data sources, including genetic and clinical data, to identify potential therapeutic targets and repurpose existing drugs.[121]
Regulatory Compliance and Safety Surveillance Aid in regulatory compliance by automating the analysis of safety data, detecting adverse events, and monitoring post-market drug safety. This can improve pharmacovigilance and regulatory decision-making processes.[122]
Regulatory processes Streamline regulatory processes in drug development, including drug approval and post-marketing surveillance, by automating data analysis, identifying safety signals, and facilitating regulatory decision-making.[123]
Repositioning for Rare Diseases Identify potential drug candidates for treating rare diseases by analyzing molecular profiles, gene expression data, and clinical information.[124]
Repositioning, or drug repositioning or repurposing, is finding new uses for already-existing drugs. Predict new therapeutic uses for already-on-The-market drugs by studying data on drug structures, actions, and targets. Repositioning Help identify new uses for existing drugs, a process known as drug repositioning or repurposing. By analyzing data on drug structures, effects, and targets, predict new therapeutic uses for drugs already on the market.[125]
Reproduction Reproduce discontinued or scarce drugs by analyzing their chemical structures and properties. This can help ensure the continuous availability of essential drugs.[126]
Repurposing drugs Find new uses for existing drugs, a process known as drug repurposing or repositioning. This can be a quicker and less costly way of developing new treatments than traditional drug development processes.[127]
Repurposing for Rare Diseases Identify potential drug candidates for treating rare diseases by analyzing drug databases, genomic data, and clinical information. This can expedite the development of therapies for rare conditions.[128]
Resistance prediction Predict drug resistance in pathogens and cancer cells, aiding in developing strategies to overcome resistance and optimize treatment.[129]
Response Prediction for Personalized Medicine Predict individual drug responses based on patient-specific factors, such as genetics, clinical data, and environmental factors. This can enable personalized treatment selection and optimization.[130]

(*Continued*)

Table 3 (Continued).

Safety and Adverse Event Prediction Predict and prevent adverse drug events by analyzing data from electronic health records, clinical notes, and other sources. Machine learning models can identify patterns and risk factors associated with adverse events, enabling early detection and prevention.[131]
Structure-based design Ai/mL models can be utilized in structure-based drug design, which involves drug design based on the drug target's molecular structure. These models can predict how well potential drugs will bind to the target, helping to identify promising drug candidates.[132]
Supply chain optimization Optimize the supply chain of pharmaceutical products, including inventory management, demand forecasting, and distribution logistics. This can help improve efficiency, reduce costs, and ensure timely availability of drugs.[133]
Synergistic combinations Identify combinations of drugs that work synergistically; their combined effect is greater than the sum of their individual effects. This could lead to the development of more effective treatments, particularly for complex diseases like cancer.[70]
Synthesis planning Assist in planning the synthesis of new drugs by proposing the most effective and feasible routes and steps. This can save time and resources in drug development.[134]
Synthetic Biology and Biotechnology Optimize the production of bioactive compounds through synthetic biology and biotechnology. For example, machine learning models can optimize the genetic engineering of microorganisms to produce drugs.[135]
Target binding affinity prediction Small molecule target protein binding affinity facilitates the invention and improvement of therapeutic candidates.[136]
Target druggability prediction Druggability of potential drug targets, helping to prioritize targets for further investigation and optimizing drug discovery efforts.[137]
Target identification Identifying potential drug targets by analyzing genomic, proteomic, and other biological data. Machine learning models can uncover associations between targets and diseases, facilitating the discovery of new therapeutic targets.[138]
Target validation Also assist in target validation, which involves confirming that a biological target (such as a protein or gene) is related to disease and can be acted upon by a drug. Machine learning models can analyze and interpret complex biological data to uncover and validate new drug targets.[139]
Toxicology predictions Potential drug toxicity is a critical component in drug development. Early detection of adverse effects can reduce the likelihood of expensive failures during the latter phases of drug development.[140]
Vaccine design Machine learning models can provide immune responses to different vaccine candidates, which can help design more effective vaccines.[141]
Virtual screening Accelerate the virtual screening process to identify potential drug candidates from large compound libraries. Machine learning models can increase the likelihood of a compound binding to a specific target, aiding drug discovery. [142]
Withdrawal prediction Likelihood and severity of withdrawal symptoms when discontinuing certain drugs. This can help healthcare providers develop tapering strategies and support patients in safely discontinuing drugs.[143]

Drug Target Identification Technologies Before AI/ML

- Genomics and Proteomics: Genomics and proteomics helped identify disease-associated genes and proteins. By understanding which genes were involved in a particular disease process, scientists could target the proteins those genes encoded.[144]

- Biochemical Assays: These were used to identify and understand the function of potential targets. Scientists would isolate a particular protein or enzyme involved in a disease process and test compounds that might inhibit or enhance its activity.[145]
- Animal Models: By studying disease progression in animal models, researchers could identify the proteins and pathways involved in a disease, giving insight into potential targets.[146]
- Cell-Based Assays: Using cultured diseased cells, researchers would test how different compounds affected the cells. To identify potential targets, they would look at how the compounds impacted cell growth, death, and other processes.[147]
- High-Throughput Screening (HTS): This involved testing thousands of compounds against a particular protein or enzyme to see which ones had the desired effect. Although not as efficient as modern AI-driven methods, it was a vital tool for drug discovery.[148]
- Structure-Based Drug Design: If the 3D structure of a target protein was known, scientists could design new molecules that fit into the protein's active site. This was an essential method in rational drug design.[149]
- Phenotypic Screening: This approach looked at the effects of a compound on the phenotype of a cell or organism. Instead of focusing on a specific target, it aimed to identify compounds that had a desired effect, like killing cancer cells. The targets were then identified later using various methods.[150]
- Literature and Knowledge Mining: Accumulated knowledge and scientific literature were used to draw connections between disease mechanisms and potential targets. Collaboration and information sharing across the scientific community were vital in this process.[151]
- Patient Data Analysis: Clinical observations and epidemiological studies helped identify correlations between molecules or pathways and diseases, leading to hypotheses about potential targets.[152]
- Collaboration with Academia: Many times, insights into potential drug targets came from academic research into the fundamental biology of a disease.[153]

Drug Target Identification in the AI/ML Era

The advent of artificial intelligence (AI) and machine learning (ML) has revolutionized the field of drug target identification by enhancing predictive modeling, expediting the screening process, integrating vast and complex data sets, and fostering personalized medicine approaches. Integrating AI/ML in drug target identification has undeniably transformed the field, providing new avenues and methodologies for accelerating drug development.

- Enhanced Predictive Modeling: AI/ML algorithms can learn from existing data to predict how potential drug targets might behave, even before being experimentally tested.[154]
- Speeding Up Screening Processes: AI has dramatically enhanced high-throughput screening, which can analyze vast numbers of compounds more quickly and accurately.[155]
- Data Integration and Mining: AI/ML has enabled the integration of various data sources like genomics, proteomics, and clinical data, resulting in a more comprehensive understanding of disease biology and potential targets.[156]
- Improving Structure-Based Drug Design: AI/ML techniques can analyze the 3D structures of proteins and predict how different molecules might interact with them, enhancing drug design.[157]
- Personalized Medicine Approaches: AI/ML can analyze patient-specific information to tailor drug therapies, allowing for more personalized and effective treatments.[158]
- Collaborative Drug Discovery: AI-driven platforms enable collaborations across different institutions, combining various expertise and data, leading to more innovative drug target identification.[159]
- Fostering Drug Repurposing: AI can identify new potential targets for existing drugs, a process known as drug repurposing or repositioning, thus maximizing the therapeutic potential of already approved drugs.[160]
- Cost-Effectiveness: By enhancing the efficiency and accuracy of drug target identification, AI/ML reduces the overall cost of the drug discovery process.[161]
- Network Analysis for Disease Understanding: AI/ML algorithms can analyze complex biological networks, revealing underlying mechanisms and potential drug targets.[162]

- Enhanced Safety Profiling: AI/ML models are used to predict the toxicological properties of compounds, contributing to the selection of safer drug targets.[132]
- Improvement in Clinical Trials: AI and ML facilitate the design and monitoring of clinical trials, making predicting patient responses to specific drugs possible.[163]
- Integration with Real-World Evidence (RWE): The use of AI to analyze real-world data from electronic health records and other sources enhances the identification of drug targets.[164]
- Identification of Biomarkers: AI/ML techniques help in the title of disease-related biomarkers, which can be crucial for targeted therapies.[165]
- Facilitating Drug Combination Strategies: AI helps design drug combinations that can be more effective in treating complex diseases.[166]
- Understanding Drug Resistance: AI/ML models help understand the mechanisms of drug resistance, aiding in designing more effective treatments.[167]

Applications in the Development of Small Molecules

AI and ML have been applied extensively to developing and identifying small molecule targets. The convergence of AI and ML with chemistry, biology, and pharmacology fosters an era of unprecedented innovation in developing and identifying small molecule targets. Through intricate modeling, prediction, and optimization, these techniques push the boundaries of traditional approaches, paving the way for safer, more effective, and more personalized therapies.

In summary, the application of AI/ML in identifying and developing small molecule targets spans various stages and aspects of the drug discovery process. By integrating information on molecular structure, biological function, pharmacokinetics, and synthesis, these computational tools enable more efficient and targeted development of small molecule therapeutics.

These applications can be broken down into specific areas:

- Virtual Screening: AI/ML models can screen millions of small molecules, predicting their binding affinity to specific biological targets, a critical factor in drug development.[168]
- Structure-Based Drug Design: By understanding the 3D structures of proteins, AI/ML algorithms facilitate the design of small molecules that can effectively interact with those targets.[157]
- Prediction of ADMET Properties: AI/ML can predict the Absorption, Distribution, Metabolism, Excretion, and Toxicity (ADMET) of small molecules, which are critical in drug development.[169]
- Biomarker Identification: AI/ML assists in identifying biomarkers that small molecules might target, providing insights into potential therapeutic applications.[165]
- Chemoinformatics and Compound Property Prediction: AI/ML models predict various physicochemical properties of small molecules, such as solubility and stability.[170]
- Drug Repurposing: Identifying new targets for existing small molecules is another area where AI/ML shines. This approach can significantly reduce the time and cost associated with drug development.[160]
- Synthesis Prediction: AI/ML can predict the synthetic routes to small molecules, a crucial step in drug development.[134]
- Understanding Drug Resistance: AI/ML models help in understanding mechanisms of drug resistance to small molecules, aiding in the design of more effective treatments.[167]
- Multi-Target Prediction: AI/ML models can predict the interaction of small molecules with multiple targets, which is essential for understanding off-target effects and polypharmacology.[171]
- Designing Small Molecule Libraries: AI/ML helps design compound libraries with desirable properties, accelerating drug discovery.[172]
- Molecular Dynamics Simulations: AI/ML algorithms can run and analyze molecular dynamics simulations, providing insights into the behavior of small molecules in biological systems.[173]
- Patient-Specific Treatment: AI/ML helps identify small molecules that could be particularly effective in individual patients, contributing to personalized medicine.[158]

- Quantum Mechanics (QM) and QM/MM Calculations: AI/ML has been employed to facilitate and interpret QM and QM/MM calculations, offering detailed insights into molecular interactions at the atomic level.[174]
- Toxicity Prediction: AI/ML models can accurately predict the toxicity of small molecules, an essential aspect of drug safety.[140]
- Enhancing Collaboration: AI/ML tools enhance collaboration between chemists, biologists, and pharmacologists, allowing for a more integrated approach to small molecule discovery.[114]
- Optimizing Clinical Trials for Small Molecules: AI/ML models can help design and monitor clinical trials specifically for small molecules, reducing the time and cost of development.[175]
- Developing New Therapeutic Strategies: AI/ML can be used to understand complex disease mechanisms, leading to the development of novel therapeutic strategies involving small molecules.[176]

Applications in the Development of Large Molecules

Identification, Selection, and Prioritization of Drug Targets

A crucial step in the early stages of drug development is identifying an appropriate biological target for therapeutic prospects. To begin with, AI/ML can be used to evaluate and synthesize large amounts of data from recent scientific papers, research studies, and other data sources to identify biological targets and elucidate disease correlations. By enhancing transcriptomic, proteomic, and other data sources from healthy individuals and those with a particular condition of interest, a significant opportunity to influence biological target selection is given by the datasets that usually involve complexity and originate from a variety of sources them excellent for use with AI/ML approaches.[177] To provide reliable, potential structure and function of biological targets and predict their involvement in a disease pathway, AI/ML can help mine and assess these huge multi-omics and other datasets, drawing on existing validated data.[114,178] Although the early target discovery and prioritization stage is vital where AI/ML could improve the efficacy and efficiency of drug development, it is necessary to confirm the biological target's involvement in the pertinent disease through further studies.[177] Only when AI/ML tools are combined with such applications can be made quickly evolving databases can such applications be made.

Compound Screening and Design

To uncover possible therapeutic candidates that affect the function of the indicated biological targets of interest, substantial in silico or experimental screening of chemical libraries is required to improve the specificity and selectivity of a chemical for the biological target. Potential AI/ML uses for compound screening include predicting compounds' chemical properties, bioactivity, effectiveness, and land effects based on their specificity and affinity for a target. [179,180]

To forecast the toxicity of a chemical based on its recognized properties, AI/ML can detect drug-target interactions and provide predictions about classes of drugs that may interact with the same targets or have a similar mechanism of action. By leveraging previously described molecules, this technique further directs drug repurposing. The greater accessibility to pertinent real-world data (RWD) from many sources, such as electronic health records (EHRs, registries, and DHTs), is highly beneficial for drug repurposing initiatives using AI/ML to find previously unrecognized impacts of drugs on disease pathways.[181]

Finally, de novo drug design is also advancing because AI/ML may be used,[182] for example, to assist in predicting the 3D structure of target proteins to direct chemical production and the total impact of a drug candidate on the target, including predicting affinity and probable toxicity.[114,179,183] When using AI/ML to predict potential structures, one must employ caution because many proteins made for pharmaceutical purposes are codon-optimized (with several synonymous mutations included). The effects of these codon optimizations on protein structure are still poorly understood.

Structure Prediction

The first protein structure prediction algorithm was reported in the late 1960s, yet protein structure prediction has long remained a paradox for longer. Levinthal's dilemma was put forth in a seminal study by Cyrus Levinthal in 1969 titled "How to Fold Graciously" in Science.[184] The paradox demonstrated the enormous range of potential conformations a protein might adopt, indicating that it would be unrealistic to sample all potential conformations and determine the native structure of the protein by random search. Even though Levinthal did not suggest a precise algorithm to predict protein structures, his study generated considerable interest and served as a springboard for further investigation.

Numerous strategies and methods have been created to address the protein folding issue and forecast protein structures. These include approaches using machine learning techniques, potentials based on knowledge, and simulations based on physics. The creation of the first successful protein folding algorithm, named "DREIDING", by Richard Corey and Irving Kuntz in 1974, represents a significant turning point in the history of protein structure prediction algorithms. DREIDING predicted the folding of tiny proteins using a distance geometry technique.[185]

The following steps were homology modeling,[186] ad initio folding,[187] energy minimization, molecular dynamics simulations,[188,189] fragment assembly,[190,191] machine learning, and deep learning. Since then, many methods have emerged to investigate the conformational space and find the natural state of proteins that is energetically advantageous.

However, the AlphaFold 2 (AF2) report,[183] which won the 14th Critical Assessment of Protein Structure Prediction (CASP14) with predictability, sparked a lot of interest and gave rise to the idea that we might now be able to tackle problems with protein structure folding that go beyond simple structure prediction from a sequence. The performance of 3D-based predictors has been consistently superior to 1D-based predictors;[192–194] therefore, a pool of excellent 3D predicted structures could mark a significant shift in applications for understanding protein structure. The consequences of mutations on protein stability and several protein interactions, such as protein-protein, protein-ligand, and protein-DNA/RNA, were also expected to be predicted by these algorithms. As a result, developers have continued to discover applications of predicted 3D structures,[195] enabling assessment of the stability and function of a mutation associated with many diseases.[196] After all, if there is a reported confidence score in the predicted structure, it must somehow be related to the structure; it was assumed, despite the disclaimer by AF2, that it "has not been validated for predicting the effect of mutations".[197] While AI/ML was anticipated to yield a paradigm shift to overcome the Levinthal paradox, it seems that the paradox stays. The current algorithms can regurgitate on reported structures, with added errors, nothing more.

Comparison of Modeling Methods Between Small and Large Molecules

The comparison showcases specialized techniques and methodologies applied to small and large molecules. It underscores the complexity of extensive molecule modeling, often requiring multi-scale approaches and integration with other biological disciplines. The cited literature provides the basis for these insights and differences.

Table 4 lists a comparative analysis of the most common modeling approaches used in small and large molecule discovery. The table summarizes the similarities and differences between the methods and challenges in small and large-molecule drug discovery. It also includes references to specific citations for each aspect.

Table 4 Comparison of applications of modeling methods in small molecules and large molecules

Aspect	Small Molecules	Large Molecules
Ligand Efficiency	Focuses on binding modes and ligand efficiency[198]	Specialized modeling for complex structures like protein-protein interfaces.[199]
Scalability & Complexity	Less complex, allowing for exhaustive computational methods[200]	Requires specialized techniques due to higher complexity[201]
Fragment-Based Drug Design	Focus on stepwise assembly of small fragments[202]	It needs to be integrated with other methods for large targets[203]
Computational Requirements	Generally requires fewer computational resources[204]	Requires substantial computational power and specialized software[205]
Flexibility Analysis	Analysis of small molecule adaptability within binding pockets[206]	Focus on conformational changes in large biological contexts[207]
High-Throughput Screening	A common method for quickly assessing the activity of small molecules[208]	More complex due to larger structural variability[148]

(*Continued*)

Table 4 (Continued).

Aspect	Small Molecules	Large Molecules
Machine Learning & AI	Used for predictive modeling, virtual screening, and QSAR[114]	Requires more sophisticated algorithms for large molecule interactions[157]
Structure-Activity Relationship	The key to optimizing small molecule leads[209]	More intricate analysis focusing on domains or epitopes[210]
Molecular Dynamics Simulations	Studies binding mechanisms and predicts free energy landscapes[211]	Crucial for understanding folding, conformational changes, and large-scale motions[212]
Experimental Integration	Complemented with in vitro assays to validate predictions[213]	Requires integration with high-resolution methods like cryo-EM or X-ray crystallography[214]
Quantum Mechanics (QM) Methods[215]	Used for detailed electronic structure analysis	Often integrated with Molecular Mechanics (MM) for large systems (QM/MM)
Pharmacophore Modeling[216]	Identification of essential features for biological activity	Used for mapping interaction sites within large complex structures
Cheminformatics[217]	Extensive use for managing chemical information, similarity searching, clustering	Used for large compound libraries, handling biological sequences
Virtual Screening[218]	Utilized for identifying potential hits from chemical libraries	Applied for identifying lead candidates in large molecular libraries
Bioinformatics Integration[219]	Interconnected with genomics and proteomics for target identification	Essential for understanding large molecule interaction networks
Free Energy Calculations[220]	Used for predicting binding affinities, often with simplified models	Requires advanced methodologies due to the complexity of large molecules
Antibody Modeling[221]	Less relevant for small molecules	Specialized methods for modeling antibody-antigen interactions
Docking & Simulation[149]	Essential in predicting how small molecules bind to their targets	Highly complex and requires refined algorithms for large-molecule docking
Multi-Scale Modeling[189]	Used in context with multiple scales (eg, QM/MM)	Essential to capture different levels of structural detail in large molecules

Nonclinical Investigation

Nonclinical research is intended to advance novel drugs for further human clinical study. It includes in vitro and in vivo experiments. Nonclinical studies that promote the development of new drugs can be carried out at any stage of the development process, including before clinical trials, concurrently with those trials, and even in post-marketing settings. Data from animal-based pharmacokinetic, pharmacodynamic, and toxicologic studies, exploratory in vitro and in vivo mechanistic studies, organ-on-chip and multi-organ chip systems, and cell assay platforms may be used to evaluate toxicity, explore mechanical models, and develop in vivo predictive models.[222–226]

Pharmacokinetics (PK) studies show drugs are absorbed, distributed, metabolized, and excreted over time. Pharmacodynamics (PD) investigates how drugs affect the body biologically. A model incorporating PK and PD can forecast how a drug's effects will change depending on dosage or dosing regimen. Drug development has long employed pharmacokinetic/pharmacodynamic (PK/PD) modeling, which can be used in nonclinical and clinical stages.[227] Along with improvements in computational tools and technology and the availability of modeling platforms, the use of physiologically-based pharmacokinetic (PBPK) and physiologically-based PK/PD (PBPK-PD) modeling is growing.[228] Other applications use more cutting-edge AI/ML techniques for PK/PD modeling, including tree-based and artificial neural network models. For instance, in excessively complex PK/PD data analysis, a recurrent neural network, a machine

learning (ML) approach widely used for time series data analysis, may be used with traditional PK/PD models, perhaps increasing accuracy for nonclinical and clinical applications.[229]

In addition to the FDA's studies on AI/ML, the Agency recently released a pivotal guideline. "Generally Accepted Scientific Knowledge (GASK) in Applications for Drug and Biological Products: Nonclinical Information",[230] In two scenarios, sponsors can rely on GASK to satisfy applicable approval requirements: (1) One can depend on GASK on a substance's known effects on bi when a product contains one (whether synthetically created or organically derived). For instance, a drug may occasionally have either on- or off-target effects on a biological pathway or molecular mechanism of action known to have harmful effects at clinically significant quantities depending on how the biological system functions. Therefore, it may be preferable to rely on GASK regarding the influence of the pathway rather than doing specific pharmacological and toxicological tests to determine the impact of the path. Therapeutic proteins bind to receptors, triggering a pharmacodynamic response that results in the pharmacological response, the basis of which is the clinical response. The FDA's progressive thinking has resulted in more logical approaches to drug development. However, optimal use of the GASK will need AI/ML applications since finding what is "scientific" and what is "knowledge" requires extensive data analysis.

Clinical Studies

The FDA encourages cutting-edge clinical trial designs based on RWD analytics and digital health data (DHT). The FDA has noticed a sharp increase in regulatory filings mentioning the. Use of AI/ML in recent years. In 2021,[231] there will be more than 100 submissions across all pharmacological and biological product categories, including those involving AI/ML. These submissions span various stages of the drug development process and include the uses of AI/ML in a range of therapeutic areas, from the creation of new drugs and the enhancement of clinical trials through the assessment of endpoints and post-market safety monitoring.

In summary, integrating AI and ML into clinical trials is a multifaceted approach that continues to revolutionize various aspects of clinical research. From patient recruitment to real-time adaptation, predictive modeling, and ensuring ethical conduct, these technologies offer a comprehensive set of tools to make the development of new medical interventions faster, more efficient, and patient-centric. They provide the opportunity to bridge the gap between traditional methods and modern needs, creating a more streamlined and responsive approach to medical research and healthcare innovation.

Clinical development in research phases continues to be the significant and most frequent application of AI/ML. Clinical research often entails a progression of clinical trial phases involving increasing numbers of human volunteers to evaluate a drug's efficacy and safety, with significant help from AI/ML. For instance, AI/ML assists in analyzing substantial volumes of data from both interventional and non-interventional trials to conclude the efficacy and safety of a treatment. AI/ML may also impact the layout and efficiency of non-traditional trials, such as decentralized clinical trials and studies using RWD obtained from electronic health records (EHRs), medical claims, or other data sources. The analysis and interpretation of data obtained from DHTs used in clinical research are also aided by AI/ML. Finally, AI/ML enhances the efficiency of trials.

Trial Design

Natural Language Processing (NLP) techniques assist in rapidly mining information from medical literature, clinical notes, and other textual data, thus expediting the design and conceptualization of trials.[232] Predictive analytics using AI models can predict patient responses to treatment. This helps in designing more effective and personalized treatment strategies.[233] AI also enables adaptive trial designs where ongoing results may influence the course of the trial, such as modifying dosages or allocating more patients to more promising treatment arms.

AI/ML gathers data from previous clinical trials to assist in designing new clinical trials through a combination of techniques, including data mining, predictive modeling, natural language processing (NLP), and machine learning algorithms. These examples illustrate the multifaceted ways in which AI/ML can leverage data from previous clinical trials to support the design of new ones.

- Data Mining from Existing Repositories and Databases: AI can mine data from previous clinical trials stored in various databases, such as ClinicalTrials.gov, to analyze trends, outcomes, and patient characteristics.[234]
- Natural Language Processing for Information Extraction: NLP can extract relevant information from unstructured clinical documents, publications, and trial protocols to inform the design of new trials.[235]
- Predictive Modeling for Identifying Success Factors: Machine learning models can analyze previous clinical trial data to identify factors contributing to success or failure, assisting in designing new, optimized trials.[236]
- Integrating Real-World Evidence (RWE): AI can integrate and analyze real-world and clinical trial data to comprehensively understand patient populations and treatment effects.[164]
- Adaptive Trial Design Using Reinforcement Learning: Reinforcement learning algorithms can dynamically adapt trial designs based on interim results and other data, potentially improving efficiency.[237]
- Utilizing Public Databases for Biomarker Discovery: Machine learning models can analyze public genomic databases to discover new biomarkers that can be targeted in clinical trials.[238]

Recruitment

AI/ML helps match people with clinical trials for experimental treatments where participants may benefit from these trials.[239] AI/ML mines vast data to compare patients to trials, including clinical trial databases, trial announcements, social media, medical literature, registries, and structured and unstructured data in EHRs. Even though these algorithms are trained on massive amounts of patient data and enrollment criteria from prior trials, it is crucial to ensure adequate representation of populations that are likely to use the drug (eg, gender, race, and ethnicity) as matching algorithms are created and, when used, to confirm that equitable inclusion was achieved during the recruitment process. If properly vetted, these technologies might play a more significant part in connecting people with experimental treatments in the future.

AI algorithms can analyze vast amounts of patient data to identify the best candidates for clinical trials more quickly and accurately. This helps to reduce the time spent on the recruitment phase.[240]

Trial Participants' Selection and Stratification

In clinical studies intended to show the efficacy of drugs, enrichment strategies can help with participant selection based on baseline information, eg, demographic data, clinical data, vital signs, labs, medical imaging data, and genomic data.[239,241–244] AI/ML has made predicting a participant's clinical outcome possible. These prediction models can improve clinical trials by identifying participants at high risk or more likely to respond to the treatment. Using these AI/ML algorithms for patient screening and selection before randomization may reduce variability and increase study power.[245]

For example, if an AI/ML model could predict the likelihood of a severe adverse event before administering an investigational treatment, participants could be divided into different groups and subsequently monitored (or excluded depending on the predicted severity of the adverse event). Such predictive models can also be used for participant stratification and enrichment strategies.

AI can match patients with suitable clinical trials by analyzing their medical records and conditions. This not only expedites recruitment but ensures the most relevant candidates are chosen.[246] AI can be used to ensure that clinical trials are conducted ethically and that the patient population is diverse and representative, thus improving the fairness and inclusivity of trials.[247]

AI-powered sentiment analysis can understand patient feedback, enhance patient experience, and ensure patient voices are considered during clinical trial phases.[248]

Optimization of the Dose/Dosing Regimen

After administering a drug, PK profiles can be described and predicted using AI/ML. It can also examine confounding factors when examining the relationship between drug exposure and response.[229,249] The ideal dose and dosing regimen for a study can be chosen using these models. This can involve assisting with dose optimization in particular populations

where data may be scarce, such as those with rare diseases, children, and pregnant women. AI/ML can be used to design optimal trial protocols, reducing potential errors, inconsistencies, and inefficiencies in the clinical trial process.[250]

Adherence

AI/ML can be used to monitor and improve adherence during a clinical study with tools like smartphone notifications and reminders, eTracking of drugs (eg, intelligent pillboxes and tools for visual confirmation),[251] and eTracking of missing clinical sessions, which generates non-adherence alerts. Examples of how AI/ML is used in clinical research to improve drug adherence include applications that use digital biomarkers, such as facial and vocal expressivity, to monitor drug adherence remotely. Machine learning algorithms can monitor patient data in real time, allowing early detection of adverse effects and ensuring patient safety. This real-time monitoring helps in quickly adapting the trial processes.[252] AI-driven wearables and sensors can enable continuous remote monitoring of patient health. This data feeds into real-time algorithms that can adapt trials, ensuring safety and efficacy.[253,254]

Retention

AI/ML can improve participants' access to important trial information by enabling features like AI chatbots, voice help, and intelligent search. AI/ML can minimize the strain on participants by adopting passive data collection techniques and extracting more information from the data already collected during clinical practice or study activities.[178] Data from DHTs and other systems can be used to construct patient profiles that can be used to predict participant retention, adverse events, and dropout rates.

Site Choice

Operational trial conduct could be enhanced by utilizing AI/ML to identify which sites have the best chances of a successful trial and to help sites identify process deficiencies. For instance, algorithms can evaluate site performance and help identify which sites may be more likely to run behind schedule using data from prior trials at a site. AI/ML can significantly reduce costs and time by automating various aspects of the clinical trial process. This efficiency may enable more trials to be run simultaneously.[158]

Data Gathering

Clinical trials increasingly utilize DHTs, including wireless and smartphone-connected gadgets, wearables, implantable, and ingestible, to gather objective, quantitative, longitudinal, and continuous physiological data.[255] Furthermore, a lot of these DHTs support the use of AI/ML, either as embedded algorithms within the DHT or applied to the data generated after the DHT's data are collected, and have been used to predict the status of a chronic disease and its prognosis for treatment[256] or to identify novel traits of an underlying condition. The vast and varied data created from continuously watching patients using these technologies can only be analyzed using AI/ML. This might involve utilizing AI/ML to evaluate multimodal data and composite measures incorporating distinct measurements gathered through various DHTs.[257]

Deep learning algorithms can analyze medical imaging data such as MRI and CT scans more accurately and quickly than traditional methods, providing critical insights into patient response to treatments.[258]

Data Management

AI/ML can perform several data cleaning and curation activities, including locating duplicate participants, imputing values for missing data,[259] and standardizing regulated terminology throughout drug development projects. AI/ML could significantly enhance data integration efforts by using supervised and unsupervised learning to support integrating data submitted in various forms and carrying out data quality assessments. AI/ML can also be utilized for data curation by creating metadata, masking and de-identifying personally identifiable information, and searching and retrieving stored data. These programs might increase the accuracy of the data and the efficiency with which the data are processed for analysis.

Analysis of Data

AI/ML has analyzed large volumes of complex and variable RWD from various sources, including sickness registries, medical claims, and EHRs. Predictive modeling and counterfactual simulation utilizing AI/ML are also used to study clinical trial designs. For instance, computer simulation and modeling are used in silico clinical trials to evaluate drug candidates using a virtual cohort of simulated people with traits representing the desired participant population.[260] In these circumstances, AI/ML could help estimate counterfactual simulations and forecast trial results before human trials.

Digital twins of patients, a new technique that may be employed in clinical research, are another way that AI/ML can be applied more individually. To create digital twins of patients, AI/ML can create in silico representations or duplicates of people that can dynamically reflect molecular and physiological status across time.[261–263] Compared to a clinical trial participant who received an innovative treatment, the digital twin may provide a thorough, longitudinal, and computationally generated clinical record that details what may have happened to that subject if they had received a placebo.

Automated systems powered by AI can handle the vast paperwork involved in clinical trials, ensuring regulatory compliance and thus speeding up administrative tasks.[264]

Machine learning algorithms can efficiently analyze large and complex clinical datasets. By automating data processing and analysis, AI/ML helps draw quicker insights and aids decision-making.[265]

Evaluation of Clinical Endpoints

The clinical endpoint is critical to evaluating the efficacy and safety of medical treatments in clinical trials. A potential safety signal might be identified using groupings of symptoms and indicators by AI/ML-enabled algorithms, which could also help with the real-time detection of situations posing a safety risk.[266,267] AI/ML could help evaluate clinical trial outcomes gathered from many sources (such as DHTs and social media), especially those with enormous amounts of data for which manual assessment might be impossible.

Algorithms for AI/ML Application in Clinical Trials

- IBM Watson Health: Offers various AI-powered healthcare solutions, including patient engagement, medical imaging, and clinical trial management. https://www.ibm.com/watson-health
- Siemens Healthineers: Provides AI-based medical imaging solutions, including deep learning applications. https://www.siemens-healthineers.com/en-us
- Flatiron Health: Offers platforms that use AI and machine learning for oncology research and patient matching in clinical trials. https://flatiron.com
- Deep 6 AI: Specializes in patient recruitment for clinical trials using AI and natural language processing. https://deep6.a
- Tempus Labs: Utilizes machine learning to provide personalized cancer care, genomic sequencing, and other data-driven healthcare solutions. https://www.tempus.com
- Medidata Solutions: Offers AI-powered solutions for clinical trial design, patient recruitment, and data management. https://www.medidata.com
- Google Health: Works on applying AI to healthcare problems, including predictive modeling, medical imaging, and disease detection. https://health.google
- GE Healthcare: Provides various AI-driven medical imaging and monitoring solutions. https://www.gehealthcare.com
- Bayer's AI-driven Drug Discovery Platform: Collaborates with AI technology companies like Exscientia to accelerate drug discovery using AI algorithms. https://www.bayer.com
- Owkin: Focuses on developing machine learning algorithms for medical research, including predictive modeling and treatment analysis. https://www.owkin.com/
- Philips: Offers AI-driven solutions for diagnostics, patient monitoring, and healthcare informatics. https://www.usa.philips.com/healthcare/solutions/ai-in-healthcare
- NVIDIA Clara™: Provides AI-powered medical imaging and genomics capabilities, supporting the development of AI algorithms in diagnostics and treatment planning. https://developer.nvidia.com/clara
- Quartic.ai: Offers AI-powered pharmaceutical manufacturing and research solutions, including drug discovery and development analytics. https://quartic.ai/

- Roche's NAVIFY Decision Support: Uses AI algorithms to enhance clinical decision-making through its portfolio, including tools for tumor board solutions and trial matching. https://diagnostics.roche.com/
- GNS Healthcare: Employs machine learning for drug discovery and development, helping to create precision medicines and optimize clinical trials. https://www.gnshealthcare.com/
- DataRobot: Provides enterprise AI solutions that can be applied to healthcare for predictive modeling, risk stratification, and other data-driven insights. https://www.datarobot.com/
- BlackThorn Therapeutics: Uses a data-driven approach to neurobehavioral health, leveraging AI and machine learning for patient selection and clinical trial design. https://www.blackthornrx.com/
- CureMetrix: Provides AI-driven mammography solutions, which can be utilized in clinical trials related to breast cancer screening and diagnosis. https://curemetrix.com/
- Bioclinica: Offers specialized technology and services for clinical trials, including medical imaging, patient recruitment, and risk-based monitoring using AI. https://www.bioclinica.com/
- CloudMedx: Utilizes AI to generate real-time insights across the healthcare continuum, including predictive risk models that can be applied to clinical trials. https://www.cloudmedxhealth.com/
- TriNetX: Offers a global health research network that uses AI to enhance clinical trial design and patient recruitment. https://www.trinetx.com/
- Zebra Medical Vision: Specializes in reading medical imaging, providing AI solutions that may be useful for radiological assessments in clinical trials. https://www.zebra-med.com/
- Insilico Medicine: Utilizes AI for drug discovery and aging research, offering solutions for target identification and biomarker development. https://insilico.com/

Post-Marketing Safety Monitoring

In the post-approval period, post-marketing safety reporting of adverse events related to drug utilization is part of post-marketing safety monitoring or pharmacovigilance (PV) activities. The research and practices involved in identifying, evaluating, comprehending, and preventing adverse events or other drug-related difficulties (such as drug errors and product quality concerns) are known as pharmacovigilance (PV). Adverse occurrences after a product's launch are reported to the FDA using an individual case safety report (ICSR). It is a crucial source of information about potential drug safety risks for monitoring post-market safety. A patient's medical history, clinical course, and result can all be included in the clinical lead in ICSRs, as well as any questionable items or products and temporal data relating to product usage and the incidence of adverse events. Complete and accurate reporting of ICSRs is crucial to understanding a drug's safety profile.[268]

AI in pharmacovigilance helps in the early detection and reporting of adverse drug reactions, contributing to patient safety during clinical trials.[269]

The FDA is looking into AI/ML applications to assist regulatory bodies in processing and evaluating ICSR submissions, anticipating a tremendous increase in the ICSR reported.

Processing of Cases

There may be opportunities to automate the ISCR processing process using AI/ML. There are many complex data sources on adverse events for ICSRs, including spontaneous reports, clinical trials, EHRs, social media, phone calls, emails, patient registries, claims data, and post-approval safety studies. Identifying negative occurrences for ICSR submission could be aided by using AI/ML to extract information from source documents. For instance, utilizing AI/ML to filter social media for adverse events and find and assess drug event connections from literature has been investigated.[67,270–272]

After determining an adverse event from a data source, AI/ML could be used for case validity, prioritization, duplicate check, coding, and quality control. AI/ML can assist in determining whether a case is authentic if it satisfies the minimal reporting requirements, such as having an identifiable patient, a suspect drug or biological product, adverse event(s), and an identifiable reporter.[273,274] To aid in prioritizing cases, AI/ML has been used to classify adverse events by expectedness (if an adverse event is known and on the product labeling).[275] Automated duplicate checks leveraging

AI/ML are being carried out to determine whether the issue is an accurate duplicate, a follow-up version of an earlier case, or a new case. Another area where AI/ML has been employed is to code adverse events from ICSRs to structured medical vocabulary phrases for quality control.

Case Analysis

Cases of adverse events are evaluated clinically. The case review procedure includes determining the possibility of a causal connection between a drug and an adverse event and the case's outcome.[270] Based on relevant variables used in causality assessments, an AI model was developed, validated, and tested to classify patients according to the likelihood that a causal association between drugs and adverse events occurs. The gravity of the results of ICSRs has also been assessed using AI/ML, supporting case evaluation, and the timely filing of individual cases that call for expedited reporting.[267]

Submission of a Case

The filing of ICSRs is generally the last action after case processing. To send ICSRs to the FDA, reporting rules have been automated using AI/ML algorithms. It is necessary to report ICSRs both individually and collectively.[276] Gathering safety information for a product and submitting it according to the required schedule regularly is the bulk reporting of adverse occurrences. For reporting reasons, aggregate reports incorporating numerous adverse events for items that happen over time can be created using AI/ML.[277]

Modern Manufacturing

The methods, conditions, and controls employed in the manufacture, processing, packaging, and storage of a drug are essential elements of drug development because they help to ensure that the drug satisfies requirements for quality, purity, and identity as well as safety and effectiveness standards. Applications for advanced analytics leveraging AI/ML in the pharmaceutical manufacturing sector include but are not limited to improving process control, increasing equipment reliability and throughput, monitoring for early warning signs or signals that the manufacturing process is out of control, identifying recurrent problem clusters, and preventing batch losses.

A novel pharmaceutical manufacturing technique or strategy that has the potential to increase the supply chain's resilience, as well as the process's robustness and dependability, is referred to as "advanced manufacturing." Advanced manufacturing might incorporate cutting-edge technical solutions, creatively employ tried-and-true procedures, or utilize production processes in an uncharted field. It applies to both new and already available large and small-molecule pharmaceutical drugs.

With the Emerging Technology Program (ETP),[278] which the FDA started promoting in 2014, sophisticated manufacturing technologies have quickly emerged. To promote prompt technological adoption, adjustments in regulatory rules were also necessary. The FDA uses its risk-based regulatory framework for artificial intelligence (AI) technologies in drug manufacturing, including both chemical and biological drugs. This could entail looking at the use of AI in bioinformatics pipelines as a step in upstream manufacturing to create and select candidates for complex biological products utilized in precision medicine, such as cancer vaccines, cellular therapies, and gene therapies. Stakeholder input on this issue is something we are interested in hearing.

AI/ML could forecast product demand, examine production schedules, and estimate and mitigate the industrial supply chain's reliability risks.

Artificial intelligence (AI) has several potential uses in the pharmaceutical industry, including but not limited to process design and control optimization, intelligent monitoring and maintenance, and trend tracking for continuous development. AI for pharmaceutical manufacture can be used with other sophisticated manufacturing technologies to obtain desired results. The implementation of an Industry 4.0 paradigm, which might lead to a tightly managed, highly connected, digitalized ecosystem and pharmaceutical value chain for the producer, is made possible by AI.

Design of the Process and Scale-Up

By leveraging process development data to construct AI models like machine learning, it is possible to find the best processing parameters and scale up processes more quickly, cutting down on waste and development time. Process design optimization can use digital twins,[279] digital representations of actual processes used to assess, predict, and improve performance. The digital twin may be helpful when analyzing manufacturing processes with little development data. AI/ML models may then use prior product and process knowledge (from, for instance, studies, programs, and scientific literature) to identify the best processing parameters more quickly, reducing design time and waste.

APC (Advanced Process Control)

APC enables manufacturing process dynamic control to produce the required result. AI techniques can also create process controls that can foresee a process unfolding by combining AI with in-The-moment sensor data. APC methods, which combine an understanding of the underlying chemical, physical, and biological transformations occurring in the production process with AI techniques, have been used by several pharmaceutical producers in the past. Conventional process controls maintain the input process parameters' constants. They cannot concurrently change several input parameters to keep the process output parameters at the correct levels.

On the other hand, advanced APC permits dynamic control of the process to achieve the desired outcome.[280] Like neural networks, APC can be implemented using AI/ML, with real-time process data as inputs. By combining real-time sensor data with AI/ML tools, these techniques can also be used to create process controls that can predict whether a process operates within a state of control, such as in conjunction with smart production line monitoring to boost current manufacturing line output and efficiency. It is anticipated that APC methods, which combine physical and chemical knowledge with AI/ML approaches, will soon be employed more frequently. Several pharmaceutical manufacturers have already reported using APC methods. High-quality model inputs influence these APC applications' structure and process comprehension. These reliable inputs enable the generation of model parameters when used in conjunction with data-driven modeling.[281] These models increase model resilience while utilizing the data needed for model development.

Process Observation and Fault Finding

AI techniques can monitor equipment and identify deviations from the norm that demand maintenance, cutting down on process downtime. For instance, AI-based software is used in vision-based quality control to scan images of packaging, labels, or glass vials to find deviations from the standards of a product's specified quality feature. AI methods can be applied to monitor product quality, including packaging quality. Without sacrificing the quality of the final product, automation and real-time monitoring of manufacturing processes can lead to more efficient inventory control, shorter lead times, and higher production output. By monitoring equipment and detecting irregularities in performance, AI/ML algorithms can suggest maintenance processes and reduce process downtime. Another example is computer vision-based quality control, which uses photographs (such as images of labels, packaging, or glass vials) that are inspected by AI/ML-based software to spot discrepancies and guarantee that pictures correspond to specifications for a given quality attribute of a product.

Trend Recognition

To highlight areas for ongoing improvement, AI can be used to analyze deviation reports and consumer complaints that contain enormous amounts of text. To provide a more thorough root cause analysis, this identifies trends in manufacturing-related aberrations. Manufacturing operations can be proactively monitored for trends using AI approaches and KPIs for process performance and capabilities. These techniques can also forecast the points at which effectiveness assessments of corrective and preventive actions will be triggered. AI/ML can enhance manufacturing, including increased output, decreased waste, better-informed decisions, and improved quality control. Analysis of process irregularities is primarily the responsibility of quality professionals and applicable subject matter experts. When examining deviation reports containing enormous amounts of data or text, AI/ML could help identify manufacturing-related deviation

patterns, cluster trouble regions, and highlight areas for continuous proactive development. This makes root causes more straightforward to identify because the human evaluation of deviation patterns can be time-consuming. AI/ML approaches can predict thresholds for starting CAPA effectiveness evaluations and proactively monitor manufacturing operations for trends and out-of-control situations. These indicators include process performance (Ppk) and process capability (Cpk) indicators.

Industry 4.0

AI/ML can be used with other cutting-edge manufacturing technologies (such as process analytical technology and continuous manufacturing) to achieve the desired results. AI/ML can be used to implement Industry 4.0. This phrase describes the fourth industrial revolution, which combines rapidly evolving technology and may provide a well-controlled, hyper-connected, and digitalized pharmaceutical value chain for the producer.[282] "Industry 4.0" refers to the fourth industrial revolution, which combines several rapidly evolving technologies to modify the manufacturing environment profoundly. Integrated, autonomous, and self-organizing production systems define Industry 4.0.[282] The National Academies of Sciences, Engineering, and Medicine recognized advancements in integrated pharmaceutical manufacturing processes in a report from 2021 titled Innovation in Pharmaceutical Manufacturing on the Horizon Technical Challenge, Regulatory Issues, and Recommendations. Since these developments require measurement, modeling, and control technology, AI can significantly contribute to them.

8.1 5.6 Real-Time Improvement

AI can control industrial processes in APC applications by altering process parameters in response to real-time data. Aside from supporting analytical procedures for in-process or final product testing, AI models can provide real-time release testing, forecast in-progress product quality attributes, and support in-process or final product testing. Few industry standards and FDA recommendations are available for the development and validation of models that affect the quality of the products, which may make it difficult to prove the validity of a model in a particular application.

Verification and Validation (V&V 40)

American Society of Mechanical Engineers (ASME) (American Society of Mechanical Engineers, 2018) is the organization that first created this method for assessing the reliability of computational models used in medical devices. Later, it was utilized in the modeling process used to create drugs.[283] Given that AI/ML is also used for computational models, the V&V 40 framework may reveal information on how trustworthy the AI/ML model is for drug development.[284] The V&V 40 Standard has been altered for medical devices and model-informed drug development. It is not AI/ML-specific and does not specify activities or establish criteria required to demonstrate model credibility for a specific context of usage or application.[283,285] In addition to the V&V 40 Standard for determining the prognosticative power of computational models for medical devices, the FDA, Health Canada, and the Drugs and Healthcare Products Regulatory Agency (MHRA) of the United Kingdom jointly published ten guiding principles to direct the development of Good Machine Learning Practices (GMLP) for medical devices that use AI/ML.[286] Some guiding principles include using a complete product life cycle approach, leveraging various information during product development, and thoroughly understanding how the model is integrated into the clinical workflow. The framework suggested a predetermined change control plan mechanism to set the stage for AI/ML-enabled devices with improved adaptability. This mechanism would enable a sponsor to proactively specify intended modifications to devise software incorporating AI/ML and the procedures used to ensure their safety and efficacy.

Regulatory Conformity

A growing amount of information and a clear understanding of the advantages and disadvantages of using AI/ML in drug development support the FDA's consideration of strategies to give regulatory clarity surrounding this practice. The use of AI/ML in drug development may present unique issues that could draw attention to other considerations, even while some standards and methods may be modified to suit them. Considerations or important areas of practice for implementing AI/ML in drug development can be aided by AI. In addition to meeting current requirements to support regulatory

decision-making regarding a drug's safety and effectiveness, the use of AI/ML in drug development raises challenges related to human-led AI/ML governance, accountability, and transparency; data considerations; and model development, performance, monitoring, and validation. Through openness and documentation across the product life cycle, AI/ML can make it easier to develop trust. Consideration of the pre-specification and documentation of the target or research question, the usage context, risk, and AI/ML development may be essential. Regarding data quality, relevance, and dependability, several issues about using AI/ML in the drug development are not particular.

Improvement of AI/ML Utility

To support AI innovation and adoption, the federal government and the international community have increased their commitment (Executive Order No. 13859, Maintaining American Leadership in Artificial Intelligence, February 11, 2019;[287] Exec. Order No. 13960,[288,289] Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government, December 3, 2020).[289] This includes encouraging trustworthy and ethical AI.

One of the ways the FDA has supported the development of creative and effective AI/ML is through the CDER AI Steering Committee (AISC), which coordinates the work surrounding AI/ML utilization throughout therapeutic research. Utilizing its commitment to advancing novel approaches and fostering collaborative efforts across the Agency, CDRH, including the DHCoE, has offered consultations for drug submissions involving AI/ML[290] and is currently developing a framework for AI/ML-based devices, including predetermined change control plans for devices incorporating AI/ML, as well as a foundation for Good Machine Learning Practices for medical device development.[290] To evaluate the security and effectiveness of new AI/ML systems, the FDA has also financed regulatory science research on robustness, user-centered transparency, and bias identification and management through external academic and clinical partnerships. Additionally, the FDA has conducted several workshops and a Patient Engagement Advisory Committee (PEAC) meeting on subjects relevant to DHT and AI/ML. [291,292]

The CDER-created Innovative Science and Technology Approaches for New Drugs (ISTAND) Pilot Program aims to broaden the categories of drug development tools (DDTs) covered by DDT certification processes, including those that employ DHTs. Applications of AI/ML could hasten the delivery of innovative drugs to patients by enhancing the data at hand. Applications of AI/ML may represent new DDTs or aid in the interpretation and analysis of established DDTs (such as clinical outcome evaluations or biomarkers).[293] The FDA's CDER and CBER have developed a MIDD Pilot Program to simplify the development and application of exposure-based, biological, and statistical models derived from nonclinical and clinical data sources.[294] In the context of MIDD, AI/ML could improve mechanistic or predictive safety assessments, dose selection or calculations, and clinical trial simulations.

The FDA's Sentinel Initiative is looking into AI/ML techniques to improve the current systems for post-market safety surveillance. These systems include the Sentinel System[295] from CDER, the Biologics Effectiveness and Safety (BEST) system[296] from CBER, and the National Evaluation System for Health Technology (NEST) initiatives from CDRH.[297] The FDA outlined its goals for using connected claims and EHR data supported by sophisticated analytics in the five-year Sentinel System strategic plan.[298] The Sentinel System Innovation Center has proposed a four-pronged approach to implementing this plan by incorporating emerging data science innovations and EHR data for medical product safety surveillance. The four pillars of this approach are data infrastructure, feature engineering, causal inference, and detection analytics.[299] Examples of AI/ML applications in this strategy include natural language processing (NLP), automated feature extraction from unstructured EHR clinical notes for computable phenotyping, and enhanced confounding adjustment from EHR-based variables. The "Super Learner" algorithms, targeted maximum likelihood estimates, and other advanced statistical and ML techniques are used in these methods.[300]

Better data sources, approaches, strategies, resources, skills, and infrastructure are all part of CBER's BEST system, intended for surveillance and epidemiological research. This program endeavors to forecast or better understand adverse outcomes related to utilizing biological products and other items that CBER regulates by applying AI/ML tools to evaluate EHRs. The FDA's perspective of using AI/ML techniques for creating valid data about product efficacy will expand through these efforts.[301]

CDER is also considering using AI to enhance the evaluation of ICSRs submitted to the FDA Adverse Event Reporting System (FAERS).[302] With the aid of AI/ML, the Information Visualization Platform (InfoViP) was developed

to recognize duplicate ICSRs, classify ICSRs based on the quality of the information they include, and produce timeline representations of clinical events to aid in the analysis of reported adverse events.[303–305]

AI/ML techniques have been investigated to automate the detection of adverse events in drug product labeling, which will help safety reviewers prioritize ICSRs to facilitate the identification of unrecognized or unexpected safety concerns.[306,307] Another AI-based technology that focuses on drug product labeling is the Computerized Labeling Assessment technology (CLAT), which automates the analysis of labels and labeling (such as prescribing information, carton, and container labeling). NLP and ML are also being researched to classify free-text narratives in FAERS ICSRs into structured medical vocabulary drug error terms, improving the human evaluation of coding quality. Additionally, through the FDA Quality Metrics Reporting Program,[308] CDER's Emerging Technology Program, and CBER's Advanced Technologies Team (CATT) Program,[308] FDA engaged the industry and gathered crucial feedback on AI/ML use cases in pharmaceutical manufacture.

The FDA also uses methods like a Broad Agency Announcement[309] to obtain extramural submissions that address new regulatory and scientific goals. These proposals may use outside infrastructure and expertise to illuminate the approaches to integrating and assessing AI/ML in drug development.

The AI model's decisions are influenced by how reliable the underlying data is. Due to this, preventing unintentional biases during model building and validation can be difficult. To speed up model building. AI programs can store knowledge obtained while working on a particular use case and subsequently use it to construct a related use case. Applications and manufacturers could be unsure how the potential for learning transfer from one AI model to another can be considered during the model design and validation process. As AI techniques become more complex, it becomes more difficult to describe how changes in model inputs affect model outputs. In these circumstances, applicants can have trouble articulating standards that verify the model, uphold the model's output's explicability, and affect product quality.

To address these concerns and predict the future, the CDER, CBER, and CDRH at the FDA and DHCoE are looking into human-led governance, accountability, and transparency, as well as model development, performance, monitoring, and validation. They achieve this by modifying the General Accountability Office AI accountability framework's general concepts.[289] In each of these groups, a risk-based approach is recommended.

Software as a Medical Device (SaMD)

The ability of AI/ML software to improve performance and learn from actual use and experience is a crucial benefit. The power of AI/ML software to learn from real-world feedback (training) and improve its performance (adaptation) puts it in a favorable position for software as a medical device (SaMD) and a fast-expanding field of study and development. SaMD enabled by AI/ML can provide precise and practical software functionality that improves patient care standards. Meetings for the Critical Path Innovation Meetings (CPIM),[310] ISTAND Pilot Program,[311] Emerging Technology Program,[312] and Real-World Evidence Program[313] are other forums for exchanging ideas and discussing a suitable AI/ML methodology or technology to improve drug development's efficiency and quality. These programs are among other significant projects.

Conclusions

The AI/ML models are now extensively used to expedite drug discovery; the FDA has taken proactive actions to ensure that machine-made decisions are at least as robust and reliable as human analysis; in certain instances, AI/ML is now used to enhance human decision-making ability. Dozens of regulations, guidelines, and White Papers have been introduced to ensure that machine-driven decisions are continuously challenged.

Of the tremendous significance of AI/ML, the first is its ability to process extensive data that is impossible to achieve otherwise. The entry of quantum computers is waiting to make these analyses more efficient. Second, data analysis in the research and manufacturing to optimize the process and validate findings. An excellent example of how this aspect has been the mainstream is the 21CFR Part 11[314] compliance that ensures digital data storage remains accurate and unalterable. Third, AI/ML allows experimental design creation that can be the most creative application to avoid following traditional paths, despite their efficiency or inaccuracies. The design space includes statistical modeling that

can help reduce the study size by calculation from large databases that would generally not be possible using human inquiry alone.

The AI/ML applications will expand exponentially; while, in 2021, more than hundreds of FDA filings deployed AI/ML tools, I anticipate most filings to engage these tools somehow; thus, scientists must receive proper training in using these tools. The FDA and related agencies have published many guidelines, policies, and data handling suggestions that require personnel trained to identify redundancies in implementing AI/ML at every stage of development and manufacturing.

The role of AI/ML in manufacturing brings the traditional models to a different stage—advanced manufacturing that has benefitted significantly from automation, machine learning, and decision-making in continuous process control. None of these improvements were possible by human interaction alone.

While AI/ML enhance productivity, this will not impact the job market since humans will still be needed to create the algorithms, supervise their operation, and finally judge their conclusions.

To introduce the many applications of AI/ML, the FDA is open to earlier meetings to discuss the nature and extent of reliance on AI/ML in drug development or manufacturing. Developers should capitalize on these opportunities. It should be noted that the FDA guidelines admit that they are learning too when these interactions arise; the goal is to make the discovery more efficient, and thus more affordable; and manufacturing more reliable and efficient. All these plans will help reduce the cost of drugs to patients, and that alone should be a sufficient reason to remove mindsets about AI/ML, perhaps using another AI/ML model.

However, it all starts with assuring that the myths around AI/ML are curbed at the onset. The industry should trust the FDA's efforts to make drug discovery more productive, manufacturing more compliant, and clinical testing more ethical.

Disclosure

The author reports no conflicts of interest in this work.

References

1. IMDRF/AIMD WG/N67 Machine Learning-enabled Medical Devices Key Terms and Definitions, final document; 2022. Available from: Https://www.imdrf.org/documents/machine-learning-enabled-medical-devices-key-terms-and-definitions. Accessed July 10, 2023.
2. Turing AM. Computing machinery and intelligence. *Mind*. 1950;59(236):433–460. doi:10.1093/mind/LIX.236.433
3. McCarthy J, Minsky ML, Rochester N, Shannon CE. A proposal for the Dartmouth Summer Research Project on Artificial Intelligence. *AI Mag*. 1956;27(4):12–14.
4. Newell A, Simon HA. The Logic Theorist—A case study in heuristics. Proceedings of the Western Joint Computer Conference Contrasts in Scientific Style. 1956. 74–90.
5. Rosenblatt F. The perceptron A probabilistic model for information storage and organization in the brain. *Psychol Rev*. 1958;65(6):386–408. doi:10.1037/h0042519
6. Gugerty L. Newell and Simon's Logic Theorist: Historical Background and Impact on Cognitive Modeling. *Proce Human Factors Ergonomics Society Ann Meeting*. 2006;50(9):880–884. doi:10.1177/154193120605000904.
7. Colmerauer A, Roussel P. The birth of Prolog. Proceedings of the Logic Programming Workshop. 1972. 1–4. Available from: https://groups.seas.harvard.edu/courses/cs252/2016fa/10.pdf. Accessed July 10, 2023.
8. Shortliffe EH, Buchanan BG. A model of inexact reasoning in medicine. *Math Biosci*. 1975;23(3–4):351–379. doi:10.1016/0025-5564(75)90047-4
9. Davis R, Lenat D. *Knowledge-Based Systems in Artificial Intelligence*. McGraw-Hill; 1982.
10. McCarthy J, Brayton RK, Eds. *LISP 1.5 Programmer's Manual*. MIT Press; 1983.
11. Lenat DB. CYC A large-scale investment in knowledge infrastructure. *Commun ACM*. 1995;38(11):33–38. doi:10.1145/219717.219745
12. Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors. *Nature*. 1986;323(6088):533–536. doi:10.1038/323533a0
13. Werbos PJ. Generalization of backpropagation with application to a recurrent gas market model. *Neural Networks*. 1988;1(4):339–356. doi:10.1016/0893-6080(88)90007-X
14. Campbell M, Hoane AJ, Hsu FH. Deep Blue. *Artif Intell*. 2002;134(1–2):57–83. doi:10.1016/S0004-3702(01)00129-1
15. LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. *Procee IEEE*. 1998;86(11):2278–2324. doi:10.1109/5.726791
16. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L. ImageNet A large-scale hierarchical image database. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Miami, FL, USA. 2009. 248–255. doi:10.1109/CVPR.2009.5206848.
17. Ferrucci D, Brown E, Chu-Carroll J, et al. Building Watson An overview of the DeepQA project. *AI Mag*. 2010;31(3):59–79. doi:10.1609/aimag.v31i3.2303
18. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM*. 2017;60;84–90. doi:10.1145/3065386.

19. Silver D, Huang A, Maddison CJ, et al. Mastering the game of Go with deep neural networks and tree search. *Nature*. 2016;529(7587):484–489. doi:10.1038/nature16961
20. Silver D, Schrittwieser J, Simonyan K, et al. Mastering the game of Go without human knowledge. *Nature*. 2017;550(7676):354–359. doi:10.1038/nature24270
21. Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. *Adv Neural Inf Process Syst*. 2017;5998–6008.
22. Devlin J, Chang MW, Lee K, Toutanova K. BERT Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics. Human Language Technologies (NAACL-HLT). 2018. 4171–4186.
23. Vinyals O, Babuschkin I, Czarnecki WM, et al. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature*. 2019;575 (7782):350–354. doi:10.1038/s41586-019-1724-z
24. Jumper J, Evans R, Pritzel A, et al. Applying and improving AlphaFold at CASP14. *Proteins Structure Function Bioinformatics*. 2021;89 (12):1711–1721. doi:10.1002/prot.26257
25. Partnership on AI. Available from: https://www.partnershiponai.org/about/. Accessed July 10, 2023.
26. OpenAI. Charter. Available from: Https://www.openai.com/charter/. Accessed July 10, 2023.
27. IEEE. The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Available from: https://ethicsinaction.ieee.org/. Accessed July 10, 2023.
28. Deng J, Lin Y. The Benefits and Challenges of ChatGPT. An Overview. *Front Computing Intelligent Sys*. 2023;2(2):81–83. doi:10.54097/fcis.v2i2.4465
29. Available from: Https://www.fda.gov/science-research/science-and-research-special-topics/artificial-intelligence-and-machine-learning-aiml-drug-development. Accessed July 10, 2023.
30. Available from: Https://openai.com/. Accessed July 10, 2023.
31. Jongsma KR, Van Solinge WW, Haitjema S. Acht misvattingen over AI in de zorg [Eight misconceptions about AI in healthcare]. *Ned Tijdschr Geneeskd*. 2023;167:D7578.
32. Fjelland R. Why general artificial intelligence will not be realized. *Humanities Social Sci Commun*. 2020;7(1):1–9. doi:10.1057/s41599-020-0494-4
33. Bohannon J. Fears of an AI pioneer. *Science*. 2015. doi:10.1126/science.349.6245.252
34. Available from: https://www.nytimes.com/2023/06/10/technology/ai-humanity.html. Accessed July 10, 2023.
35. Rydell J, Racey P. Street lamps and the feeding ecology of insectivorous bats. *J Mammal*. 1995;76(2):430–446.
36. Narayan R, McClintock JE. Observational Evidence for a Correlation Between Jet Power and Black Hole Spin. *Class Quantum Grav*. 2013;30 (24):L69–L73. doi:10.1111/j.1745-3933.2011.01181.x.
37. Crouch CH, Mazur E. Peer Instruction: ten years of experience and results. *Am J Phys*. 2001;69(9):970–977. doi:10.1119/1.1374249
38. Kahneman D. *Thinking, Fast and Slow*. Farrar, Straus and Giroux; 2011.
39. Mauro JC, Varshneya AK, Gupta PK, et al. Enthalpy landscapes and the glass transition, in: Fluid Structure Interaction II. *Fluid Stru Interaction II*. 2008;80(9):241–281. doi:10.1007/978-1-4020-9741-6_15.
40. Galilei G. *Two New Sciences. Translated by Stillman Drake, 1974*. University of Wisconsin Press; 1638.
41. Mindell DP. *The Evolving World: Evolution in Everyday Life*. Harvard University Press; 2013.
42. Radford B. The Ten-Percent Myth. *Sci Am Mind*. 2014;25(2):28–33.
43. Rakov VA, Uman MA. *Lightning: Physics and Effects*. Cambridge University Press; 2003.
44. Lewandowsky S, Gignac GE, Vaughan S. The pivotal role of perceived scientific consensus in acceptance of science. *Nat Clim Chang*. 2013;3 (4):399–404. doi:10.1038/nclimate1720
45. Bishop BA, Anderson CW. Student conceptions of natural selection and its role in evolution. *J Research Sci Teach*. 1990;27(5):415–427. doi:10.1002/tea.3660270503
46. National Geographic Society. *Ostrich*. National Geographic Society; 2010.
47. Beacham RC. *Spectacle Entertainments of Early Imperial Rome*. Yale University Press; 1999.
48. Byrne CJ. *The Far Side of the Moon: A Photographic Guide*. Springer Science & Business Media; 2007.
49. Gordon JE. *Structures: Or Why Things Don't Fall Down*. Da Capo Press; 1978.
50. Valtin H. Drink at least eight glasses of water a day. Really? Is there scientific evidence for "8x8"? *Am J Phys*. 2002;283(5):R993–R1004. doi:10.1152/ajpregu.00365.2002
51. Goss KD. *The Salem Witch Trials: A Reference Guide*. Greenwood Press; 2008.
52. Sears ME, Kerr KJ, Bray RI, et al. Arsenic, Cadmium, Lead, and Mercury in Sweat: a Systematic Review. *J Environ Public Health*. 2012;2012:1–10. doi:10.1155/2012/184745
53. Available from: https://www.thefountaininstitute.com/blog/chat-gpt-ethics. Accessed July 10, 2023.
54. Rottman D Available from: https://www.technologyreview.com/2023/03/25/1070275/chatgpt-revolutionize-economy-decide-what-looks-like/. Accessed July 10, 2023.
55. Available from: https://blog.enterprisedna.co/is-chat-gpt-safe/#~text=Using%20ChatGPT%20Safely-,Is%20ChatGPT%20Safe%20to%20Use%3F,that%20it%20sounds%20human%2Dlike. Accessed July 10, 2023.
56. Available from: https://www.forbes.com/sites/forbestechcouncil/2023/03/10/the-unforeseen-consequences-of-chatgpt/?sh=39d2e3b91eea. Accessed July 10, 2023.
57. ElZarrad MK, Lee AY, Purcell R, Steele SJ. Advancing an agile regulatory ecosystem to respond to the rapid development of innovative technologies. *Clin Transl Sci*. 2022;15(5):1332–1339. doi:10.1111/cts.13267
58. Chen M, Hao Y, Sun Y, Zhang S, Tian H. Artificial intelligence in drug safety assessment Focus on drug-induced liver injury. *Drug Metab Rev*. 2020;52(2):250–265.
59. Stokes JM, Yang K, Swanson K. A deep learning approach to antibiotic discovery. *Cell*. 2020;180(4):688–702. doi:10.1016/j.cell.2020.01.021
60. Pappalardo F, Russo G, Candido S, Pennisi M, Cavalieri S. Artificial intelligence for the development of individualized treatment in autoimmune diseases. *J Clin Med*. 2020;9(2):438. doi:10.3390/jcm9020438

61. Liu Y, Zhang W, Li L. Deep learning-based transcriptome data classification for drug response prediction. *BMC Genomics*. 2018;19(Suppl 7):682. doi:10.1186/s12864-018-5071-5
62. Finkbeiner S, Frumkin M, Kao J. 7.14 - High-Content Screening (HCS) as a Discovery Tool for Neurological Diseases. In: Gereau AK IV, Sarah AL, editors. *Neurotherapeutics*. Academic Press; 2015:237–262.
63. Subudhi SK, Aparicio A. Predictive biomarkers for checkpoint inhibitor-based immunotherapy. *Lancet Oncol*. 2019;20(8):e431–e441.
64. Haque A, Engel J, Teichmann SA, Lönnberg T. A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Med*. 2017;9(1):75. doi:10.1186/s13073-017-0467-4
65. Zhou Y, Zhang Z, Tay FEH. Drug-target interaction prediction a deep learning-based approach. *Neurocomputing*. 2019;321:321–330.
66. Papin C, Aransay NR, Isambert N, et al. Optimising patient selection for oncology Phase 1 trials using kernel methods A data-driven feasibility study. *Br J Clin Pharmacol*. 2018;84(12):2877–2886. doi:10.1111/bcp.13753
67. Wang W, Haerian K, Salmasian H, Harpaz R, Chase H, Friedman C. A drug-adverse event extraction algorithm to support pharmacovigilance knowledge mining from PubMed citations. *AMIA Annual Symposium Proceedings*. 2011. 2011. 1464–1470.
68. Paraskevopoulou MD, Maniou S, Panagiotopoulou M, Vlachakis D, Iliopoulos I. DrugCombi a web application for analyzing and visualizing drug combinations. *Cancer Res*. 2020;80(16 Supplement):4804.
69. Kluger Y, Basri R, Chang JT. Machine learning tools for drug combination discovery and optimization. *Mol Syst Biol*. 2020;16(9):e9687.
70. Zitnik M, Agrawal M, Leskovec J. Modeling polypharmacy side effects with graph convolutional networks. *Bioinformatics*. 2018;34(13):i457–i466. doi:10.1093/bioinformatics/bty294
71. Kumar V, Medabalimi S. Applications of artificial intelligence in pharmaceutical continuous manufacturing. *J Pharm Innov*. 2021;16(2):189–198.
72. Gómez-Bombarelli R, Wei JN, Duvenaud D, et al. Automatic chemical design using a data-driven continuous representation of molecules. *ACS Central Science*. 2018;4(2):268–276. doi:10.1021/acscentsci.7b00572
73. Shen L, Du W, Liang XJ. Intelligent drug delivery systems inspired by biomimetic design and advancing technologies. *Adv Drug Deliv Rev*. 2021;176:113882.
74. Du Toit LC, Pillay V, Danckwerts MP, Fang S, Crews AL, Adler KB. Tuberculosis chemotherapy current drug delivery approaches. *Respir Res*. 2011;12(1):118. doi:10.1186/1465-9921-12-118
75. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*. 2017;542(7639):115–118. doi:10.1038/nature21056
76. Hodos RA, Kidd BA, Shameer K, Readhead BP, Dudley JT. In silico methods for drug repurposing and pharmacology. *Wiley Interdiscip Rev Syst Biol Med*. 2016;8(3):186–210. doi:10.1002/wsbm.1337
77. Yuan Q, Li H, Zhang J, Liu X. A review of in silico ADME/T modeling approaches. *J Pharmacol Toxicol Methods*. 2021;113:106881.
78. Gottlieb A, Stein GY, Oron Y. Prediction of intracellular drug concentrations in the central nervous system. *Eur J Pharmacol*. 2012;674(2–3):285–292.
79. Kostal J, Voutchkova-Kostal A, Anastas PT, Zimmerman JB. Identifying and designing chemicals with minimal acute aquatic toxicity. *Proce National Acad Sci*. 2012;112(20):6289–6294. doi:10.1073/pnas.1314991111.
80. Chen C, Guo Y, Grossman RL. Elucidating the landscape of aberrant DNA methylation in hepatocellular carcinoma through epigenomic data analysis. *Front Genet*. 2018;9:15. doi:10.3389/fgene.2018.00015
81. Vora LK, Gholap AD, Jetha K, et al. Artificial Intelligence in Pharmaceutical Technology and Drug Delivery Design. *Pharmaceutics*. 2023;15(7):1916. doi:10.3390/pharmaceutics15071916.
82. Gupta A, Bishnoi A, Sharma AK. Artificial intelligence in pharmaceutical product development. *Int J Pharm Sci Res*. 2019;10(3):1012–1021.
83. Anzalone AV, Randolph PB, Davis JR, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature*. 2019;576(7785):149–157. doi:10.1038/s41586-019-1711-4
84. Zhou J, Troyanskaya OG. Predicting effects of noncoding variants with deep learning-based sequence model. *Nat Methods*. 2015;12(10):931–934. doi:10.1038/nmeth.3547
85. Moore CB, Wallace JR, Frase AT, Pendergrass SA, Ritchie MD. BioBin: a bioinformatics tool for automating the binning of rare variants using publicly available biological knowledge. *BMC Med Genomics*. 2013;6(Suppl 2):S6. doi:10.1186/1755-8794-6-S2-S6
86. Unterthiner T, Mayr A, Klambauer G, et al. Deep learning as an opportunity in virtual screening. *Proce Deep Learning Workshop NIPS*. 2014;27(1):1–9.
87. Wang S, Summers RM. Machine learning and radiology. *Med Image Anal*. 2012;16(5):933–951. doi:10.1016/j.media.2012.02.005
88. Krieg C, Nowicka M, Guglietta S, et al. High-dimensional single-cell analysis predicts response to anti-PD-1 immunotherapy. *Nat Med*. 2018;24(2):144–153. doi:10.1038/nm.4466
89. Pestian J, Glauser T, Matykiewicz P, Forestier G, Moreau-Gaudry A, Jannin P. Using natural language processing and machine learning algorithms to categorize medical devices for multiple taxonomies. *J Biomed Inform*. 2017;67:34–43. doi:10.1016/j.jbi.2017.02.001
90. Chen J, Xu X, Xiao W, et al. An ensemble machine learning model for predicting sites of metabolism of xenobiotics. *Sci Rep*. 2019;9(1):1–12. doi:10.1038/s41598-018-37186-2
91. Robinson O, Want E, Coen M, Kennedy S, Spector T, Holmes E. Metabolic phenotyping and systems biology approaches to understanding metabolic syndrome and fatty liver disease. *Gastroenterology*. 2014;146(1):46–62. doi:10.1053/j.gastro.2013.11.001
92. Pasolli E, Truong DT, Malik F, Waldron L, Segata N. Machine learning meta-analysis of large metagenomic datasets: tools and biological insights. *PLoS Comput Biol*. 2016;12(7):e1004977. doi:10.1371/journal.pcbi.1004977
93. Baek M, DiMaio F, Anishchenko I, et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science*. 2021;373(6557):871–876. doi:10.1126/science.abj8754
94. Akram MU, Shah MA, Han J. Mining big data: current status, and forecast to the future. *Int J Inf Manage*. 2019;44:47–63.
95. Zhang L, Zhang X, Ju H, Chen Z, Xu T. Integrating machine learning and metagenomics for drug discovery. *Front Genet*. 2019;10:446. doi:10.3389/fgene.2019.00446
96. Bento AP, Gaulton A, Hersey A, et al. The ChEMBL bioactivity database: an update. *Nucleic Acids Res*. 2014;42(D1):D1083–D1090. doi:10.1093/nar/gkt1031

97. Hamp T, Getz G, Khalili A, Cesnik AJ. Artificial intelligence in neurodegenerative disease research: current status and future perspectives. *Alzheimer's Dementia*. 2020;16(11):1576–1590.
98. Paulus MP, Thompson WK. The Challenges and Opportunities of Small Data in Mental Health. *Biol Psychiatry*. 2019;4(9):772–775.
99. Burlina P, Joshi N, Pekala M, Pacheco KD, Freund DE, Bressler NM. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. *JAMA Ophthalmol*. 2017;135(11):1170–1176. doi:10.1001/jamaophthalmol.2017.3782
100. Chen Y, Chu F, Chen H, Wang Y, Wang Z. Orphan drug discovery: an automated screening approach on a heterogeneous network-based inference method. *J Biomed Inform*. 2019;94:103184. doi:10.1016/j.jbi.2019.103184
101. Jalali A, Shirazi FH, Chetty G. Computational approaches for identifying novel analgesics: A review. *Front Pharmacol*. 2020;11:607. doi:10.3389/fphar.2020.00607
102. Modamio S, López-Coronado M, Jódar-Sánchez F. Analysis of machine learning techniques for heart failure readmissions. *Artif Intell Med*. 2018;90:43–54.
103. Korotkevich G, Sukhov V, Budin B. Fast gene set enrichment analysis. *bioRxiv*. 2016. doi:10.1101/060012.
104. Chen C, Liaw A, Breiman L. Using random forest to learn imbalanced data. *Univ California Berkeley*. 2004;110(1):24.
105. Ezzat A, Wu M, Li XL, Kwoh CK. Drug combination sensitivity scoring facilitates the discovery of synergistic and efficacious drug combinations in cancer. *PLoS Comput Biol*. 2019;15(3):e1006752. doi:10.1371/journal.pcbi.1006752
106. Gálvez-Peralta M, Camacho-Molina A, Medeiros M. Artificial intelligence applied to medicine: Current trends and future directions. *Front Med*. 2019;6:139. doi:10.3389/fmed.2019.00139
107. Aliper A, Plis S, Artemov A, Ulloa A, Mamoshina P, Zhavoronkov A. Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data. *Mol Pharm*. 2016;13(7):2524–2530. doi:10.1021/acs.molpharmaceut.6b00248
108. Ryan PB, Schuemie MJ, Welebob E, Duke J, Valentine S, Hartzema AG. Defining a reference set to support methodological research in drug safety. *Drug Safety*. 2013;36(1):33–47. doi:10.1007/s40264-013-0097-8
109. Hicks JK, Swen JJ, Gaedigk A. Challenges in CYP2D6 phenotype assignment from genotype data: A critical assessment and call for standardization. *Curr Drug Metab*. 2020;21(8):665–681.
110. Rostami-Hodjegan A, Tucker GT. Simulation and prediction of in vivo drug metabolism in human populations from in vitro data. *Nat Rev Drug Discov*. 2007;6(2):140–148. doi:10.1038/nrd2173
111. Zhang L, Beal SL, Sheiner LB. Simultaneous vs. sequential analysis for population PK/PD data I: best-case performance. *J Pharmacokinet Pharmacodyn*. 2003;30(6):387–404. doi:10.1023/B:JOPA.0000012998.04442.1f
112. Harpaz R, DuMouchel W, LePendu P, Bauer-Mehren A, Ryan P, Shah NH. Performance of pharmacovigilance signal-detection algorithms for the FDA adverse event reporting system. *Clin Pharmacol Ther*. 2013;93(6):539–546. doi:10.1038/clpt.2013.24
113. Mervin LH, Afzal AM, Drakakis G, Lewis R, Engkvist O, Bender A. Target prediction utilising negative bioactivity data covering large chemical space. *J Cheminform*. 2015;7(1):51. doi:10.1186/s13321-015-0098-y
114. Vamathevan J, Clark D, Czodrowski P, et al. Applications of machine learning in drug discovery and development. *Nat Rev Drug Discov*. 2019;18(6):463–477. doi:10.1038/s41573-019-0024-5
115. Costello JC, Heiser LM, Georgii E, et al. A community effort to assess and improve drug sensitivity prediction algorithms. *Nat Biotechnol*. 2014;32(12):1202–1212. doi:10.1038/nbt.2877
116. Rajput MKS, Prasad R, Bhatt P, Sharma A. Deep neural network models for predicting drug-induced liver injury using structural and bioactivity data. *J Chem Inf Model*. 2019;59(7):3128–3138. doi:10.1021/acs.jcim.9b00105
117. Vilar S, Uriarte E, Santana L, et al. Similarity-based modeling in large-scale prediction of drug-drug interactions. *Nat Protoc*. 2016;9(9):2147–2163. doi:10.1038/nprot.2014.151
118. Zhang W, Liu F, Luo L, Zhang J. Predicting drug side effects by multi-label learning and ensemble learning. *BMC Bioinform*. 2017;16(1):365. doi:10.1186/s12859-015-0774-y
119. Ching T, Himmelstein DS, Beaulieu-Jones BK, et al. Opportunities and obstacles for deep learning in biology and medicine. *J Royal Soc Interface*. 2018;15(141):20170387. doi:10.1098/rsif.2017.0387
120. Cherkasov A, Muratov EN, Fourches D, et al. QSAR modeling: where have you been? Where are you going to?. *J Med Chem*. 2014;57(12):4977–5010. doi:10.1021/jm4004285
121. Cui Y, Zhang Y, Zhao Y, Zhang L, Li D. Artificial intelligence in rare diseases: applications, challenges, and future directions. *Hum Genet*. 2021;140(7):963–979.
122. Rathore S, Gupta RD, Sharma A. Impact of artificial intelligence in pharma and life sciences: demystifying the future. *Front Artificial Intelligence*. 2019;2:26. doi:10.3389/frai.2019.00026
123. Forsström T, Axelson H, Stertman L, Virding S. Artificial intelligence for regulatory decision support. *Regulatory Toxicol Pharmacol*. 2017;91:S45–S49.
124. Cheng F, Desai RJ, Handy DE, et al. Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nat Commun*. 2018;9(1):1–12. doi:10.1038/s41467-018-05116-5
125. Napolitano F, Zhao Y, Moreira VM, et al. Drug repositioning a machine-learning approach through data integration. *J Cheminform*. 2013;5(1):1–14. doi:10.1186/1758-2946-5-30
126. Gupta MK, Chauhan S. Artificial intelligence in formulation development. *Curr Opin Chem Eng*. 2020;28:107–114.
127. Luo H, Zhang P, Huang H, et al. D3Targets-2019-nCoV a webserver for predicting drug targets and for multi-target and multi-site based virtual screening against COVID-19. *Acta Pharmaceutica Sinica B*. 2020;10(7):1239–1248. doi:10.1016/j.apsb.2020.04.006
128. Maffucci I, Contini A. In Silico Drug Repurposing for SARS-CoV-2 Main Proteinase and Spike Proteins. *J Proteome Res*. 2020;19(11):4637–4648. doi:10.1021/acs.jproteome.0c00383.
129. Nath N, Remya R, Bhaumik S, Muthusamy A. Machine learning-based drug resistance prediction in cancer A systematic literature review. *Brief Bioinform*. 2019;20(1):300–316.
130. Costello JC, Heiser LM. Artificial intelligence for drug response prediction in cancer. *Mol Oncol*. 2019;13(4):759–766.
131. Tatonetti NP, Ye PP, Daneshjou R, Altman RB. Data-driven prediction of drug effects and interactions. *Science*. 2012;337(6104):1625–1628.
132. Zhang L, Tan J, Han D, Zhu H. From machine learning to deep learning: progress in machine intelligence for rational drug discovery. *Drug Discov Today*. 2017;22(11):1680–1685. doi:10.1016/j.drudis.2017.08.010

133. Garg A, Bhargava A. AI-based inventory management of perishable products in the healthcare sector. *Comput Ind Eng*. 2020;148:106652.
134. Segler MH, Preuss M, Waller MP. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature*. 2018;555(7698):604–610. doi:10.1038/nature25978
135. Carbonell P, Jervis AJ, Robinson CJ, et al. An automated Design-Build-Test-Learn pipeline for enhanced microbial production of fine chemicals. *Commun Biol*. 2018;1(1):1–10. doi:10.1038/s42003-018-0076-9
136. Jiménez J, Skalic M, Martínez-Rosell G, De Fabritiis G, Valencia A. DeepDock: a novel deep-learning architecture for protein–ligand interactions. *J Chem Inf Model*. 2018;58(4):816–823. doi:10.1021/acs.jcim.7b00717
137. Mendez D, Gaulton A, Bento AP, et al. ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Res*. 2019;47(D1):D930–D940. doi:10.1093/nar/gky1075
138. Menden MP, Iorio F, Garnett M, McDermott U, Benes CH. Machine learning prediction of cancer cell sensitivity to drugs based on genomic and chemical properties. *PLoS One*. 2013;8(4):e61318. doi:10.1371/journal.pone.0061318
139. Lee I, Nam H, Kim K, Ghang H. Recent advances in genome mining of secondary metabolites in Aspergillus terreus. *Front Microbiol*. 2019;9:3235. doi:10.3389/fmicb.2018.03235
140. Mayr A, Klambauer G, Unterthiner T, Hochreiter S. DeepTox: Toxicity Prediction using Deep Learning. *Front Environ Sci*. 2016;3:80. doi:10.3389/fenvs.2015.00080
141. Ong E, Wong MU, Huffman A, He Y. COVID-19 Coronavirus Vaccine Design Using Reverse Vaccinology and Machine Learning. *Front Immunol*. 2020;11:1581. doi:10.3389/fimmu.2020.01581
142. Nath A, Lee AC, Engler M, et al. Machine learning-assisted virtual screening A review. *Mol Inform*. 2021;40(3):2000133.
143. Marzullo L, Schölkopf B. Predicting drug withdrawal symptoms from social media data. *PLoS Comput Biol*. 2020;16(12):e1008352.
144. Collins FS, Morgan M, Patrinos A. The Human Genome Project: Lessons from Large-Scale Biology. *Science*. 2003;300(5617):286–290. doi:10.1126/science.1084564
145. Copeland RA. Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis. 2nd ed. Wiley; 2000.
146. Gordon JW, Singer GA. The mouse: the laboratory animal of the year. *Genome Biol*. 2004;5(4):117. doi:10.1186/gb-2004-5-10-117
147. Fang Y, Fu D, Shen XZ, Aburatani H, Hayes JD, Yamamoto M. Quantitative Cell-based Protein Degradation Assays to Identify and Classify Drugs that Target the Ubiquitin-Proteasome System. *J Biol Chem*. 2008;283(46):33554–33562. doi:10.1074/jbc.M804597200
148. Mayr LM, Bojanic D. Novel trends in high-throughput screening. *Curr Opin Pharmacol*. 2009;9(5):580–588. doi:10.1016/j.coph.2009.08.004
149. Kuntz ID, Blaney JM, Oatley SJ, Langridge R, Ferrin TE. A Geometric Approach to Macromolecule-Ligand Interactions. *J Mol Biol*. 1982;161(2):269–288. doi:10.1016/0022-2836(82)90153-X
150. Swinney DC, Anthony J. How were new medicines discovered?. *Nat Rev Drug Discov*. 2011;10(7):507–519. doi:10.1038/nrd3480
151. Jensen LJ, Saric J, Bork P. Literature mining for the biologist: from information retrieval to biological discovery. *Nat Rev Genet*. 2006;7(2):119–129. doi:10.1038/nrg1768
152. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology. 3rd ed. Lippincott Williams & Wilkins; 2008.
153. Kneller R. The Importance of New Companies for Drug Discovery: origins of a Decade of New Drugs. *Nat Rev Drug Discov*. 2010;9(11):867–882. doi:10.1038/nrd3251
154. Schneider G. Automating drug discovery. *Nat Rev Drug Discov*. 2018;17(2):97–113. doi:10.1038/nrd.2017.232
155. Chen H, Engkvist O, Wang Y, Olivecrona M, Blaschke T. The rise of deep learning in drug discovery. *Drug Discov Today*. 2018;23(6):1241–1250. doi:10.1016/j.drudis.2018.01.039
156. Mamoshina P, Vieira A, Putin E, Zhavoronkov A. Applications of Deep Learning in Biomedicine. *Mol Pharm*. 2016;13(5):1445–1454. doi:10.1021/acs.molpharmaceut.5b00982
157. Wallach I, Heifets A. Most Ligand-Based Classification Benchmarks Reward Memorization Rather than Generalization. *J Chem Inf Model*. 2018;58(5):916–932. doi:10.1021/acs.jcim.7b00403
158. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med*. 2019;25(1):44–56.
159. Tari LB, Patel JH. Systematic drug repositioning through mining adverse event data in ClinicalTrials.gov. *PeerJ*. 2014;2:e304. doi:10.7717/peerj.304
160. Dudley JT, Deshpande T, Butte AJ. Exploiting drug-disease relationships for computational drug repositioning. *Brief Bioinform*. 2011;12(4):303–311. doi:10.1093/bib/bbr013
161. Mak KK, Pichika MR. Artificial intelligence in drug development: present status and future prospects. *Drug Discov Today*. 2019;24(3):773–780. doi:10.1016/j.drudis.2018.11.014
162. Zitnik M, Nguyen F, Wang B, Leskovec J, Goldenberg A, Hoffman MM. Machine learning for integrating data in biology and medicine: principles, practice, and opportunities. *Information Fusion*. 2019;50:71–91. doi:10.1016/j.inffus.2018.09.012
163. Collins FS, Varmus H. A new initiative on precision medicine. *N Eng J Med*. 2015;372(9):793–795. doi:10.1056/NEJMp1500523
164. Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-World Evidence — What Is It and What Can It Tell Us?. *N Eng J Med*. 2016;375(23):2293–2297. doi:10.1056/NEJMsb1609216
165. Camacho DM, Collins KM, Powers RK, Costello JC, Collins JJ. Next-Generation Machine Learning for Biological Networks. *Cell*. 2018;173(7):1581–1592. doi:10.1016/j.cell.2018.05.015
166. Preuer K, Lewis RPI, Hochreiter S, Bender A, Bulusu KC, Klambauer G. DeepSynergy: predicting anti-cancer drug synergy with Deep Learning. *Bioinformatics*. 2018;34(9):1538–1546. doi:10.1093/bioinformatics/btx806
167. Jang H, Pluciński L, Taniguchi R, Gross SP, Julaiti A. A Machine Learning Approach to Predicting Chemical Reactions. *Chem Sci*. 2018;9(8):2262–2270.
168. Wang L, Wu Y, Deng Y, et al. Accurate and reliable prediction of relative ligand binding potency in prospective drug discovery by way of a modern free-energy calculation protocol and force field. *J Am Chem Soc*. 2015;137(7):2695–2703. doi:10.1021/ja512751q
169. Xu Y, Dai Z, Chen F, Gao S, Pei J, Lai L. Deep Learning for Drug-Induced Liver Injury. *J Chem Inf Model*. 2015;55(10):2085–2093. doi:10.1021/acs.jcim.5b00238
170. Duvenaud D, Maclaurin D, Iparraguirre J, et al. Convolutional Networks on Graphs for Learning Molecular Fingerprints. *Adv Neural Inf Process Syst*. 2015;28:2224–2232.

171. Ramsundar B, Liu B, Wu Z, et al. Is Multitask Deep Learning Practical for Pharma?. *J Chem Inf Model*. 2017;57(8):2068–2076. doi:10.1021/acs.jcim.7b00146
172. Ekins S, Puhl AC, Zorn KM, et al. Exploiting Machine Learning for End-to-End Drug Discovery and Development. *Nat Mater*. 2019;18(5):435–441. doi:10.1038/s41563-019-0338-z
173. Ribeiro JML, Bravo P, Wang Y, Tiwary P. Reweighted autoencoded variational Bayes for enhanced sampling (RAVE). *J Chem Phys*. 2018;149(7):072301. doi:10.1063/1.5025487
174. Noé F, Tkatchenko A, Müller KR, Clementi C. Machine learning for quantum mechanical properties of atoms in molecules. *J Chem Phys*. 2020;152(19):194104. doi:10.1063/1.5143268
175. Bycroft C, Freeman C, Petkova D, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562(7726):203–209. doi:10.1038/s41586-018-0579-z
176. Chen B, Butte AJ, Butte MJ. Data-Mining for Detecting Signals of Adverse Drug Reactions of Novel Drugs. *J Med Internet Res*. 2018;20(11):e11148.
177. Fumagalli SE, Padhiar NH, Meyer D, et al. Analysis of 3.5 million SARS-CoV-2 sequences reveals unique mutational trends with consistent nucleotide and codon frequencies. *Virol J*. 2023;20(1):31. doi:10.1186/s12985-023-01982-8
178. Weissler EH, Naumann T, Andersson T, et al. The role of machine learning in clinical research transforming the future of evidence generation. *Trials*. 2021;22(1):537. doi:10.1186/s13063-021-05489-x
179. Chan HCS, Shan H, Dahoun T, Vogel H, Yuan S. Advancing Drug Discovery via Artificial Intelligence. *Trends Pharmacol Sci*. 2019;40(8):592–604. doi:10.1016/j.tips.2019.06.004
180. Schneider P, Walters WP, Plowright AT, et al. Rethinking drug design in the artificial intelligence era. *Nat Rev Drug Discov*. 2020;19(5):353–364. doi:10.1038/s41573-019-0050-3
181. Liu Z, Chen X, Carter W, et al. AI- powered drug repurposing for developing COVID-19 treatments. *Reference Module Biomed Sci*. 2022. doi:10.1016/B978-0-12-824010-6.00005-8.
182. Mouchlis VD, Afantitis A, Serra A, et al. Advances in de Novo Drug Design From Conventional to Machine Learning Methods. *Int J Mol Sci*. 2021;22(4). doi:10.3390/ijms22041676
183. Jumper J, Evans R, Pritzel A, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583–589. doi:10.1038/s41586-021-03819-2
184. Levinthal C. How to Fold Graciously. Mossbauer Spectroscopy in Biological Systems: Proceedings of a Meeting Held at Allerton House, Monticello, Illinois. 1969. 22–24.
185. Corey RB, Kuntz ID. ENCEPP: a program for predicting the conformational geometry of organic molecules. *J Comput Chem*. 1974;2(3):287–303.
186. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. *J Mol Biol*. 1993;234(3):779–815. doi:10.1006/jmbi.1993.1626
187. Rose GD, Geselowitz AR, Lesser GJ, Lee RH, Zehfus MH. Hydrophobicity of amino acid residues in globular proteins. *Science*. 1985;229(4716):834–838. doi:10.1126/science.4023714
188. Levitt M, Warshel A. Computer simulation of protein folding. *Nature*. 1975;253(5494):694–698. doi:10.1038/253694a0
189. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. *Nat Struct Biol*. 2002;9(9):646–652.
190. ones DT, Taylor WR, Thornton JM. A new approach to protein fold recognition. *Nature*. 1992;358(6381):86–89. doi:10.1038/358086a0
191. Simons KT, Kooperberg C, Huang E, Baker D. Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions. *J Mol Biol*. 1997;268(1):209–225. doi:10.1006/jmbi.1997.0959
192. Montanucci L, Capriotti E, Frank Y, Ben-Tal N, Fariselli P. DDGun: an untrained method for the prediction of protein stability changes upon single and multiple point variations. *BMC Bioinform*. 2019;20(S14):S14. doi:10.1186/s12859-019-2923-1
193. Savojardo C, Fariselli P, Martelli PL, Casadio R. INPS-MD: a web server to predict stability of protein variants from sequence and structure. *Bioinformatics*. 2016;32(16):2542–2544. doi:10.1093/bioinformatics/btw192
194. Lv X, Chen J, Lu Y, Chen Z, Xiao N, Yang Y. Accurately predicting mutation-caused stability changes from protein sequences using extreme gradient boosting. *J Chem Inf Mod*. 2020;60(4):2388–2395. doi:10.1021/acs.jcim.0c00064
195. Bertoline LMF, Lima AN, Krieger JE, Teixeira SK. Before and after AlphaFold2: An overview of protein structure prediction. *Front Bioinform*. 2023;3:1120370. doi:10.3389/fbinf.2023.1120370
196. Milla´n C, Keegan RM, Pereira J, et al. Assessing the utility of CASP14 models for molecular replacement. *Proteins*. 2021;89(12):1752–1769. doi:10.1002/prot.26214
197. Available from: https://alphafold.ebi.ac.uk/faq. Accessed July 10, 2023.
198. Hopkins AL, Groom CR, Alex A, et al. Ligand efficiency: A useful metric for lead selection. *Drug Discov Today*. 2004;9(10):430–431. doi:10.1016/S1359-6446(04)03069-7
199. Kozakov D, Hall DR, Xia B, et al. The ClusPro web server for protein-protein docking. *Nat Protoc*. 2017;12(2):255. doi:10.1038/nprot.2016.169
200. Cheng T, Li Q, Zhou Z, et al. Structure-based virtual screening for drug discovery: a problem-centric review. *AAPS J*. 2012;14(1):133–141. doi:10.1208/s12248-012-9322-0.
201. Rao VS, Srinivas K, Sujini GN, et al. Protein-protein interaction detection: methods and analysis. *Int J Proteomics*. 2014;2014:147648. doi:10.1155/2014/147648.
202. Erlanson DA, Davis BJ, Jahnke W, et al. Fragment-Based Drug Discovery: Advancing Fragments in the Absence of Crystal Structures. *Cell Chem Biol*. 2016;26(1):9–15. doi:10.1016/j.chembiol.2018.10.001.
203. Metallo SJ. Intrinsically disordered proteins are potential drug targets. Curr Opin Chem Biol. 2010;14(4):481. doi:10.1016/j.cbpa.2010.06.169.
204. Macalino SJY, Gosu V, Hong S, et al. Role of computer-aided drug design in modern drug discovery. *Arch Pharm Res*. 2015;38(9):1686–1701. doi:10.1007/s12272-015-0640-5
205. Karplus M, Kuriyan J. Molecular dynamics and protein function. *Proce National Acad Sci*. 2005;102(19):6679–6685. doi:10.1073/pnas.0408930102
206. Teague SJ. Implications of protein flexibility for drug discovery. *Nat Rev Drug Discov*. 2003;2(7):527–541. doi:10.1038/nrd1129

207. Lexa KW, Carlson HA. Protein flexibility in docking and surface mapping. *Q Rev Biophys*. 2012;45(3):301. doi:10.1017/S0033583512000066
208. Macarron R, Banks MN, Bojanic D, et al. Impact of high-throughput screening in biomedical research. *Nat Rev Drug Discov*. 2011;10(3):188–195. doi:10.1038/nrd3368
209. Hansch C, Fujita T. ρ-σ-π Analysis. A Method for the Correlation of Biological Activity and Chemical Structure. *J Am Chem Soc*. 1964;86(8):1616–1626. doi:10.1021/ja01062a035
210. Kuroda D, et al. Structure-based drug design for combating influenza virus by targeting the PA-PB1 interaction. *Sci Rep*. 2012;2:1166.
211. Dror RO, Dirks RM, Grossman JP, et al. Biomolecular simulation: a computational microscope for molecular biology. *Annu Rev Biophys*. 2012;41(1):429–452. doi:10.1146/annurev-biophys-042910-155245
212. Shaw DE, Maragakis P, Lindorff-Larsen K, et al. Atomic-level characterization of the structural dynamics of proteins. *Science*. 2010;330(6002):341–346. doi:10.1126/science.1187409
213. Lock EF, Abdo N, Huang R, et al. Quantitative high-throughput screening for chemical toxicity in a population-based in vitro model. *Toxicol Sci*. 2012;126(2):578. doi:10.1093/toxsci/kfs023
214. Kühlbrandt W. Cryo-EM enters a new era. *eLife*. 2014;3:e03678. doi:10.7554/eLife.03678
215. Senn HM, Thiel W. QM/MM methods for biomolecular systems. *Angewandte Chemie*. 2009;48(7):1198–1229. doi:10.1002/anie.200802019
216. Fons N. Textbook of Drug Design and Discovery, Fifth Edition. *Yale J Biol Med*;2017;90:160.
217. Bender A, Glen RC. Molecular similarity: A key technique in molecular informatics. *Org Biomol Chem*. 2004;2(22):3204–3218.
218. Shoichet BK. Virtual screening of chemical libraries. *Nature*. 2004;432(7019):862–865. doi:10.1038/nature03197
219. Jensen LJ, Kuhn M, Stark M, et al. STRING 8—a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res*. 2006;37(suppl_1):D412–D416. doi:10.1093/nar/gkn760
220. Chipot C, Pohorille A, Eds. *Free Energy Calculations*. Springer Science & Business Media; 2007.
221. Fischman S, Ofran Y. Computational design of antibodies. *Curr Opin Struct Biol*. 2018;45:156–162. doi:10.1016/j.sbi.2018.04.007.
222. Bulitta JB, Hope WW, Eakin AE, et al. Generating Robust and Informative Nonclinical In Vitro and In Vivo Bacterial Infection Model Efficacy Data To Support Translation to Humans. *Antimicrob Agents Chemother*. 2019;63(5):e02307–02318. doi:10.1128/AAC.02307-18
223. Harrison LI, Gibaldi M. Physiologically based pharmacokinetic model for digoxin disposition in dogs and its preliminary application to humans. *J Pharm Sci*. 1977;66(12):1679–1683. doi:10.1002/jps.2600661206
224. Hsu V, de LTV, Zhao M, et al. Towards quantitation of the effects of renal impairment and probenecid inhibition on kidney uptake and efflux transporters, using physiologically based pharmacokinetic modelling and simulations. *Clin Pharmacokinet*. 2014;53(3):283–293. doi:10.1007/s40262-013-0117-y
225. Mager DE, Woo S, Jusko WJ. Scaling pharmacodynamics from in vitro and preclinical animal studies to humans. *Drug Metab Pharmacokinet*. 2009;24(1):16–24. doi:10.2133/dmpk.24.16
226. Shroff T, Aina K, Maass C, et al. Studying metabolism with multi-organ chips new tools for disease modelling, pharmacokinetics and pharmacodynamics. *Open Biol*. 2022;12(3):210333. doi:10.1098/rsob.210333
227. Daryaee F, Tonge PJ. Pharmacokinetic-pharmacodynamic models that incorporate drug-target binding kinetics. *Curr Opin Chem Biol*. 2019;50:120–127. doi:10.1016/j.cbpa.2019.03.008
228. Sager JE, Yu J, Ragueneau-Majlessi I, Isoherranen N. Physiologically based pharmacokinetic (PBPK) modeling and simulation approaches A systematic review of published models, applications, and model verification. *Drug Metab Dispos*. 2015;43(11):1823–1837. doi:10.1124/dmd.115.065920
229. Liu X, Liu C, Huang R, et al. Long short-term memory recurrent neural network for pharmacokinetic-pharmacodynamic modeling. *Int J Clin Pharmacol Ther*. 2021;59(2):138–146. doi:10.5414/CP203800
230. Available from: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/generally-accepted-scientific-knowledge-applications-drug-and-biological-products-nonclinical. Accessed July 10, 2023.
231. Liu Q, Huang R, Hsieh J, et al. Landscape analysis of the application of artificial intelligence and machine learning in regulatory submissions for drug development from 2016 to 2021. *Clin Pharmacol Ther*. 2022. doi:10.1002/cpt.2668
232. Wang Y, Liu S, Afzal N, et al. A Comparison of Word Embeddings for the Biomedical Natural Language Processing. *J Biomed Inform*. 2018;87:12–20. doi:10.1016/j.jbi.2018.09.008
233. Beam AL, Kohane IS. Big Data and Machine Learning in Health Care. *JAMA*. 2018;319(13):1317–1318. doi:10.1001/jama.2017.18391
234. Enticott J, Johnson A, Teede H, et al. Learning health systems using data to drive healthcare improvement and impact: a systematic review. *BMC Health Serv Res*. 2021;21(1):200. doi:10.1186/s12913-021-06215-8.
235. Demner-Fushman D, Elhadad N. Aspiring to Unintended Consequences of Natural Language Processing: a Review of Recent Developments in Clinical and Consumer-Generated Text Processing. *IMIA Yearbook Med Informatics*. 2016;25(01):224–233. doi:10.15265/IY-2016-017
236. Elki ME, Zhu X. Predictive modeling of clinical trial terminations using feature engineering and embedding learning. *Sci Rep*. 2021;11:3446. doi:10.1038/s41598-021-82840-x
237. Thall PF, Lee SJ. Practical model-based dose-finding in Phase I clinical trials: methods based on toxicity. *Int J Radiation Oncol Biol Phys*. 2015;61(3):691–698. doi:10.1046/j.1525-1438.2003.13202.x.
238. Dara S, Dhamercherla S, Jadav SS, et al. Machine Learning in Drug Discovery: A Review. *Artificial Intelligence Rev*. 2022;55(1):1947–1999. doi:10.1007/s10462-021-10058-4
239. Harrer S, Shah P, Antony B, Hu J. Artificial Intelligence for Clinical Trial Design. *Trends Pharmacol Sci*. 2019;40(8):577–591. doi:10.1016/j.tips.2019.05.005
240. Luo G, et al. Machine learning in patient recruitment and retention for clinical trials. *Trials*. 2020;21:934. doi:10.1186/s13063-020-04884-0
241. Aerts HJWL, Grossmann P, Tan Y, et al. Defining a Radiomic Response Phenotype A Pilot Study using targeted therapy in NSCLC. *Sci Rep*. 2016;6:33860. doi:10.1038/srep33860
242. Athreya AP, Neavin D, Carrillo-Roa T, et al. Pharmacogenomics-Driven Prediction of Antidepressant Treatment Outcomes A Machine-Learning Approach With Multi-trial Replication. *Clin Pharmacol Ther*. 2019;106(4):855–865. doi:10.1002/cpt.1482
243. Dercle L, Fronheiser M, Lu L, et al. Identification of non-small cell lung cancer sensitive to systemic cancer therapies using radiomics. *Clin Cancer Res*. 2020;26(9):2151–2162. doi:10.1158/1078-0432.CCR-19-2942

244. Kawakami E, Tabata J, Yanaihara N, et al. Application of artificial intelligence for preoperative diagnostic and prognostic prediction in epithelial ovarian cancer based on blood biomarkers. *Clin Cancer Res*. 2019;25(10):3006–3015. doi:10.1158/1078-0432.CCR-18-3378
245. Wang Y, Carter BZ, Li Z, Huang X. Application of machine learning methods in clinical trials for precision medicine. *JAMIA Open*. 2022;5(1): ooab107. doi:10.1093/jamiaopen/ooab107
246. Luo J, et al. Automatically explaining machine learning prediction results: a demonstration on type 2 diabetes risk prediction. *Health Inform Sci Systems*. 2016;4:2. doi:10.1186/s13755-016-0015-4.
247. Char DS, Shah NH, Magnus D. Implementing Machine Learning in Health Care — Addressing Ethical Challenges. *N Eng J Med*. 2018;378 (11):981–983. doi:10.1056/NEJMp1714229.
248. Pavelko RL, Myrick JG, Verghese RS, et al. Public reactions to celebrity cancer disclosures via social media: Implications for campaign message design and strategy. *Health Educ J*. 2017;76:492–506, doi:10.1177/0017896917696122.
249. Lu J, Deng K, Zhang X, Liu G, Guan Y. Neural-ODE for pharmacokinetics modeling and its advantage to alternative machine learning models in predicting new dosing regimens. *iScience*. 2021;24(7):102804. doi:10.1016/j.isci.2021.102804
250. Sheikh A, Cornford T, Barber N, et al. Implementation and adoption of nationwide electronic health records in secondary care in England: final qualitative results from prospective national evaluation in "early adopter" hospitals. *BMJ*. 2016;343(oct17 1):d6054. doi:10.1136/bmj.d6054
251. Mason M, Cho Y, Rayo J, Gong Y, Harris M, Jiang Y. Technologies for drug adherence monitoring and technology assessment criteria Narrative review. *JMIR Mhealth Uhealth*. 2022;10(3):e35157. doi:10.2196/35157
252. Angermueller C, Pärnamaa T, Parts L, et al. Deep learning for computational biology. *Mol Syst Biol*. 2016;12(7):878. doi:10.15252/ msb.20156651
253. Dunn J, Runge R, Snyder M, et al. Wearables and the medical revolution. *Per Med*. 2018;15(5):429–448. doi:10.2217/pme-2018-0044
254. Labovitz DL, Shafner L, Reyes Gil M, et al. Using Artificial Intelligence to Reduce the Risk of Nonadherence in Patients on Anticoagulation Therapy. *Stroke*. 2017;48(5):1416–1419. doi:10.1161/STROKEAHA.116.016281.
255. The draft guidance for industry, investigators, and other stakeholders. Digital Health Technologies for Remote Data Acquisition in Clinical Investigations; 2021. When final, this guidance will represent FDA's current thinking on this topic. Available from: https://www.fda.gov/media/ 155022/download. Accessed July 10, 2023.
256. Stehlik J, Schmalfuss C, Bozkurt B, et al. Continuous Wearable Monitoring Analytics Predict Heart Failure Hospitalization The LINK-HF Multicenter Study. *Circ Heart Fail*. 2020;13(3):e006513. doi:10.1161/circheartfailure.119.006513
257. Cohoon TJ, Bhavnani SP. Toward precision health applying artificial intelligence analytics to digital health biometric datasets. *Per Med*. 2020;17(4):307–316. doi:10.2217/pme-2019-0113
258. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60–88. doi:10.1016/j. media.2017.07.005
259. Zhang X, Yan C, Gao C, Malin BA, Chen Y. Predicting Missing Values in Medical Data via XGBoost Regression. *J Healthc Inform Res*. 2020;4 (4):383–394. doi:10.1007/s41666-020-00077-1
260. Pappalardo F, Russo G, Tshinanu FM, Viceconti M. In silico clinical trials concepts and early adoptions. *Brief Bioinform*. 2019;20 (5):1699–1708. doi:10.1093/bib/bby043
261. European Drugs Agency. DRAFT Qualification opinion for Prognostic Covariate Adjustment (PROCOVA™). European Drugs Agency; 2022. Available from: https://www.ema.europa.eu/documents/other/draft-qualification-opinion-prognostic-covariate-adjustment-procovatm_en.pdf. Accessed July 10, 2023.
262. Laubenbacher R, Sluka JP, Glazier JA. Using digital twins in viral infection. *Science*. 2021;371(6534):1105–1106. doi:10.1126/science.abf3370
263. Schuler A, Walsh D, Hall D, Walsh J, Fisher C. Increasing the efficiency of randomized trial estimates via linear adjustment for a prognostic score. *Int J Biostat*. 2021. doi:10.1515/ijb-2021-0072
264. Kotsiantis SB. Supervised machine learning: A review of classification techniques. *Informatica*. 2007;31:249–268.
265. Weng SF, Bolton T, Di Angelantonio E, et al. Can machine-learning improve cardiovascular risk prediction using routine clinical data?. *PLoS One*. 2019;14(4):e0213653. doi:10.1371/journal.pone.0213653
266. Pierce CE, Bouri K, Pamer C, et al. Evaluation of Facebook and Twitter monitoring to detect safety signals for medical products. An analysis of recent FDA safety alerts. *Drug Saf*. 2017;40(4):317–331. doi:10.1007/s40264-016-0491-0
267. Routray R, Tetarenko N, Abu-Assal C, et al. Application of augmented intelligence for pharmacovigilance case seriousness determination. *Drug Saf*. 2020;43(1):57–66. doi:10.1007/s40264-019-00869-4
268. Bate A, Hobbiger SF. Artificial Intelligence, Real-World Automation and the Safety of Drugs. *Drug Saf*. 2021;44(2):125–132. doi:10.1007/ s40264-020-01001-7
269. Sarker A, et al. A corpus for large-scale machine learning and microservices to generate daily Adverse Drug Event counts from Twitter. *J Am Med Informatics Assoc*. 2020;27(4):476–487.
270. Comfort S, Dorrell D, Meireis S, Fine J. MOdified NARanjo Causality Scale for ICSRs (MONARCSi) A Decision Support Tool for Safety Scientists. *Drug Saf*. 2018;41(11):1073–1085. doi:10.1007/s40264-018-0690-y
271. Negi K, Pavuri A, Patel L, Jain C. A novel method for drug-adverse event extraction using machine learning. *Inform Med Unlocked*. 2019;17. doi:10.1016/j.imu.2019.100190
272. Wang SV, Schneeweiss S, Berger ML, Brown J, de Vries F, Douglas I. Reporting to improve reproducibility and facilitate validity assessment for healthcare database studies V1.0. *Pharmacoepidemiol Drug Saf*. 2017;26(9):1018–1032. doi:10.1002/pds.4295
273. Abatemarco D, Perera S, Bao SH, et al. Training Augmented Intelligent Capabilities for Pharmacovigilance Applying Deep-learning Approaches to Individual Case Safety Report Processing. *Pharmaceut Med*. 2018;32(6):391–401. doi:10.1007/s40290-018-0251-9
274. Schmider J, Kumar K, LaForest C, Swankoski B, Naim K, Caubel PM. Innovation in pharmacovigilance. Use of artificial intelligence in adverse event case processing. *Clin Pharmacol Ther*. 2019;105(4):954–961. doi:10.1002/cpt.1255
275. Kassekert R, Grabowski N, Lorenz D, et al. Industry Perspective on Artificial Intelligence/Machine Learning in Pharmacovigilance. *Drug Saf*. 2022;45(5):439–448. doi:10.1007/s40264-022-01164-5
276. Ghosh R, Kempf D, Pufko A, Barrios Martinez LF, Davis CM, Sethi S. Automation opportunities in pharmacovigilance An industry survey. *Pharmaceut Med*. 2020;34(1):7–18. doi:10.1007/s40290-019-00320-0

277. Lewis DJ, McCallum JF. Utilizing advanced technologies to augment pharmacovigilance systems Challenges and opportunities. *Ther Innov Regul Sci*. 2020;54(4):888–899. doi:10.1007/s43441-019-00023-3
278. FDA Emerging Technology Program. Available from: https://www.fda.gov/about-fda/center-drug-evaluation-andresearch-cder/emerging-technology-program. Accessed July 10, 2023.
279. IBM: What is a Digital Twin. Available from: Https//www.ibm.com/topics/what-is-A-digital-twin. Accessed July 10, 2023.
280. Huang J, O'Connor T, Ahmed K, et al. AIChE PD2M Advanced Process Control workshop-moving APC forward in the pharmaceutical industry. *J Adv Manufacturing Processing*. 2021;3(1):e10071. doi:10.1002/amp2.10071
281. National Academies of Sciences, E., and Medicine. *Innovations in Pharmaceutical Manufacturing on the Horizon Technical Challenges, Regulatory Issues, and Recommendations*. National Academies of Sciences, E; 2021.
282. Arden NS, Fisher AC, Tyner K, Yu LX, Lee SL, Kopcha M. Industry 4.0 for pharmaceutical manufacturing. Preparing for the smart factories of the future. *Int J Pharm*. 2021;602:120554. doi:10.1016/j.ijpharm.2021.120554
283. Promoting Innovation in Medical Product Assessment. A Risk-based Framework for Evaluating Computational Models for Regulatory Decision-Making; 2020. Available from: https://www.fda.gov/drugs/news-events-human-drugs/promoting-innovation-medical-product-assessment-risk-based-framework-evaluating-computational-models. Accessed July 10, 2023.
284. A V&V 70 Subcommittee has been established for Verification and Validation of Machine Learning.
285. The draft guidance for industry and FDA staff. Assessing the Credibility of Computational Modelling Simulation in Medical Device Submissions; 2021. When final, this guidance will represent FDA's current thinking on this topic. Available from: https://www.fda.gov/media/154985/download. Accessed July 10, 2023.
286. Good Machine Learning Practice for Medical Device Development Guiding Principles; 2021. Available from: https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles. Accessed July 10, 2023.
287. Exec. Order No. 13859, Maintaining American Leadership in Artificial Intelligence. (2019- 02544). (February 11, 2019). 84 Fed. Reg. 3967 Available from: https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence. Accessed July 10, 2023.
288. Exec. Order No. 13960, Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government. (2020-27065). (December 3, 2020). 85 Fed. Reg. 78939. Available from: https://www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-The-use-of-trustworthy-artificial-intelligence-in-The-federal-government. Accessed July 10, 2023.
289. Lander E, Nelson A. ICYMI WIRED (Opinion) Americans Need a Bill of Rights for an AI-Powered World; 2021. Available from: https://www.whitehouse.gov/ostp/news-updates/2021/10/22/icymi-wired-opinion-americans-need-A-bill-of-rights-for-an-ai-powered-world/. Accessed July 10, 2023.
290. Proposed Regulatory Framework for Modifications to Artificial Intelligence/Machine Learning (AI/ML)- Based Software as a Medical Device (SaMD) – Discussion Paper and Request for Feedback; 2019. Available from: https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf. Accessed July 10, 2023.
291. the Public Workshop – evolving Role of Artificial Intelligence in Radiological Imaging; 2020. Available from: https://www.fda.gov/medical-devices/workshops-conferences-medical-devices/public-workshop-evolving-role-artificial-intelligence-radiological-imaging-02252020-02262020. Accessed July 10, 2023.
292. the Virtual Public Workshop – Transparency of Artificial Intelligence/Machine Learning-enabled Medical Devices; 2021. Available from: https://www.fda.gov/medical-devices/workshops-conferences-medical-devices/virtual-public-workshop-transparency-artificial-intelligencemachine-learning-enabled-medical-devices. Accessed July 10, 2023.
293. The guidance for industry and FDA. staff Qualification Process for Drug Development Tools; 2020. Available from: https://www.fda.gov/media/133511/download. Accessed July 10, 2023.
294. the Model-Informed Drug Development Paired Meeting Program; 2022. Available from: https://www.fda.gov/drugs/development-resources/model-informed-drug-development-pilot-program. Accessed July 10, 2023.
295. FDA's Sentinel Initiative; 2022. Available from: https://www.fda.gov/safety/fdas-sentinel-initiative. Accessed July 10, 2023.
296. The CBER Biologics Effectiveness and Safety (BEST) System; 2022. Available from: https://www.fda.gov/vaccines-blood-biologics/safety-availability-biologics/cber-biologics-effectiveness-. Accessed July 10, 2023.
297. the National Evaluation System for health Technology (NEST); 2019. Available from: https://www.fda.gov/about-fda/cdrh-reports/national-evaluation-system-health-technology-nest. Accessed July 10, 2023.
298. The FDA Sentinel System Five-Year Strategy; 2019. Available from: https://www.fda.gov/media/120333/download. Accessed July 10, 2023.
299. Desai RJ, Matheny ME, Johnson K, et al. Broadening the reach of the FDA Sentinel system. A roadmap for integrating electronic health record data in a causal analysis framework. *NPJ Digit Med*. 2021;4(1):170. doi:10.1038/s41746-021-00542-0
300. Balzer LB, Zheng W, van der Laan MJ, Petersen ML. A new approach to hierarchical data analysis. Targeted maximum likelihood estimation for the causal effect of a cluster-level exposure. *Stat Methods Med Res*. 2019;28(6):1761–1780. doi:10.1177/0962280218774936
301. the CBER BEST System; 2022. Available from: https://www.fda.gov/vaccines-blood-biologics/safety-availability-biologics/cber-biologics-effectiveness-and-safety-best-system. Accessed July 10, 2023.
302. Ball R, Dal Pan G. Artificial Intelligence for Pharmacovigilance Ready for Prime Time?. *Drug Saf*. 2022;45(5):429–438. doi:10.1007/s40264-022-01157-4
303. Kreimeyer K, Dang O, Spiker J, et al. Increased Confidence in Deduplication of Drug Safety Reports with Natural Language Processing of Narratives at the US Food and Drug Administration. *Front Drug Safety Regulation*. 2022;2. doi:10.3389/fdsfr.2022.918897
304. Kreimeyer K, Dang O, Spiker J, et al. Feature engineering and machine learning for causality assessment in pharmacovigilance Lessons learned from application to the FDA Adverse Event Reporting System. *Comput Biol Med*. 2021;135:104517. doi:10.1016/j.compbiomed.2021.104517
305. Spiker J, Kreimeyer K, Dang O, et al. Information visualization platform for postmarket surveillance decision support. *Drug Saf*. 2020;43(9):905–915. doi:10.1007/s40264-020-00945-0
306. Bayer S, Clark C, Dang O, et al. An Evaluation of Text Processing Systems for Adverse Event Extraction from Drug Labels for Pharmacovigilance. *Drug Saf*. 2021;44(1):83–94. doi:10.1007/s40264-020-00996-3

307. Ly T, Pamer C, Dang O, et al. Evaluation of Natural Language Processing (NLP) systems to annotate drug product labeling with MedDRA terminology. *J Biomed Inform*. 2018;83:73–86. doi:10.1016/j.jbi.2018.05.019
308. The Quality Metrics for Drug Manufacturing; 2022. Available from: https://www.fda.gov/drugs/pharmaceutical-quality-resources/quality-metrics-drug-manufacturing. Accessed July 10, 2023.
309. Available from: https://www.fda.gov/science-research/advancing-regulatory-science/regulatory-science-extramural-research-and-development-projects. Accessed July 10, 2023.
310. CPIM; 2022. Available from: https://www.fda.gov/drugs/new-drugs-fda-cders-new-molecular-entities-and-new-therapeutic-biological-products/critical-path-innovation-meetings-cpim. Accessed July 10, 2023.
311. the ISTAND Pilot Program; 2021. Available from: https://www.fda.gov/drugs/drug-development-tool-ddt-qualification-programs/innovative-science-and-technology-approaches-new-drugs-istand-pilot-program. Accessed July 10, 2023.
312. Emerging Technology Program; 2022. Available from: https://www.fda.gov/about-fda/center-drug-evaluation-and-research-cder/emerging-technology-program. Accessed July 10, 2023.
313. Framework for FDA's Real World Evidence Program; 2020. Available from: Https:/fda.gov/media/120060/download. Accessed July 10, 2023.
314. Available from: https://www.fda.gov/media/75414/download. Accessed July 10, 2023.