FDA U.S. FOOD & DRUG ADMINISTRATION

Using Artificial Intelligence & Machine Learning in the Development of Drug & Biological Products

Discussion Paper and Request for Feedback

May 2023
(Revised February 2025)

Table of Contents

I. Background and Scope

To fulfill its mission of protecting, promoting, and advancing public health, the Food and Drug Administration's (FDA's) Center for Drug Evaluation and Research (CDER), in collaboration with the Center for Biologics Evaluation and Research (CBER) and the Center for Devices and Radiological Health (CDRH), including the Digital Health Center of Excellence (DHCoE), is publishing this document to facilitate a discussion with stakeholders on the use of ***artificial intelligence*** (***AI***)[1] and ***machine learning*** (***ML***)[2] in drug development,[3,4] including in the development of medical devices intended to be used with drugs, to help inform the regulatory landscape in this area.

FDA helps to ensure that drugs are safe and effective while facilitating innovations in their development. Recent, rapid technological innovations in data collection and generation tools, combined with robust information management and exchange systems and advanced computing abilities, may transform the way drugs are developed and used (ElZarrad, Lee, Purcell, & Steele, 2022). This evolving ecosystem presents unique opportunities and challenges, and FDA is committed to working across its medical product centers with partners domestically and internationally to ensure that the full potential of these innovations is realized for the benefit of the public.

Developers, manufacturers, regulators, academic groups, and other stakeholders are working to develop a shared understanding of where and how specific innovations, such as AI and ML, can best be used throughout the drug development process. FDA is publishing this discussion paper as part of a multifaceted approach to enhance mutual learning and to establish a dialogue with FDA stakeholders on this topic. AI can generally be described as a branch of computer science, statistics, and engineering that uses algorithms or models to perform tasks and exhibit behaviors such as learning, making decisions, and making predictions.[5] ML is considered a subset of AI that allows ML models to be developed by ML training algorithms through analysis of data, without models being explicitly programmed.[6] Additionally, there are a variety of ML methods and different types of algorithms that may be utilized in a given context. For purposes of this document, AI and ML will be referenced together as AI/ML, and references to

[1] Words and phrases in ***bold italics*** are defined in the Glossary.
[2] There are multiple definitions for AI and ML, and the Glossary includes several definitions from federal legislation and agencies.
[3] For purposes of this discussion paper, all references to *drug* or *drugs* include both human drugs and biological products.
[4] FDA is focusing this discussion paper on drug development. However, many of the AI/ML scientific and regulatory science principles outlined in this document may be applicable across all medical products, including in the development of medical devices intended to be used with drugs (including, but not limited to, combination products, companion devices, and complementary devices). Some medical devices intended to be used with drugs are intended for use only in clinical investigations; others are intended to be marketed for use outside of clinical investigations. Examples include medical devices that help identify side effects of drugs as well as medical devices that assist in drug dosing.
[5] See IMDRF/AIMD WG/N67 Machine Learning-enabled Medical Devices: Key Terms and Definitions, final document, May 6, 2022. https://www.imdrf.org/documents/machine-learning-enabled-medical-devices-key-terms-and-definitions
[6] *Ibid.*

drug development and the drug development process include a wide scope of activities and phases, including manufacturing and postmarket drug safety monitoring, among others.[7,8]

This discussion paper, which considers the application of AI/ML in the broad context of the drug development process, is not FDA guidance or policy and does not endorse a specific AI/ML use or approach in drug development. Rather, this discussion paper is an initial communication with stakeholders, including academic groups, researchers, and technology developers, that is intended to promote mutual learning and discussion. It is particularly beneficial for those new to drug development and human subjects research, to recognize some of the initial thinking and considerations involved with utilizing these technologies, including having familiarity with FDA's current activities, initiatives, practices, and potentially applicable regulations. FDA is soliciting feedback on the opportunities and challenges with utilizing AI/ML in the development of drugs, as well as in the development of medical devices intended to be used with drugs. This feedback will provide an additional resource to help inform the regulatory landscape in this area.

In this discussion paper, three main topics are discussed:

- **Landscape of current and potential uses of AI/ML**: FDA recognizes the potential for AI/ML to enhance drug development in many ways, including to help bring safe and effective drugs to patients faster; provide broader access to drugs and thereby improve health; increase the quality of manufacturing; enhance drug safety; and develop novel drugs and drug classes, as well as personalized treatment approaches. Section II provides examples of the use of AI/ML to highlight the potential impact of AI/ML across the drug development process and includes a brief description of FDA's experience with AI/ML in drug development. The list of examples in this section is not comprehensive of all AI/ML uses, and it includes uses where FDA oversight may or may not be applicable. The purpose of this section is to promote shared learning and to identify areas where future regulatory clarity may be helpful.

- **Considerations for the use of AI/ML**: FDA is also aware of the potential concerns and risks with emerging innovations such as AI/ML and will share initial considerations and solicit feedback on how to help ensure the responsible utilization of AI/ML in drug development. Section III briefly describes several key efforts to develop general principles, standards, and practices for the use of AI/ML across diverse applications and then explores the principles and considerations that may be particularly applicable when using AI/ML for drug development activities. FDA understands that AI/ML use in drug development is

[7] See The Drug Development Process, January 2018. https://www.fda.gov/patients/learn-about-drug-and-device-approvals/drug-development-process
[8] In this discussion paper, the topic of clinical investigations focuses on the drug development process, however, many other activities and phases included as part of the drug development process may also be part of the development process for other medical products; see footnote 4.

diverse, and careful assessments that consider the specific ***context of use*** are needed. Taking a risk-based approach to evaluate and manage the use of AI/ML can help facilitate innovations and protect public health.

- **Next steps and stakeholder engagement**: FDA is interested in mutual opportunities to learn and engage with all stakeholders to establish a shared understanding of AI/ML systems and their rapidly evolving potential uses and considerations in drug development. As part of this ongoing effort, FDA welcomes feedback on this discussion paper and any AI/ML-related issues pertaining to drug development. Specifically, to initiate a broader dialogue with stakeholders, Section III includes several key questions to which interested parties can provide perspectives and Section IV outlines opportunities for future engagement.

II. Current and Potential Uses of AI/ML in the Drug Development Process

This section provides a high-level overview of the diverse and evolving uses of AI/ML being employed throughout the drug development process. These examples are not comprehensive of all AI/ML uses and include uses where FDA oversight may or may not be applicable.[9] Additionally, while some of the uses of AI/ML described in this section may also have utility in clinical practice, this paper is focused on uses of AI/ML in the drug development process. The purpose of this section is to promote shared learning and to identify areas where future FDA regulatory clarity may be beneficial.

Although the overall drug development process is an iterative continuum of activities and not strictly linear in nature, for simplicity, this section utilizes different phases of drug development to highlight several uses of AI/ML, ranging from drug discovery and clinical research to postmarket safety surveillance and advanced pharmaceutical manufacturing. The section also includes references to how AI/ML is being applied to ***real-world data*** (***RWD***) and data from ***digital health technologies*** (***DHTs***) in support of drug development. Some of the general challenges and considerations with utilizing AI/ML in different drug development use cases are discussed in **Section III**.

A. Drug Discovery

Early drug discovery is one of the areas with significant interest and activity in utilizing AI/ML. Included below is a brief discussion of the current and potential uses of AI/ML for drug target identification, selection, and prioritization, as well as compound screening and drug design in drug discovery.

[9] The examples listed were not necessarily submitted to FDA for review or approval and are not meant to suggest an endorsement of any specific approach. The FDA does not endorse any particular use of AI/ML.

1. Drug Target Identification, Selection, and Prioritization

The early stages of drug development generally rely on the initial identification of a suitable biological target for drug candidates. As a starting point, the process of identifying biological targets and elucidating disease relationships can utilize AI/ML to analyze and synthesize significant amounts of information from existing scientific research, publications, and other data sources. The growth of available genomic, transcriptomic, proteomic, and other data sources from healthy persons and those with a specific disease of interest provide a significant opportunity to inform biological target selection. These datasets are often complex and originate from disparate sources, which can be well-suited for the utilization of AI/ML approaches (Fumagalli et al., 2023). Building from existing validated data, AI/ML can be applied to mine and analyze these large multi-omics and other datasets to provide information on the potential structure and function of biological targets to predict their role in a disease pathway (Vamathevan et al., 2019; Weissler et al., 2021). While early target identification and prioritization is a critical step where AI/ML could help improve the efficiency and effectiveness of drug development, it is important to validate the role of the biological target in the disease of interest through subsequent studies (Fumagalli et al., 2023).

2. Compound Screening and Design

The discovery of potential drug candidates that modify the function of the identified biological targets of interest generally involves significant *in silico* or experimental screening of compound libraries, followed by subsequent refinement of a compound's specificity and selectivity for the biological target. In the area of compound screening, potential AI/ML uses include predicting the chemical properties and bioactivity of compounds and predicting efficacy and potential adverse events based on the compound's specificity and affinity for a target (Chan, Shan, Dahoun, Vogel, & Yuan, 2019; Schneider et al., 2020).

AI/ML approaches used to further elucidate drug-target interactions could also help provide predictions about classes of drugs potentially interacting with the same targets or having a similar mechanism of action, which may help predict the toxicity of a molecule based on specific known features. This strategy can help guide drug repurposing efforts that could utilize previously characterized compounds. Drug repurposing efforts utilizing AI/ML can also potentially benefit from the increased availability of suitable RWD from a variety of sources (e.g., electronic health records (EHRs), registries, and DHTs) to identify previously unknown effects of drugs on disease pathways (Z. Liu et al., 2022).

Finally, AI/ML could accelerate advances in *de novo* drug design (Mouchlis et al., 2021). For example, AI/ML may be applied to help predict the 3D structure of target proteins, informing chemical synthesis and the potential effect of a drug candidate on the target, including predicting affinity and potential toxicity (Chan et al., 2019; Jumper et al., 2021; Vamathevan et al., 2019). It is worth noting that one must be cautious with the use of AI/ML in 3-D structure prediction, as many proteins that are developed for pharmaceutical applications are codon optimized (with many synonymous mutations incorporated), the impact of which on protein structure is still an area of active research (Fumagalli et al., 2023; Jumper et al., 2021).

B. Nonclinical Research

Nonclinical research refers to *in vitro* and *in vivo* studies and is designed to further advance potential therapeutics towards clinical research in humans. Nonclinical studies, in support of new drug development, can be conducted at all phases of development: prior to clinical studies, in parallel with clinical development, and even in postmarketing environments. Data from pharmacokinetic, pharmacodynamic, and toxicologic studies conducted in animals; exploratory *in vitro* and *in vivo* mechanistic studies conducted in animal models; organ-on-chip and multi-organ chip systems; and cell assay platforms may be leveraged using AI/ML (e.g., computational modeling and simulation techniques) for evaluating toxicity, exploring mechanistic models, and developing *in vivo* predictive models (Bulitta et al., 2019; Harrison & Gibaldi, 1977; Hsu et al., 2014; Mager, Woo, & Jusko, 2009; Shroff et al., 2022).

Pharmacokinetics (PK) describes the time course of drug absorption, distribution, metabolism, and excretion. Pharmacodynamics (PD) explores the body's biological response to drugs. When PK and PD are integrated in a model, the model can describe how the drug effect will change with time when a certain dose or dosing regimen is used. Pharmacokinetic/pharmacodynamic (PK/PD) modeling has been used in drug development for decades and can be applied at both the nonclinical and clinical stages (Daryaee & Tonge, 2019). Along with the advances in computational tools and technology and the availability of modeling platforms, use of physiologically-based pharmacokinetic (PBPK) and physiologically-based PK/PD (PBPK-PD) modeling is also increasing (Sager, Yu, Ragueneau-Majlessi, & Isoherranen, 2015). There are current efforts to explore the use of more novel AI/ML algorithms (e.g., artificial ***neural network*** models and tree-based models) for PK/PD modeling. For example, a ***recurrent neural network****,* an ML algorithm commonly used for analyzing time series data, may be used to complement traditional PK/PD models in the area of highly complex PK/PD data analysis, and possibly lead to improved ***accuracy*** for nonclinical and clinical applications (Liu et al., 2021).

C. Clinical Research

Clinical research typically involves a series of phases of clinical trials in increasing numbers of human subjects to assess the safety and effectiveness of a drug. One of the most significant applications of AI/ML in drug development is in efforts to streamline and advance clinical research. For example, AI/ML is being utilized to analyze vast amounts of data from both interventional studies (also referred to as clinical trials) and non-interventional studies (also referred to as observational studies) to make inferences regarding the safety and effectiveness of a drug. Additionally, AI/ML has the potential to inform the design and efficiency of non-traditional trials such as ***decentralized clinical trials***, and trials incorporating the use of RWD extracted from EHRs, medical claims, or other data sources. AI/ML may also have a role in analyzing and interpreting data collected from DHTs used in clinical studies. Finally, AI/ML could also be used to improve the conduct of clinical trials and augment operational efficiency. The following subsections will highlight some of the uses and potential uses of AI/ML during the design and conduct of clinical research.

1. Recruitment

AI/ML is increasingly being developed and used to connect individuals to trials for investigational treatments from which participants may benefit. Specifically, AI/ML is being used to mine vast amounts of data, such as data from clinical trial databases, trial announcements, social media, medical literature, registries, and structured and unstructured data in EHRs, which can be used to match individuals to trials (Harrer, Shah, Antony, & Hu, 2019). While these algorithms are trained on high volumes of patient data and enrollment criteria from past trials, it is important to ensure adequate representation of populations that are likely to use the drug. In the future, these technologies, if properly validated, may continue to play an increasing role in matching individuals with investigational treatments.

2. Selection and Stratification of Trial Participants

Enrichment strategies can aid participant selection in clinical investigations designed to demonstrate the effectiveness of drug and biological products.[10] AI/ML has been explored and used as part of a clinical investigation in the prediction of an individual participant's clinical outcome based on baseline characteristics (e.g., demographic information, clinical data, vital signs, labs, medical imaging data, and genomic data) (Aerts et al., 2016; Athreya et al., 2019; Dercle et al., 2020; Harrer et al., 2019; Kawakami et al., 2019). Such predictive models can be used to enrich clinical trials (e.g., identifying high-risk participants or participants more likely to respond to the treatment). When these types of AI/ML algorithms are used for patient evaluation and selection before randomization, it may be possible to reduce variability and increase study power (Y. Wang, Carter, Li, & Huang, 2022).

In addition to utilization in enrichment strategies, such predictive models can also be used for participant stratification, for example, if an AI/ML model could predict the probability of a serious adverse event before an investigational treatment is administered. Based on their predicted risk for these serious adverse events, participants can be stratified into different groups and then monitored accordingly (or excluded depending on predicted severity of the adverse event).

---

[10] See the guidance for industry *Enrichment Strategies for Clinical Trials to Support Determination of Effectiveness of Human Drugs and Biological Products* (March 2019). https://www.fda.gov/media/121320/download

3. Dose/Dosing Regimen Optimization

AI/ML can be used to characterize and predict PK profiles after drug administration. It can also be used to study the relationship between drug exposure and response, taking into consideration confounding factors. These kinds of models can be used to optimize the dose/dosing regimen selection for a study (Liu et al., 2021; Lu, Deng, Zhang, Liu, & Guan, 2021). This could potentially include aiding in dose optimization in special populations where there may be limited data (e.g., rare disease studies, pediatric and pregnant populations).

4. Adherence

AI/ML can be used to monitor and improve adherence during a clinical trial through tools, such as smartphone alerts and reminders, eTracking of medication (e.g., smart pillboxes and tools for visual confirmation) (Mason et al., 2022), and eTracking of missed clinical visits, which trigger non-adherence alerts. Examples of AI/ML used in clinical research to improve medication adherence include applications using digital biomarkers, such as facial and vocal expressivity, to monitor adherence remotely.

5. Retention

AI/ML has the potential to improve the participants' access to relevant trial information by enabling tools, such as AI chatbots, voice assistance, and intelligent search. AI/ML can also be used to reduce the burden for participants by using passive data collection techniques and by extracting more information from available data generated during clinical practice or by study activities (Weissler et al., 2021). Additionally, data from DHTs and other systems can be used to develop patient profiles to potentially predict dropouts and adverse events to ensure participant retention.

6. Site Selection

Trial operational conduct could also be optimized by utilizing AI/ML to help identify which sites have the greatest potential for a successful trial and to aid sites in identifying process gaps. For example, algorithms can be used to evaluate site performance and to help determine which sites may have a higher risk of running behind schedule based on data from other trials at that site.

7. Clinical Trial Data Collection, Management, and Analysis

a. Data Collection

DHTs, such as wireless and smartphone-connected products, wearables, implantables, and ingestibles, are increasingly being used in clinical trials to collect objective, quantifiable, longitudinal, and continuous physiological data.[11] In addition, many of these DHTs enable the use of AI/ML, either as embedded algorithms within the DHT or employed upon the data generated after the data are collected from the DHT, and have been used to predict the status of a chronic disease and its response to treatment (Stehlik et al., 2020) or to identify novel characteristics of an underlying condition (Avram et al., 2020). AI/ML can be utilized to analyze the large and diverse data generated from the continuous monitoring of persons using these technologies. This could include using AI/ML to aid in the evaluation of multimodal data and composite measures that may combine individual measures collected through multiple DHTs (Cohoon & Bhavnani, 2020).

b. Data Management

AI/ML can be used for a range of data cleaning and curation purposes, including duplicate participant detection and imputation of missing data values (Zhang, Yan, Gao, Malin, & Chen, 2020), as well as the ability to harmonize ***controlled terminology*** across drug development programs. Use of AI/ML could also significantly enhance data integration efforts by using supervised and unsupervised learning to help integrate data submitted in various formats and perform data quality assessments. Additionally, AI/ML can be used for data curation via masking and de-identification of personal identifiable information, metadata creation, and search and retrieval of stored data. These applications can potentially increase data accuracy and improve the speed at which data are prepared for analyses.

c. Data Analysis

AI/ML has been used to analyze high volumes of diverse and complex RWD extracted from EHRs, medical claims, and disease registries, among other sources. Additionally, the use of AI/ML in predictive modeling and counterfactual simulation to inform clinical trial designs is being actively explored. For example, *in silico* clinical trials utilize computational modeling and simulation to evaluate drug candidates using a virtual cohort of simulated participants with realistic variability of traits representing the desired participant population (Pappalardo, Russo, Tshinanu, & Viceconti, 2019). AI/ML could be employed in these situations to aid in evaluating a vast number of counterfactual simulations and to predict trial outcomes before human trials.

At an even more personalized level, AI/ML can also be used in the context of digital twins of patients, an emerging method that could potentially be used in clinical research. To create digital twins of patients, AI/ML can be utilized to build *in silico* representations or replicas of an individual that can dynamically reflect molecular and physiological status over time (European Medicines Agency, 2022; Laubenbacher, Sluka, & Glazier, 2021; Schuler et al., 2021). In comparison to a participant in a clinical trial that received an investigational treatment, the digital twin could potentially provide a comprehensive,

[11] See the draft guidance for industry, investigators, and other stakeholders *Digital Health Technologies for Remote Data Acquisition in Clinical Investigations* (December 2021). When final, this guidance will represent FDA's current thinking on this topic. https://www.fda.gov/media/155022/download

longitudinal, and computationally generated clinical record that describes what may have happened to that specific participant if they had received a placebo.

8. Clinical Endpoint Assessment

Clinical ***endpoint*** assessment is a key part of evaluating safety and efficacy of medical interventions in clinical trials. AI/ML-enabled algorithms could detect clusters of signs and symptoms to identify a potential safety signal, as well as help detect cases with safety issues in real time (Pierce et al., 2017; Routray et al., 2020). AI/ML could be used to assist in the assessment of outcomes captured from diverse sources (e.g., DHTs, social media) during a clinical trial, including those consisting of large amounts of data for which manual review may be impractical.

D. Postmarketing Safety Surveillance

For purposes of this paper, pharmacovigilance (PV) refers to the science and activities related to the detection, assessment, understanding, and prevention of adverse events or any other drug-related problems (including medication errors and product quality issues).[12] Postmarketing safety surveillance, or PV activities in the post-approval period, includes postmarketing safety reporting of adverse events associated with use of human drug and biological products. An individual case safety report (ICSR) is used, as applicable, for the postmarketing reporting of adverse events to FDA and serves as an important data source of potential drug safety issues for postmarket safety surveillance. The clinical information in ICSRs can include suspect product or products, and temporal information related to use of the product and occurrence of the adverse event(s) in the patient's medical history, clinical course, and outcome. Complete and accurate reporting of ICSRs is critical to the understanding of a drug's safety profile. For reasons including increases in ICSR volume, AI/ML applications are being explored to help process and evaluate ICSR submissions within regulatory agencies (Ball & Dal Pan, 2022; Bate & Hobbiger, 2021).

1. Case Processing

There are potential opportunities to use AI/ML for automation during ISCR processing. The number and complexity of data sources of adverse events for ICSRs have increased, including from spontaneous reports, clinical trials, EHRs, social media, phone calls, emails, literature, patient registries, claims data, and post-approval safety studies (Beninger, 2020). The use of AI/ML to detect information from source documents could help identify adverse events for ICSR submission. For instance, the use of AI/ML to detect and evaluate drug event associations from literature and to screen social media for adverse events has been explored (Comfort, Dorrell, Meireis, &

[12] See the guidance for industry *Good Pharmacovigilance Practices and Pharmacoepidemiologic Assessment* (March 2005). Accessed September 30, 2022. https://www.fda.gov/media/71546/download See also, Council for International Organizations of Medical Sciences (CIOMS) Pharmacovigilance definition. Accessed September 29, 2022. https://cioms.ch/pharmacovigilance/

Fine, 2018; Negi, Pavuri, Patel, & Jain, 2019; S. V. Wang et al., 2017; W. Wang et al., 2011).

After an adverse event is identified from a data source, AI/ML could be used for case validity, case prioritization, duplicate check, coding, and quality control. The use of AI/ML can help identify whether a case is a valid case, which includes determining whether a case contains the minimum reporting requirements, such as an identifiable patient, suspect drug or biological product, adverse event(s), and identifiable reporter (Abatemarco et al., 2018; Schmider et al., 2019). During case intake, to assist in the prioritization of cases, AI/ML has been used to classify adverse events by expectedness (whether an adverse event is known and in the product labeling) (Abatemarco et al., 2018; Routray et al., 2020). Automated duplicate checks using AI/ML are being conducted to identify whether the case is a true duplicate, a follow up version of a prior case, or a new case (Kassekert 2022). Another area in which AI/ML has been applied is the coding of adverse events described in ICSRs to structured medical dictionary terms and for quality control purposes (Ghosh 2020).

2. Case Evaluation

Adverse event cases undergo clinical assessment. Case evaluation includes assessing the possibility of a causal relationship between the drug and adverse event, as well as assessing the outcome of the case. An AI model was developed based on relevant features used in causality assessments; it was trained, validated, and tested to classify cases by the probability of a causal relationship between the drug and adverse event (Comfort et al., 2018). AI/ML has also been applied to determine seriousness of the outcome of ICSRs (Routray, et al., 2020), which not only supports case evaluation, but also the timeliness of individual case submissions that require expedited reporting.

3. Case Submission

Generally, the final step after case processing is the submission of ICSRs. AI/ML algorithms have been used to automate reporting rules for submission of ICSRs to FDA. The reporting of ICSRs is required on an individual basis, as well as in aggregate (Ghosh et al., 2020). The aggregate reporting of adverse events generally involves the compilation of safety data for a product that is submitted at regular time intervals as specified. AI/ML can be used to develop aggregate reports that include multiple adverse events for particular products that occur within a time period for reporting purposes (Lewis & McCallum, 2020).

E. Advanced Pharmaceutical Manufacturing[13]

A critical aspect of drug development includes the methods, facilities, and controls used in manufacturing, processing, packing, and holding of a drug to help ensure that the

[13] The examples in this section are based on the review of general published information that projects or forecasts how AI/ML may be currently used in the pharmaceutical manufacturing space. In the continued

drug meets the requirements of safety and effectiveness, has the identity and strength it is represented to possess, and meets quality and purity characteristics. Advanced analytics leveraging AI/ML in the pharmaceutical manufacturing industry offers many possibilities, including, but not limited to, enhancing process control, increasing equipment reliability and throughput, monitoring early warnings or signals that the manufacturing process is not in a state of control, detecting recurring problem clusters, and preventing batch losses. The use of AI/ML to support pharmaceutical manufacturing can be deployed together with other advanced manufacturing technologies (e.g., process analytical technology, continuous manufacturing) to achieve the desired benefits. AI/ML is an enabler for the implementation of Industry 4.0, a term that refers to the fourth industrial revolution that brings together rapidly evolving technologies, and could result in a well-controlled, hyper-connected, digitized ecosystem and pharmaceutical value chain for the manufacturer (Arden et al., 2021). AI/ML could also be used to improve the reliability of the manufacturing supply chain through forecasting product demand, analyzing production schedules, estimating and mitigating the impact of potential disruptions, and optimizing inventory. Use of AI/ML-based approaches in pharmaceutical manufacturing can be broadly grouped into the areas outlined below that cover the entire drug manufacturing life cycle, from design to commercial manufacturing.

1. Optimization of Process Design

Digital twins can also be used in process design optimization. In this context, a digital twin of a process is a digital replica of the physical process used to better understand, analyze, predict, and optimize process performance. The digital twin could be especially beneficial for analyzing manufacturing processes characterized by a limited amount of development data, where AI/ML models could potentially leverage prior knowledge of the product and process (e.g., from previous studies, development programs, and scientific literature) to more quickly identify the optimal processing parameters, thus reducing design time and waste.

2. Advanced Process Control

Process controls have been implemented in pharmaceutical manufacturing for several decades. Traditional process controls maintain input process parameters at set points, but are not capable of simultaneously changing multiple input parameters to maintain the output parameters at desired levels to optimize the process. On the other hand, advanced process control (APC) allows dynamic control of the process to achieve a desired output (Huang et al., 2021). AI/ML techniques such as neural networks, with real-time process data as inputs, can be used to implement APC. These methods can

---

spirit of FDA's recent engagement through the Quality Metrics Feedback Program and CDER's Emerging Technology Program, FDA has been able to solicit valuable feedback demonstrated by industry interactions on several AI/ML use cases in the pharmaceutical manufacturing space, such as optimal risk-based supply chain modeling, business forecasting, process optimization, application of natural language processing (NLP) algorithms for complaints reduction, use of predictive analytics for non-conformance (NC) reduction, and corrective and preventive action (CAPA) effectiveness.

also be used to develop process controls that can predict whether a process is performing under a state of control by using AI/ML tools in combination with real-time sensor data, including, in conjunction with smart monitoring of production lines, to improve existing manufacturing line efficiency and output. In the near term, APC approaches that combine physics and chemistry knowledge with AI/ML techniques are expected to be increasingly adopted and have already been reported by several pharmaceutical manufacturers (National Academies of Sciences, 2021). In these APC applications, high quality model inputs inform process understanding and, model structure. These robust inputs, when combined with data-driven modeling, allow derivation of model parameters. These models leverage data required for model development while improving model robustness.

3. Smart Monitoring and Maintenance

Manufacturing processes can be automated and monitored in real time, leading to more efficient inventory management with shorter lead times and increased production output, without impacting product quality. AI/ML methods can be used to monitor equipment and detect deviations from normal performance that can trigger maintenance activities, thus reducing process downtime. Another example is the use of computer vision-based quality control that uses images (e.g., images of packaging, labels, or glass vials) that are analyzed by AI/ML-based software to detect deviations and to ensure images match the requirements of a given quality attribute of a product. Augmenting human visual inspection of drug products and packaging with such AI/ML-based methods can improve the accuracy and efficiency of visual inspection controls.

4. Trend Monitoring

AI/ML can be used in many ways to make manufacturing more effective and efficient with faster output, less waste, more informed decision-making, and enhanced quality control. Current practice for the analysis of deviations in the process is primarily done by quality personnel and relevant subject matter experts. AI/ML could be utilized to assist in examination of deviation reports that mostly contain large volumes of data or text to analyze manufacturing-related deviation trends, cluster problem areas, and prioritize areas for proactive continual improvement. This offers the advantage of expediting the process of identifying root causes, as solely manual review of deviation trends can be very time-consuming. AI/ML methods integrated with process performance (Ppk) and process capability (Cpk) metrics can be used to proactively monitor manufacturing operations for trends and out-of-control events, and predict thresholds for triggering CAPA effectiveness evaluations.

F. FDA Experience with AI/ML for Drug Development

FDA recognizes the increased use of AI/ML throughout the drug development life cycle and its potential to accelerate the development of safe and effective drugs. AI/ML is increasingly integrated in areas where FDA is actively engaged, including clinical trial design, DHTs, and RWD analytics. Over the last few years, FDA has seen a rapid growth in the number of submissions that reference AI/ML. Submissions across drug and biological product applications that include AI/ML have increased over the last few years to more than 100 submissions in 2021 (Q. Liu et al., 2022). These submissions cut across a range of therapeutic areas, and the uses of AI/ML within the submissions cover the many different areas of the drug development process highlighted in this section, from drug discovery and clinical trial enrichment to endpoint assessment and postmarket safety surveillance. Inclusion of AI/ML in the clinical development/research phase represents the most common stage for AI/ML uses in submissions.

One of the ways FDA has been supporting the development of innovative and robust AI/ML is through the establishment of the CDER AI Steering Committee (AISC), which coordinates efforts around AI/ML uses across therapeutic development. Leveraging its commitment to advancing innovative approaches and promoting collaborative efforts across the Agency, CDRH, including the DHCoE, have provided consults for drug submissions that involve AI/ML, and are developing a framework for AI/ML-based devices, including predetermined change control plans for devices incorporating AI/ML,[14] as well as a foundation for Good Machine Learning Practices for medical device development.[15] In addition, FDA has organized various workshops[16,17] and held a Patient Engagement Advisory Committee (PEAC) meeting on DHT and AI/ML-related topics[18] and has fostered regulatory science research, including on robustness, user-centered transparency, and bias identification and management, through external academic and clinical partnerships to evaluate the safety and effectiveness of emerging AI/ML products.[19]

Additionally, CDER has developed the Innovative Science and Technology Approaches for New Drugs (ISTAND) Pilot Program, which is designed to expand ***drug development tool*** (DDT) types included in the DDT qualification programs, including tools that leverage DHTs. Applications of AI/ML may represent novel DDTs or could be used to aid in the interpretation and analysis of traditional DDTs (such as ***biomarkers*** or ***clinical outcome assessments***), potentially speeding novel therapeutics to patients

[14] 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, April 2019. https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf
[15] Good Machine Learning Practice for Medical Device Development: Guiding Principles, October 2021. https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles
[16] See the Virtual Public Workshop – Transparency of Artificial Intelligence/Machine Learning-enabled Medical Devices, October 14, 2021. https://www.fda.gov/medical-devices/workshops-conferences-medical-devices/virtual-public-workshop-transparency-artificial-intelligencemachine-learning-enabled-medical-devices
[17] See the Public Workshop – Evolving Role of Artificial Intelligence in Radiological Imaging, February 25–26, 2020. https://www.fda.gov/medical-devices/workshops-conferences-medical-devices/public-workshop-evolving-role-artificial-intelligence-radiological-imaging-02252020-02262020
[18] See the Patient Engagement Advisory Committee Meeting Announcement, October 22, 2020. https://www.fda.gov/advisory-committees/advisory-committee-calendar/october-22-2020-patient-engagement-advisory-committee-meeting-announcement-10222020-10222020
[19] See CERSI research projects, October 2022. https://www.fda.gov/science-research/advancing-regulatory-science/cersi-research-projects

by enhancing the evidence available for decision-making.[20] In the area of model-informed drug development (MIDD), FDA's CDER and CBER have established a MIDD Pilot Program to facilitate the development and application of exposure-based, biological, and statistical models derived from nonclinical and clinical data sources.[21] In the context of MIDD, AI/ML could be employed to help improve clinical trial simulations, optimize dose selection or estimations, or enhance predictive or mechanistic safety evaluations.

In the area of postmarket safety surveillance, the FDA's Sentinel Initiative, including CDER's Sentinel System,[22] CBER's Biologics Effectiveness and Safety (BEST) system,[23] and CDRH's National Evaluation System for health Technology (NEST)[24] efforts, are exploring AI/ML approaches to improve existing systems. The FDA outlined its goals for using linked claims and EHR data supported by advanced analytics in the 5-year Sentinel System strategic plan.[25] The Sentinel System Innovation Center has outlined a four-pronged approach to implement this plan by incorporating emerging data science innovations and EHR data for medical product safety surveillance: (1) data infrastructure, (2) feature engineering, (3) causal inference, and (4) detection analytics (Desai et al., 2021). Examples of AI/ML applications in this approach include ***natural language processing*** (***NLP***) and automated feature extraction from unstructured EHR clinical notes for computable phenotyping and improved confounding adjustment from EHR-based variables using advanced statistical and ML approaches, such as algorithms created to enhance performance or "Super Learner" and targeted maximum likelihood estimation (Naimi & Balzer, 2018).

CBER's BEST system is designed to provide better data sources, methods, tools, expertise, and infrastructure to conduct surveillance and epidemiological studies.[26] Part of this program is an effort to use AI/ML methods to analyze EHRs to predict or better understand adverse events associated with the use of biological products and other products that CBER regulates. This work may also enhance FDA's understanding of the use of AI/ML methods for generating real-world evidence about product efficacy.

CDER is also exploring the application of AI to enhance the evaluation of ICSRs submitted to the FDA Adverse Event Reporting System (FAERS) (Ball & Dal Pan, 2022). The Information Visualization Platform (InfoViP) was developed with AI/ML to

---

[20] See the guidance for industry and FDA staff *Qualification Process for Drug Development Tools* (November 2020). https://www.fda.gov/media/133511/download
[21] See the Model-Informed Drug Development Paired Meeting Program, October 2022. https://www.fda.gov/drugs/development-resources/model-informed-drug-development-pilot-program
[22] See FDA's Sentinel Initiative, December 2022. https://www.fda.gov/safety/fdas-sentinel-initiative
[23] See the CBER Biologics Effectiveness and Safety (BEST) System, March 2022. https://www.fda.gov/vaccines-blood-biologics/safety-availability-biologics/cber-biologics-effectiveness-and-safety-best-system
[24] See the National Evaluation System for health Technology (NEST), October 2019. https://www.fda.gov/about-fda/cdrh-reports/national-evaluation-system-health-technology-nest
[25] See the FDA Sentinel System Five-Year Strategy, January 2019. https://www.fda.gov/media/120333/download
[26] See the CBER BEST System, March 2022. https://www.fda.gov/vaccines-blood-biologics/safety-availability-biologics/cber-biologics-effectiveness-and-safety-best-system

detect duplicate ICSRs, classify ICSRs by level of information quality, and derive visualization of the timeline of clinical events to aid in analysis of reported adverse events (Kreimeyer et al., 2022; Kreimeyer et al., 2021; Spiker et al., 2020). AI/ML methods have been investigated to automate the identification of adverse events in drug product labeling to support safety reviewers in the triaging of ICSRs to facilitate the identification of unknown or unexpected safety issues (Bayer et al., 2021; Ly et al., 2018). Another AI-based tool that focuses on drug product labeling and is currently in use is the Computerized Labeling Assessment Tool (CLAT), which serves to automate the review of label and labeling (e.g., prescribing information, carton and container labeling). NLP and ML are also being explored to classify free-text narratives in FAERS ICSRs into structured medical dictionary medication error terminologies to support the human review of coding quality. Additionally, through the FDA Quality Metrics Reporting Program,[27] CDER's Emerging Technology Program, and CBER's Advanced Technologies Team (CATT) Program,[28] FDA has been able to engage industry and gain valuable feedback on AI/ML use cases in pharmaceutical manufacturing.

The FDA also utilizes mechanisms such as a Broad Agency Announcement to solicit extramural proposals that address emerging regulatory science priorities, including leveraging external expertise and infrastructure to provide insight on the methods used to integrate and evaluate AI/ML in drug development.

III. Considerations for the Use of AI/ML in Drug Development

As shown in **Section II**, AI/ML has been applied to a broad range of drug development activities and continues to evolve. The use of AI/ML has the potential to accelerate the drug development process and make clinical trials safer and more efficient. However, it is important to assess whether the use of AI/ML introduces specific risks and harms. For example, AI/ML algorithms have the potential to amplify errors and preexisting biases present in underlying data sources and, when the findings are extrapolated outside of the testing environment, raise concerns related to generalizability and ethical considerations. Additionally, an AI/ML system may exhibit limited explainability due to its underlying complexity or may not be fully transparent for proprietary reasons. These concerns have resulted in a focus on developing standards for trustworthy AI that address specific characteristics in areas such as explainability, reliability, privacy, safety, security, and bias mitigation. This section begins with an overview of considerations and good practices for the general application of AI/ML and ends with questions to solicit feedback from stakeholders on these considerations and to further identify potential good practices in the context of drug development. This will aid FDA in further identifying opportunities and challenges with utilizing AI/ML throughout the drug development process.

---

[27] See the Quality Metrics for Drug Manufacturing, October 2022. https://www.fda.gov/drugs/pharmaceutical-quality-resources/quality-metrics-drug-manufacturing
[28] See the CBER Advanced Technologies Team (CATT) Program, June 27, 2019. https://www.fda.gov/vaccines-blood-biologics/industry-biologics/cber-advanced-technologies-team-catt

A. Overarching Standards and Practices for the Use of AI/ML

There has been an increased commitment by the Federal Government and the international community to facilitate AI innovation and adoption, which includes promoting trustworthy and ethical AI (*Exec. Order No. 13859, Maintaining American Leadership in Artificial Intelligence*, February 11, 2019; *Exec. Order No. 13960, Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government*, December 3, 2020; Lander & Nelson, October 22, 2021; *Notice of Request for Information on Public and Private Sector Uses of Biometric Technologies*, October 8, 2021; Organisation for Economic Co-operation and Development, 2019; Vought, 2020). As a result, efforts for the development of cross-sector and sector-specific standards to facilitate the technological advancement of AI have rapidly increased in both domestic and international forums. For example, in August 2019, the National Institute for Standards and Technology (NIST) released "U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools" to help ensure the use of technical standards and to advance innovation, trust, and confidence in the use of AI (National Institute of Standards and Technology, 2019). The plan identified several areas of focus for AI standards development, including data and knowledge, performance testing and reporting methodology, risk management, and trustworthiness, among others. Other standards organizations, such as the International Organization for Standardization (ISO), the Institute of Electrical and Electronics Engineers (IEEE), and the International Electrotechnical Commission (IEC), are also developing relevant AI/ML standards and work products addressing fundamental issues of data quality, explainability, and performance, in addition to examining applications that are specific to certain industries. The Verification and Validation (V&V 40) risk-informed credibility assessment framework was initially developed by the American Society of Mechanical Engineers (ASME) for the assessment of credibility of computational models used for medical devices (American Society of Mechanical Engineers, 2018) and was later adopted into model-informed drug development[29] (Kuemmel et al., 2020; Viceconti et al., 2021). As AI/ML is also used for computational models, the V&V 40 framework potentially serves to inform whether the AI/ML model is credible for use in drug development.[30] The V&V 40 Standard, which is not specific to AI/ML and does not specify activities or define criteria required to establish model credibility for a particular context of use or application, has been adapted for medical devices and for model-informed drug development.[31,32]

[29] Promoting Innovation in Medical Product Assessment: A Risk-based Framework for Evaluating Computational Models for Regulatory Decision-Making, October 2020. https://www.fda.gov/drugs/news-events-human-drugs/promoting-innovation-medical-product-assessment-risk-based-framework-evaluating-computational-models

[30] A V&V 70 Subcommittee has been established for Verification and Validation of Machine Learning.

[31] See the draft guidance for industry and FDA staff *Assessing the Credibility of Computational Modelling Simulation in Medical Device Submissions* (December 2021). When final, this guidance will represent FDA's current thinking on this topic. https://www.fda.gov/media/154985/download

[32] Promoting Innovation in Medical Product Assessment: A Risk-based Framework for Evaluating Computational Models for Regulatory Decision-Making, October 2020. https://www.fda.gov/drugs/news-events-human-drugs/promoting-innovation-medical-product-assessment-risk-based-framework-evaluating-computational-models

In addition to the V&V 40 Standard for evaluating the predictive capability of computational models for medical devices, FDA, Health Canada, and the United Kingdom's Medicines and Healthcare products Regulatory Agency (MHRA) jointly published 10 guiding principles to inform the development of Good Machine Learning Practices (GMLP) for medical devices that use AI/ML.[33] The guiding principles include adopting a total product life cycle approach in which multidisciplinary expertise is leveraged throughout product development, with an in-depth understanding of how the model is integrated into the clinical workflow. The principles also emphasize the importance of adequate representation within the clinical study population to manage bias, improve generalizability, and provide sufficient transparency with clear and essential information, such as the product's intended use and indications, the data used to test and train the model, and known limitations. Finally, these GMLP highlight the importance of monitoring deployed models for performance while managing the risk of model retraining. FDA's CDRH had previously discussed the role of GMLP for medical devices, and in 2019 issued a proposed framework for modifications to AI/ML-based SaMD. The framework proposed a predetermined change control plan mechanism—whereby a sponsor can proactively specify intended modifications to device software incorporating AI/ML and the methods that will be used to ensure their safety and effectiveness—thereby laying the foundation for AI/ML-enabled devices with improved capacity for adaptation.[34]

Although the standards and practices described in this section were not tailored specifically for drug development, the utility and applicability of these standards to drug development and the development of medical devices intended to be used with drugs, will be explored to ensure alignment and consistency.

B. Discussion of Considerations and Practices for AI/ML in Drug Development

Informed by the diverse applications of AI/ML in drug development (see **Section II**), FDA is considering approaches to provide regulatory clarity around the use of AI/ML in drug development, supported by an expanding body of knowledge and a clear appreciation of the opportunities and challenges with utilizing AI/ML in drug development. While certain standards and practices outlined in **Section III.A** can potentially be adapted to address the use of AI/ML in the context of drug development, the use of AI/ML in drug development may raise specific challenges that could highlight additional considerations. As noted above, this document is not FDA guidance or policy and does not endorse any specific approaches for the use of AI/ML in drug development. However, the feedback and future discussions with stakeholders can help inform future regulatory activities.

---

[33] Good Machine Learning Practice for Medical Device Development: Guiding Principles, October 2021. https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles

[34] 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, April 2019. https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf

Adapting the overarching principles of the General Accountability Office AI accountability framework[35] below, FDA's CDER, CBER, CDRH, including DHCoE, aim to initiate a discussion with stakeholders and solicit feedback on three key areas in the context of AI/ML in drug development:

(1) human-led governance, accountability, and transparency;

(2) quality, reliability, and representativeness of data; and

(3) model development, performance, monitoring, and validation.

In each of these areas, a risk-based approach could include measures commensurate with the level of risk posed by the specific context of use for AI/ML.

(1) <u>Human-led governance, accountability, and transparency</u>

Human-led AI/ML governance can help ensure adherence to legal and ethical values, where accountability and transparency are essential for the development of trustworthy AI. Such governance and clear accountability may extend across the spectrum of planning, development, use, modification, and discontinuation (as applicable) of AI/ML in the drug development process.

As part of governance, a risk management plan that considers the context of use may be applied to identify and mitigate risks. This approach can help guide the level of documentation, transparency, and explainability, with tracking and recording of key steps and decisions, including the rationale for any deviations and procedures that enable vigilant oversight and auditing. Transparency and documentation can provide critical insight on the initial planning, development, function, and any modifications of the AI/ML in the specific context of use, while explainability can provide accompanying evidence or reason for the outputs.

Questions:

- In what specific use cases or applications of AI/ML in drug development are there the greatest need for additional regulatory clarity?

- What does transparency mean in the use of AI/ML in drug development (for example, transparency could be considered as the degree to which appropriate information about the AI/ML model—including its use, development,

[35] See Artificial Intelligence: An Accountability Framework for Federal Agencies and Other Entities (June 2021). https://www.gao.gov/assets/gao-21-519sp.pdf

performance, and, when available, logic—is clearly communicated to regulators and/or other stakeholders)?[36]

- In your experience, what are the main barriers and facilitators of transparency with AI/ML used during the drug development process (and in what context)?

- What are some of the good practices utilized by stakeholders for providing risk-based, meaningful human involvement when AI/ML is being utilized in drug development?

- What processes are in place to enhance and enable traceability and auditability?

- How are pre-specification activities managed, and changes captured and monitored, to ensure the safe and effective use of AI/ML in drug development?

(2) <u>Quality, reliability, and representativeness of data</u>

AI/ML is particularly sensitive to the attributes or characteristics of the data used for training, testing, and validation. Although not unique to AI/ML, missing data, bias, and data drift are typically important considerations. Ensuring data quality, reliability, and that the data are fit for use (i.e., relevant for the specific intended use and population) can be critical. Potential data-related issues to consider include:

Bias: AI/ML can potentially amplify preexisting biases that exist in the underlying input data. NIST published a document characterizing three categories of bias (human, systemic, and statistical/computational) and "how they may occur in the commission, design, development, and deployment of AI technologies that can be used to generate predictions, recommendations, or decisions (e.g., algorithmic decision systems), and how AI systems may create societal harms."[37]

Integrity: The completeness, consistency, and ***accuracy*** of data.[38]

Privacy and security: The protection and privacy of data, linked to data classifications and the technical features of the system.

Provenance: Record trail that accounts for the origin of a piece of data (in a database, document, or repository) together with an explanation of how and why it got to the present place.[39] Provenance describes "the metadata, or extra

---

[36] Adapted from ISO/IEC JTC1/SC42 DIS 25059 (draft). https://www.iso.org/standard/80655.html?browse=tc
[37] NIST Special Publication 1270, March 2022. https://doi.org/10.6028/NIST.SP.1270
[38] For additional considerations related to data integrity see the guidance for industry *Data Integrity and Compliance with Drug CGMP* (December 2018). https://www.fda.gov/media/119267/download
[39] Encyclopedia of Database Systems, definition of data provenance. https://link.springer.com/referenceworkentry/10.1007%2F978-0-387-39940-9_1305

information about data, that can help answer questions such as who created the data and when."[40]

Relevance: Adequate data are available and are appropriate for the intended use.

Replicability: Obtaining consistent results across studies aimed at answering the same question, each of which has obtained its own data.[41] It is important to clarify data access early in the process.

Reproducibility: Obtaining consistent results using the same input data, computational steps, methods and code, and conditions of analysis[42] (while not confirming validity, the transparency required to demonstrate reproducibility permits evaluation of the validity of design and operational decisions (S. V. Wang et al., 2017)).

Representativeness: Confidence that a sample from which evidence is generated is sufficiently similar to the intended population. In the context of patient experience data, representativeness includes the extent to which the elicited experiences, perspectives, needs, and priorities of the sample are sufficiently similar to those of the intended patient population.[43]

Questions:

- What additional data considerations exist for AI/ML in the drug development process?

- What practices are developers, manufacturers, and other stakeholders currently utilizing to help assure the integrity of AI/ML or to address issues, such as bias, missing data, and other data quality considerations, for the use of AI/ML in drug development?

- What are some of the key practices utilized by stakeholders to help ensure data privacy and security?

- What are some of the key practices utilized by stakeholders to help address issues of reproducibility and replicability?

- What processes are developers using for bias identification and management?

[40] 21st Century Cures Act: Interoperability, Information Blocking, and the ONC Health IT Certification Program (March 2019). https://www.federalregister.gov/documents/2019/03/04/2019-02224/21st-century-cures-act-interoperability-information-blocking-and-the-onc-health-it-certification
[41] *Ibid.*
[42] National Academies of Sciences, Engineering, and Medicine, 2019, Reproducibility and Replicability in Science. https://doi.org/10.17226/25303
[43] See discussion document for Patient-focused Drug Development Public Workshop *Collecting Comprehensive and Representative Input*, December 2017. https://www.fda.gov/media/109179/download

(3) Model development, performance, monitoring, and validation

The use of the model may be important to consider in evaluating AI/ML model development and performance, including through practices of pre-specification steps and clear documentation of criteria for developing and assessing models. It may also be important to consider the model risk and credibility; model risk drives the selection of credibility goals and activities.[44] Model risk is determined by two factors, which are shaped by the ***context of use***: model influence (the weight of the model in the totality of evidence for a specific decision) and decision consequence (the potential consequences of a wrong decision).

In balancing performance and explainability, it may be important to consider the complexity of the AI/ML model. In situations where complex models (e.g., artificial neural network models) are determined to have similar performance, there may be overall advantages to selecting the more traditional and parsimonious (i.e., fewer parameters) model.

It may also be important to monitor and document monitoring efforts of the AI/ML model to ensure it is reliable, relevant, and consistent over time. This includes documentation of the results of monitoring and any corrective action taken to ensure that the AI/ML produces intended results. Subsequent assessments (e.g., postmarket safety monitoring, surveillance) can provide valuable feedback on processes and real-world model performance. Real-world model performance includes applications that may be supported by collection and monitoring of RWD (e.g., electronic health records, product and disease registries). Potential re-training based on real-world performance could provide important insights to model performance, and following such re-training, it may be important to monitor and document the AI/ML model to appropriately manage risks.

Data considerations also include providing the details of the training dataset utilized to develop the AI/ML model, along with the performance, when employing independent, external testing data to support verification and validation ("external validity"). It is generally important for data of sufficient quality for the particular context of use to be representative of the population where the AI/ML method will be utilized. It is important to help ensure AI/ML models are validated to produce results that are credible for the model's use. Credibility activities include verification of the software code and calculations, validation of the model, and evaluation of the applicability of

---

[44] Credibility refers to trust in the predictive capability of a computational model for a particular context of use (Kuemmel et al., 2020). This includes steps to document performance and approaches to measure uncertainty at the component level (e.g., model and non-level components, including metrics and assessing performance and outcome of each component) and system level (e.g., methods for assessment, performance metrics, and outcomes), where feasible. Demonstration of credibility often includes a risk-based approach, where uses presenting the highest risk generally require the greatest standard of evidence, with a gradient of evidence needed based on the associated risk (i.e., informing early-stage drug development for non-serious medical condition versus evaluating drug safety and effectiveness for critical medical condition).

validation assessments to the context of use. These activities include considerations of measuring the level of uncertainty of the model predictions. Upon completion of credibility activities, an assessment can be made to determine whether the model is sufficiently credible for its use and whether the model may be acceptable for a given regulatory purpose.

Questions:

- What are some examples of current tools, processes, approaches, and best practices being used by stakeholders for:

  - Documenting the development and performance of AI/ML models that can be applied in the context of drug development (e.g., CONSORT-AI (Liu et al., 2020) and SPIRIT-AI (Cruz Rivera et al., 2020))?

  - Selecting model types and algorithms for a given context of use?

  - Determining when to use specific approaches for validating models and measuring performance in a given context of use (e.g., selecting relevant success criteria and performance measures)?

  - Evaluating transparency and explainability and increasing model transparency?

  - Addressing issues of accuracy and explainability (e.g., scenarios where models may provide increased accuracy, while having limitations in explainability)?

  - Selecting open-source AI software for AI/ML model development? What are considerations when using open-source AI software?

  - The use of RWD performance in monitoring AI/ML?

- What practices and documentation are being used to inform and record data source selection and inclusion or exclusion criteria?

- In what context of use are stakeholders addressing explainability, and how have you balanced considerations of performance and explainability?

- What approaches are being used to document the assessment of uncertainty in model predictions, and how is uncertainty being communicated? What methods and standards should be developed to help support the assessment of uncertainty?

As outlined above, many of the overarching principles and standards related to the characteristics of trustworthy AI can help inform considerations or key practice areas for the application of AI/ML in the context of drug development. 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. Transparency and documentation across the entire product life cycle can help build trust in the use of AI/ML. In this regard, it may be important to consider pre-specification and documentation of the purpose or question of interest, context of use, risk, and development of AI/ML. While not unique to the use of AI/ML in drug development, there are also a broad range of data quality, relevance, and reliability-related considerations. Related to the area of model development, performance, monitoring, and validation, the V&V 40 risk-informed credibility assessment framework may be a helpful guide when considering the specific use for AI/ML. In general, use of a risk-based approach may guide the level of evidence and record keeping needed for the verification and validation of AI/ML models for a specific context of use. Engagement with the FDA early in the process can also help inform and address these considerations.

IV. Next Steps: Engagement and Collaboration

The release of this initial discussion paper is part of a broader effort to communicate with a range of stakeholders and to explore the relevant considerations for the use of AI/ML in the development of human drugs and biological products. Coupled with this document, FDA has included a series of questions for feedback, and a workshop with stakeholders is planned to provide an opportunity for further engagement. The FDA will also provide several other mechanisms to engage with stakeholders, sponsors, and developers on this topic, and these can be utilized to address questions before conducting a study that utilizes AI/ML. In addition to formal meetings where these methods can be discussed, the Critical Path Innovation Meetings (CPIM),[45] ISTAND Pilot Program,[46] Emerging Technology Program,[47] and Real-World Evidence Program[48] meetings are examples of additional avenues for communicating and discussing a relevant AI/ML methodology or technology and improving efficiency and quality in drug development. Additionally, communication and engagement with patients and the public regarding considerations for AI/ML in drug development is critical to ensure patient-centered approaches and policies.

Building on this discussion paper, FDA will continue to solicit feedback and engage a broad group of stakeholders to further discuss considerations for utilizing AI/ML throughout the drug development life cycle. These discussions and future

[45] See CPIM, November 11, 2022. https://www.fda.gov/drugs/new-drugs-fda-cders-new-molecular-entities-and-new-therapeutic-biological-products/critical-path-innovation-meetings-cpim
[46] See the ISTAND Pilot Program, February 10, 2021. https://www.fda.gov/drugs/drug-development-tool-ddt-qualification-programs/innovative-science-and-technology-approaches-new-drugs-istand-pilot-program
[47] See Emerging Technology Program, February 22, 2022. https://www.fda.gov/about-fda/center-drug-evaluation-and-research-cder/emerging-technology-program
[48] See Framework for FDA's Real World Evidence Program, April 14, 2020. https:/fda.gov/media/120060/download

collaborations with stakeholders may provide a foundation for a future framework or guidance.

Glossary

Accuracy: The level of agreement between the measured value and the true value of the clinical event or characteristic.

Artificial Intelligence (**AI**): A branch of computer science, statistics, and engineering that uses algorithms or models to perform tasks and exhibit behaviors such as learning, making decisions, and making predictions.[49]

Biomarker: A defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or biological responses to an exposure or intervention, including therapeutic interventions. Biomarkers may include molecular, histologic, radiographic, or physiologic characteristics. A biomarker is not a measure of how an individual feels, functions, or survives.[50]

Clinical Outcome Assessment (COA): A measure that describes or reflects how a patient feels, functions, or survives. There are four types of COAs: patient-reported outcome, observer-reported outcome, clinician-reported outcome, and performance outcome.[51]

Context of Use: A statement that fully and clearly describes the way AI/ML is to be used and the drug development-related purpose of the use.[52]

Controlled Terminology: A finite set of values (e.g., codes, text, numeric) that represent the only allowed values for a data item. Generally, controlled terminology standards specify the key concepts that are represented as definitions, preferred terms, synonyms, and code systems.[53]

Decentralized Clinical Trial: A clinical investigation where some or all of the trial-related activities occur at a location separate from the investigator's location.[54]

Digital Health Technology (**DHT**): A system that uses computing platforms, connectivity, software, and/or sensors for health care and related uses. These technologies span a wide range of uses, from applications in general wellness to applications as a medical device. They include technologies intended for use as a

---

[49] See IMDRF/AIMD WG/N67 Machine Learning-enabled Medical Devices: Key Terms and Definitions, final document, May 6, 2022. https://www.imdrf.org/documents/machine-learning-enabled-medical-devices-key-terms-and-definitions

[50] See BEST (Biomarkers, EndpointS, and other Tools) Resource Glossary, 2016. https://www.ncbi.nlm.nih.gov/books/NBK338448

[51] See Clinical Outcome Assessment (COA), December 2020. https://www.fda.gov/about-fda/clinical-outcome-assessment-coa-frequently-asked-questions

[52] CDISC Glossary, 2022. https://evs.nci.nih.gov/ftp1/CDISC/Glossary/CDISC%20Glossary.html

[53] *Ibid.*

[54] See the draft guidance for industry, investigators, and other stakeholders *Digital Health Technologies for Remote Data Acquisition in Clinical Investigations* (December 2021). When final, this guidance will represent FDA's current thinking on this topic. https://www.fda.gov/media/155022/download

medical product, in a medical product, or as an adjunct to other medical products (devices, drugs, and biologics). They may also be used to develop or study medical products. Data captured by DHTs can often be transmitted directly to investigators, sponsors, and/or other authorized parties, with the capability to maintain blinding or masking when appropriate. The ability to transmit data remotely increases opportunities for patients to participate in clinical investigations at locations remote from the investigator's site.[55].

Digital Twins: An integrated multi-physics, multiscale, probabilistic simulation of a complex system that uses the best available data, sensors, and models to mirror the behavior of its corresponding twin. A fully developed digital twin consists of a physical component (e.g., unit operations), a virtual component, and automated data communications between the two. The development and application of digital twins are now being extended to manufacturing and complex products to assess sensitivities of material attributes and process parameters, reliability of control strategies, and effectiveness of mitigation plans for potential disturbances.[56]

Drug Development Tool (**DDT**): A biomarker, COA, or any other method, material, or measure determined to aid drug development and regulatory review. Animal models developed to be used for product development under the Animal Rule[57] have been determined by FDA to be DDTs under section 507 of the FD&C Act.[58]

Endpoint: A precisely defined variable intended to reflect an outcome of interest that is statistically analyzed to address a particular research question. A precise definition of an endpoint typically specifies the type of assessments made, the timing of those assessments, the assessment tools used, and possibly other details, as applicable, such as how multiple assessments within an individual are to be combined.[59]

Machine Learning (**ML**): A subset of AI that allows ML models to be developed by ML training algorithms through analysis of data, without being explicitly programmed.[60]

Natural Language Processing (**NLP**): The branch of computer science, specifically the branch of AI, concerned with giving computers the ability to understand text and spoken words in much the same way human beings can.[61]

---

[55] *Ibid.*

[56] See Modeling & Simulation at FDA, November 16, 2022. https://www.fda.gov/science-research/about-science-research-fda/modeling-simulation-fda

[57] See Animal Rule Approvals, June 2022. https://www.fda.gov/drugs/nda-and-bla-approvals/animal-rule-approvals

[58] See the guidance for industry and FDA staff *Qualification Process for Drug Development Tools* (November 2020). https://www.fda.gov/media/133511/download

[59] See BEST (Biomarkers, EndpointS, and other Tools) Resource Glossary, 2016. https://www.ncbi.nlm.nih.gov/books/NBK338448

[60] See IMDRF/AIMD WG/N67 Machine Learning-enabled Medical Devices: Key Terms and Definitions, final document, May 6, 2022. https://www.imdrf.org/documents/machine-learning-enabled-medical-devices-key-terms-and-definitions

[61] "What is natural language processing?" Accessed September 8, 2022. https://www.ibm.com/cloud/learn/natural-language-processing#toc-what-is-na-jLju4DjE

Neural Network: A commonly used form of AI/ML that is used for categorization applications and has been loosely likened to the way that neurons in the brain process signals. Neural networks typically consist of at least three layers of neurons: input layer (which receives information), hidden layer (responsible for extracting patterns and conducting the internal processing), and output layer (produces and presents the final network output).[62]

Real-World Data (**RWD**): The data relating to patient health status and/or the delivery of health care routinely collected from a variety of sources. Examples of RWD include data derived from electronic health records (EHRs); medical claims and billing data; data from product and disease registries; patient-generated data, including from in-home-use settings; and data gathered from other sources that can inform on health status, such as mobile devices.[63]

Real-World Evidence (**RWE**): The clinical evidence about the usage and potential benefits or risks of a medical product derived from analysis of RWD. RWD sources (e.g., registries, collections of EHRs, administrative and medical claims databases) can be used for data collection and, in certain cases, to develop analysis infrastructure to support many types of study designs to develop RWE, including, but not limited to, randomized trials (e.g., large simple trials, pragmatic clinical trials) and observational studies (prospective or retrospective).[64]

Recurrent Neural Network: A type of artificial neural network that uses sequential data or time series data to exhibit temporal dynamic behavior. These algorithms are commonly used for ordinal or temporal problems, such as language translation, NLP, speech recognition, and image captioning.[65]

---

62 See the Executive Summary for the Patient Engagement Advisory Committee Meeting: Artificial Intelligence and Machine Learning in Medical Devices, October 22, 2020. https://www.fda.gov/media/142998/download
63 See the draft guidance for industry, investigators, and other stakeholders Real-World Data: Assessing Electronic Health Records and Medical Claims Data to Support Regulatory Decision-Making for Drug and Biological Products (September 2021). https://www.fda.gov/media/152503/download
64 Ibid.
65 Adapted from https://www.ibm.com/cloud/learn/recurrent-neural-networks

References

Abatemarco, D., Perera, S., Bao, S. H., Desai, S., Assuncao, B., Tetarenko, N., . . . Mingle, E. (2018). Training Augmented Intelligent Capabilities for Pharmacovigilance: Applying Deep-learning Approaches to Individual Case Safety Report Processing. *Pharmaceut Med, 32*(6), 391-401. doi:10.1007/s40290-018-0251-9

Aerts, H. J. W. L., Grossmann, P., Tan, Y., Oxnard, G. R., Rizvi, N., Schwartz, L. H., & Zhao, B. (2016). Defining a Radiomic Response Phenotype: A Pilot Study using targeted therapy in NSCLC. *Sci Rep, 6*, 33860. doi:10.1038/srep33860

American Society of Mechanical Engineers. (2018). Assessing Credibility of Computational Modeling through Verification and Validation: Application to Medical Devices. In.

Arden, N. S., Fisher, A. C., Tyner, K., Yu, L. X., Lee, S. L., & Kopcha, M. (2021). Industry 4.0 for pharmaceutical manufacturing: Preparing for the smart factories of the future. *Int J Pharm, 602*, 120554. doi:10.1016/j.ijpharm.2021.120554

Athreya, A. P., Neavin, D., Carrillo-Roa, T., Skime, M., Biernacka, J., Frye, M. A., . . . Bobo, W. V. (2019). Pharmacogenomics-Driven Prediction of Antidepressant Treatment Outcomes: A Machine-Learning Approach With Multi-trial Replication. *Clin Pharmacol Ther, 106*(4), 855-865. doi:10.1002/cpt.1482

Avram, R., Olgin, J. E., Kuhar, P., Hughes, J. W., Marcus, G. M., Pletcher, M. J., . . . Tison, G. H. (2020). A digital biomarker of diabetes from smartphone-based vascular signals. *Nat Med, 26*(10), 1576-1582. doi:10.1038/s41591-020-1010-5

Ball, R., & Dal Pan, G. (2022). "Artificial Intelligence" for Pharmacovigilance: Ready for Prime Time? *Drug Saf, 45*(5), 429-438. doi:10.1007/s40264-022-01157-4

Bate, A., & Hobbiger, S. F. (2021). Artificial Intelligence, Real-World Automation and the Safety of Medicines. *Drug Saf, 44*(2), 125-132. doi:10.1007/s40264-020-01001-7

Bayer, S., Clark, C., Dang, O., Aberdeen, J., Brajovic, S., Swank, K., . . . Ball, R. (2021). ADE Eval: An Evaluation of Text Processing Systems for Adverse Event Extraction from Drug Labels for Pharmacovigilance. *Drug Saf, 44*(1), 83-94. doi:10.1007/s40264-020-00996-3

Beninger, P. (2020). Signal Management in Pharmacovigilance: A Review of Activities and Case Studies. *Clin Ther, 42*(6), 1110-1129. doi:10.1016/j.clinthera.2020.03.018

Bulitta, J. B., Hope, W. W., Eakin, A. E., Guina, T., Tam, V. H., Louie, A., . . . Hoover, J. L. (2019). Generating Robust and Informative Nonclinical In Vitro and In Vivo Bacterial Infection Model Efficacy Data To Support Translation to Humans. *Antimicrob Agents Chemother, 63*(5), e02307–02318. doi:10.1128/AAC.02307-18

Chan, H. C. S., Shan, H., Dahoun, T., Vogel, H., & Yuan, S. (2019). Advancing Drug Discovery via Artificial Intelligence. *Trends Pharmacol Sci, 40*(8), 592-604. doi:10.1016/j.tips.2019.06.004

Cohoon, T. J., & Bhavnani, S. P. (2020). Toward precision health: applying artificial intelligence analytics to digital health biometric datasets. *Per Med, 17*(4), 307-316. doi:10.2217/pme-2019-0113

Comfort, S., Dorrell, D., Meireis, S., & Fine, J. (2018). MOdified NARanjo Causality Scale for ICSRs (MONARCSi): A Decision Support Tool for Safety Scientists. *Drug Saf, 41*(11), 1073-1085. doi:10.1007/s40264-018-0690-y

Cruz Rivera, S., Liu, X., Chan, A. W., Denniston, A. K., Calvert, M. J., & and the SPIRIT-AI and CONSORT-AI Working Group; SPIRIT-AI and CONSORT-AI Steering Group; SPIRIT-AI and CONSORT-AI Consensus Group. (2020). Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. *Nat Med, 26*(9), 1351-1363. doi:10.1038/s41591-020-1037-7

Daryaee, F., & Tonge, P. J. (2019). Pharmacokinetic-pharmacodynamic models that incorporate drug-target binding kinetics. *Curr Opin Chem Biol, 50*, 120-127. doi:10.1016/j.cbpa.2019.03.008

Dercle, L., Fronheiser, M., Lu, L., Du, S., Hayes, W., Leung, D. K., . . . Zhao, B. (2020). Identification of non-small cell lung cancer sensitive to systemic cancer therapies using radiomics. *Clin Cancer Res, 26*(9), 2151-2162. doi:10.1158/1078-0432.CCR-19-2942

Desai, R. J., Matheny, M. E., Johnson, K., Marsolo, K., Curtis, L. H., Nelson, J. C., . . . Schneeweiss, S. (2021). Broadening the reach of the FDA Sentinel system: A roadmap for integrating electronic health record data in a causal analysis framework. *NPJ Digit Med, 4*(1), 170. doi:10.1038/s41746-021-00542-0

ElZarrad, M. K., Lee, A. Y., Purcell, R., & Steele, S. J. (2022). Advancing an agile regulatory ecosystem to respond to the rapid development of innovative technologies. *Clin Transl Sci, 15*(5), 1332–1339. doi:10.1111/cts.13267

European Medicines Agency. (2022). *DRAFT Qualification opinion for Prognostic Covariate Adjustment (PROCOVA™)*. European Medicines Agency Retrieved from https://www.ema.europa.eu/documents/other/draft-qualification-opinion-prognostic-covariate-adjustment-procovatm_en.pdf

Exec. Order No. 13859, Maintaining American Leadership in Artificial Intelligence. (2019-02544). (February 11, 2019). 84 Fed. Reg. 3967 Retrieved from https://www.federalregister.gov/documents/2019/02/14/2019-02544/maintaining-american-leadership-in-artificial-intelligence

Exec. Order No. 13960, Promoting the Use of Trustworthy Artificial Intelligence in the Federal Government. (2020-27065). (December 3, 2020). 85 Fed. Reg. 78939 Retrieved from https://www.federalregister.gov/documents/2020/12/08/2020-27065/promoting-the-use-of-trustworthy-artificial-intelligence-in-the-federal-government

Fumagalli, S. E., Padhiar, N. H., Meyer, D., Katneni, U., Bar, H., DiCuccio, M., . . . Kimchi-Sarfaty, C. (2023). Analysis of 3.5 million SARS-CoV-2 sequences reveals unique mutational trends with consistent nucleotide and codon frequencies. *Virol J, 20*(1), 31. doi:10.1186/s12985-023-01982-8

Ghosh, R., Kempf, D., Pufko, A., Barrios Martinez, L. F., Davis, C. M., & Sethi, S. (2020). Automation opportunities in pharmacovigilance: An industry survey. *Pharmaceut Med, 34*(1), 7-18. doi:10.1007/s40290-019-00320-0

Harrer, S., Shah, P., Antony, B., & Hu, J. (2019). Artificial Intelligence for Clinical Trial Design. *Trends Pharmacol Sci, 40*(8), 577-591. doi:10.1016/j.tips.2019.05.005

Harrison, L. I., & Gibaldi, M. (1977). Physiologically based pharmacokinetic model for digoxin disposition in dogs and its preliminary application to humans. *J Pharm Sci, 66*(12), 1679-1683. doi:10.1002/jps.2600661206

Hsu, V., de, L. T. V., M., Zhao, P., Zhang, L., Zheng, J. H., Nordmark, A., . . . Huang, S. M. (2014). 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, 53*(3), 283-293. doi:10.1007/s40262-013-0117-y

Huang, J., O'Connor, T., Ahmed, K., Chatterjee, S., Garvin, C., Ghosh, K., . . . Warman, M. (2021). AIChE PD2M Advanced Process Control workshop-moving APC forward in the pharmaceutical industry. *Journal of Advanced Manufacturing and Processing, 3*(1), e10071. doi:10.1002/amp2.10071

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., . . . Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), 583-589. doi:10.1038/s41586-021-03819-2

Kawakami, E., Tabata, J., Yanaihara, N., Ishikawa, T., Koseki, K., Iida, Y., . . . Okamoto, A. (2019). Application of atificial intelligence for preoperative diagnostic and prognostic prediction in epithelial ovarian cancer based on blood biomarkers. *Clin Cancer Res, 25*(10), 3006-3015. doi:10.1158/1078-0432.CCR-18-3378

Kreimeyer, K., Dang, O., Spiker, J., Gish, P., Weintraub, J., Wu, E., . . . Botsis, T. (2022). Increased Confidence in Deduplication of Drug Safety Reports with Natural Language Processing of Narratives at the US Food and Drug Administration. *Frontiers in Drug Safety and Regulation, 2*. doi:10.3389/fdsfr.2022.918897

Kreimeyer, K., Dang, O., Spiker, J., Munoz, M. A., Rosner, G., Ball, R., & Botsis, T. (2021). Feature engineering and machine learning for causality assessment in pharmacovigilance: Lessons learned from application to the FDA Adverse Event Reporting System. *Comput Biol Med, 135*, 104517. doi:10.1016/j.compbiomed.2021.104517

Kuemmel, C., Yang, Y., Zhang, X., Florian, J., Zhu, H., Tegenge, M., . . . Zineh, I. (2020). Consideration of a credibility assessment framework in model-informed drug development: Potential application to physiologically-based pharmacokinetic modeling and simulation. *CPT Pharmacometrics Syst Pharmacol, 9*(1), 21-28. doi:10.1002/psp4.12479

Lander, E., & Nelson, A. (October 22, 2021). *ICYMI: WIRED (Opinion): Americans Need a Bill of Rights for an AI-Powered World*. https://www.whitehouse.gov/ostp/news-updates/2021/10/22/icymi-wired-opinion-americans-need-a-bill-of-rights-for-an-ai-powered-world/

Laubenbacher, R., Sluka, J. P., & Glazier, J. A. (2021). Using digital twins in viral infection. *Science, 371*(6534), 1105-1106. doi:10.1126/science.abf3370

Lewis, D. J., & McCallum, J. F. (2020). Utilizing advanced technologies to augment pharmacovigilance systems: Challenges and opportunities. *Ther Innov Regul Sci, 54*(4), 888-899. doi:10.1007/s43441-019-00023-3

Liu, Q., Huang, R., Hsieh, J., Zhu, H., Tiwari, M., Liu, G., . . . Huang, S. M. (2022). Landscape analysis of the application of artificial antelligence and machine learning in regulatory submissions for drug development from 2016 to 2021. *Clin Pharmacol Ther*. doi:10.1002/cpt.2668

Liu, X., Cruz Rivera, S., Moher, D., Calvert, M. J., Denniston, A. K., & the SPIRIT-AI and CONSORT-AI Working Group. (2020). Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat Med, 26*(9), 1364-1374. doi:10.1038/s41591-020-1034-x

Liu, X., Liu, C., Huang, R., Zhu, H., Liu, Q., Mitra, S., & Wang, Y. (2021). Long short-term memory recurrent neural network for pharmacokinetic-pharmacodynamic modeling. *Int J Clin Pharmacol Ther, 59*(2), 138-146. doi:10.5414/CP203800

Liu, Z., Chen, X., Carter, W., Moruf, A., Komatsu, T. E., Pahwa, S., . . . W., T. (2022). AI-powered drug repurposing for developing COVID-19 treatments. In *Reference Module in Biomedical Sciences*: Elsevier.

Lu, J., Deng, K., Zhang, X., Liu, G., & Guan, Y. (2021). Neural-ODE for pharmacokinetics modeling and its advantage to alternative machine learning models in predicting new dosing regimens. *iScience, 24*(7), 102804. doi:10.1016/j.isci.2021.102804

Ly, T., Pamer, C., Dang, O., Brajovic, S., Haider, S., Botsis, T., . . . Ball, R. (2018). Evaluation of Natural Language Processing (NLP) systems to annotate drug product labeling with MedDRA terminology. *J Biomed Inform, 83*, 73-86. doi:10.1016/j.jbi.2018.05.019

Mager, D. E., Woo, S., & Jusko, W. J. (2009). Scaling pharmacodynamics from in vitro and preclinical animal studies to humans. *Drug Metab Pharmacokinet, 24*(1), 16-24. doi:10.2133/dmpk.24.16

Mason, M., Cho, Y., Rayo, J., Gong, Y., Harris, M., & Jiang, Y. (2022). Technologies for medication adherence monitoring and technology assessment criteria: Narrative review. *JMIR Mhealth Uhealth, 10*(3), e35157. doi:10.2196/35157

Mouchlis, V. D., Afantitis, A., Serra, A., Fratello, M., Papadiamantis, A. G., Aidinis, V., . . . Melagraki, G. (2021). Advances in de Novo Drug Design: From Conventional to Machine Learning Methods. *Int J Mol Sci, 22*(4). doi:10.3390/ijms22041676

National Academies of Sciences, E., and Medicine. (2021). *Innovations in Pharmaceutical Manufacturing on the Horizon: Technical Challenges, Regulatory Issues, and Recommendations*. Retrieved from https://doi.org/10.17226/26009.

National Institute of Standards and Technology. (2019). *U.S. Leadership in AI: A Plan for Federal Engagement in Developing Technical Standards and Related Tools*. Retrieved from https://www.nist.gov/system/files/documents/2019/08/10/ai_standards_fedengagement_plan_9aug2019.pdf

Negi, K., Pavuri, A., Patel, L., & Jain, C. (2019). A novel method for drug-adverse event extraction using machine learning. *Informatics in Medicine Unlocked, 17*, Article 100190. doi:10.1016/j.imu.2019.100190

Notice of Request for Information on Public and Private Sector Uses of Biometric Technologies. (2021-21975). (October 8, 2021). 86 Fed. Reg. 56300 Retrieved from https://www.federalregister.gov/documents/2021/10/08/2021-21975/notice-of-request-for-information-rfi-on-public-and-private-sector-uses-of-biometric-technologies

Organisation for Economic Co-operation and Development. (2019). *Recommendation of the Council on Artificial Intelligence*. Retrieved from https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

Pappalardo, F., Russo, G., Tshinanu, F. M., & Viceconti, M. (2019). In silico clinical trials: concepts and early adoptions. *Brief Bioinform, 20*(5), 1699-1708. doi:10.1093/bib/bby043

Pierce, C. E., Bouri, K., Pamer, C., Proestel, S., Rodriguez, H. W., Van Le, H., . . . Dasgupta, N. (2017). Evaluation of Facebook and Twitter monitoring to detect safety signals for medical products: An analysis of recent FDA safety alerts. *Drug Saf, 40*(4), 317-331. doi:10.1007/s40264-016-0491-0

Routray, R., Tetarenko, N., Abu-Assal, C., Mockute, R., Assuncao, B., Chen, H., . . . Mingle, E. (2020). Application of augmented intelligence for pharmacovigilance case seriousness determination. *Drug Saf, 43*(1), 57-66. doi:10.1007/s40264-019-00869-4

Sager, J. E., Yu, J., Ragueneau-Majlessi, I., & Isoherranen, N. (2015). Physiologically based pharmacokinetic (PBPK) modeling and simulation approaches: A systematic review of published models, applications, and model verification. *Drug Metab Dispos, 43*(11), 1823-1837. doi:10.1124/dmd.115.065920

Schmider, J., Kumar, K., LaForest, C., Swankoski, B., Naim, K., & Caubel, P. M. (2019). Innovation in pharmacovigilance: Use of artificial intelligence in adverse event case processing. *Clin Pharmacol Ther, 105*(4), 954-961. doi:10.1002/cpt.1255

Schneider, P., Walters, W. P., Plowright, A. T., Sieroka, N., Listgarten, J., Goodnow, R. A., . . . Schneider, G. (2020). Rethinking drug design in the artificial intelligence era. *Nat Rev Drug Discov, 19*(5), 353-364. doi:10.1038/s41573-019-0050-3

Schuler, A., Walsh, D., Hall, D., Walsh, J., Fisher, C., & for the Critical Path for Alzheimer's, D., the Alzheimer's Disease Neuroimaging Initiative and the Alzheimer's Disease Cooperative Study. (2021). Increasing the efficiency of randomized trial estimates via linear adjustment for a prognostic score. *Int J Biostat*. doi:10.1515/ijb-2021-0072

Shroff, T., Aina, K., Maass, C., Cipriano, M., Lambrecht, J., Tacke, F., . . . Loskill, P. (2022). Studying metabolism with multi-organ chips: new tools for disease modelling, pharmacokinetics and pharmacodynamics. *Open Biol, 12*(3), 210333. doi:10.1098/rsob.210333

Spiker, J., Kreimeyer, K., Dang, O., Boxwell, D., Chan, V., Cheng, C., . . . Botsis, T. (2020). Information visualization platform for postmarket surveillance decision support. *Drug Saf, 43*(9), 905-915. doi:10.1007/s40264-020-00945-0

Stehlik, J., Schmalfuss, C., Bozkurt, B., Nativi-Nicolau, J., Wohlfahrt, P., Wegerich, S., . . . Pham, M. (2020). Continuous Wearable Monitoring Analytics Predict Heart Failure Hospitalization: The LINK-HF Multicenter Study. *Circ Heart Fail, 13*(3), e006513. doi:10.1161/circheartfailure.119.006513

Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., . . . Zhao, S. (2019). Applications of machine learning in drug discovery and development. *Nat Rev Drug Discov, 18*(6), 463-477. doi:10.1038/s41573-019-0024-5

Viceconti, M., Pappalardo, F., Rodriguez, B., Horner, M., Bischoff, J., & Musuamba Tshinanu, F. (2021). In silico trials: Verification, validation and uncertainty quantification of predictive models used in the regulatory evaluation of biomedical products. *Methods, 185*, 120-127. doi:10.1016/j.ymeth.2020.01.011

Vought, R. T. (2020). *Guidance for Regulation of Artifical Intelligence Applications*. Retrieved from https://www.whitehouse.gov/wp-content/uploads/2020/11/M-21-06.pdf

Wang, S. V., Schneeweiss, S., Berger, M. L., Brown, J., de Vries, F., Douglas, I., . . . the joint ISPE-ISPOR Special Task Force on Real World Evidence in Health Care Decision Making. (2017). Reporting to improve reproducibility and facilitate validity assessment for healthcare database studies V1.0. *Pharmacoepidemiol Drug Saf, 26*(9), 1018-1032. doi:10.1002/pds.4295

Wang, W., Haerian, K., Salmasian, H., Harpaz, R., Chase, H., & Friedman, C. (2011). A drug-adverse event extraction algorithm to support pharmacovigilance knowledge mining from PubMed citations. *AMIA Annu Symp Proc, 2011*, 1464-1470. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/22195210

Wang, Y., Carter, B. Z., Li, Z., & Huang, X. (2022). Application of machine learning methods in clinical trials for precision medicine. *JAMIA Open, 5*(1), ooab107. doi:10.1093/jamiaopen/ooab107

Weissler, E. H., Naumann, T., Andersson, T., Ranganath, R., Elemento, O., Luo, Y., . . . Ghassemi, M. (2021). The role of machine learning in clinical research: transforming the future of evidence generation. *Trials, 22*(1), 537. doi:10.1186/s13063-021-05489-x

Zhang, X., Yan, C., Gao, C., Malin, B. A., & Chen, Y. (2020). Predicting Missing Values in Medical Data via XGBoost Regression. *J Healthc Inform Res, 4*(4), 383-394. doi:10.1007/s41666-020-00077-1