U.S. FOOD & DRUG ADMINISTRATION
CENTER FOR DRUG EVALUATION & RESEARCH
OFFICE OF CLINICAL PHARMACOLOGY

Application of Artificial Intelligence and Machine Learning in Drug Development and Precision Medicine

Hao Zhu, Ph.D., Mstat
Division of Pharmacometrics,
FDA/CDER/OTS/OCP

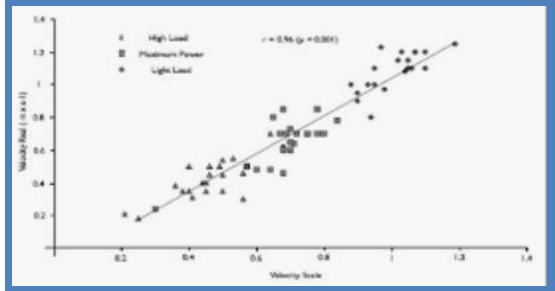
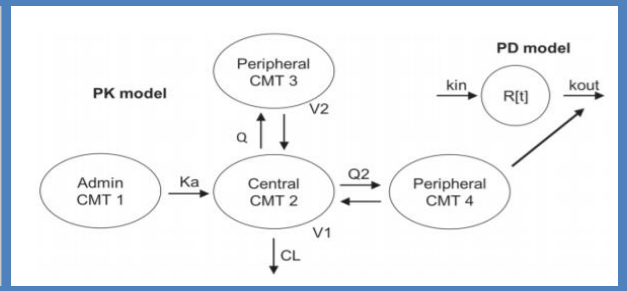
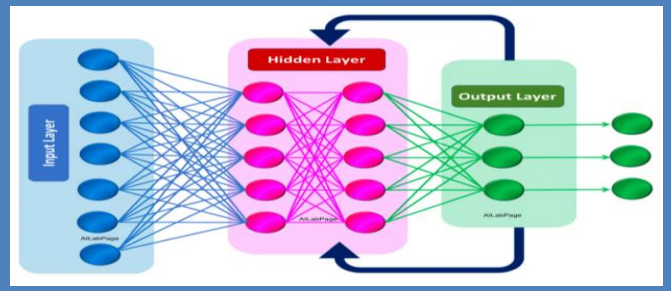
FDA/UMD CERSI Workshop
(February 2023)

Disclaimer: This presentation reflects the views of the presenter and should not be construed to represent those of the FDA or the United States Government.

Outline

- Introduction
  - AI/ML models
  - MIDD and AI/ML
  - AI/ML Review Function at OCP
- Landscape Analysis for AI/ML Related Submissions at the FDA
- Case Examples
  - Examples for AI/ML Related Submissions
  - Anakinra Review
- Examples of Ongoing Research Activities
- Take Home Messages

AI/ML Models

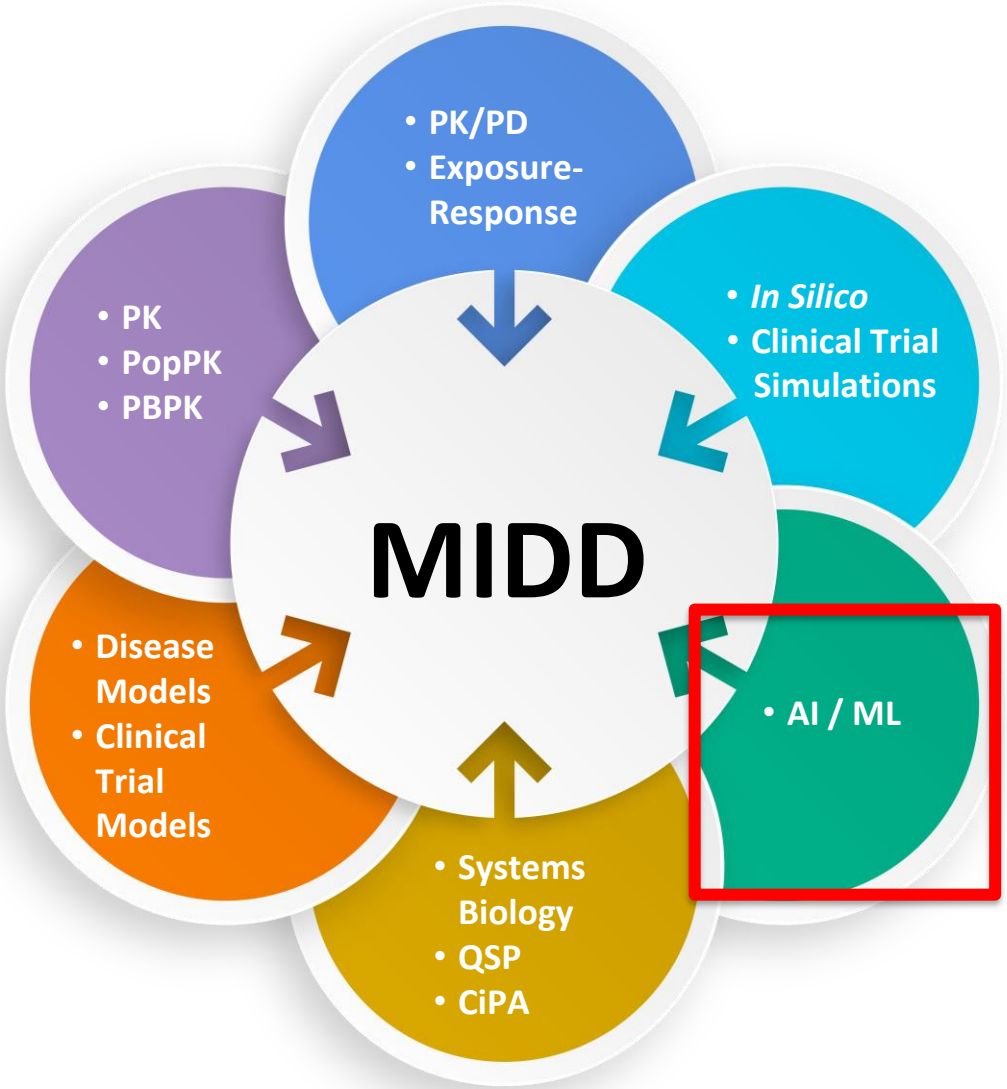
The Traditional Statistical/Pharmacometric Modeling	The Newer Machine Learning Approach
 	
Explicit model structure defined with domain knowledge and assumptions	Very general/flexible model structure, with less reliance on domain knowledge and assumptions
Tend to excel at handling data with a small or moderate number of independent variables	Tend to excel at handling data with a large number of independent variables (i.e., features), can also handle new types of data such as imaging
Easier to interpret/explain or draw statistical inference	Harder to interpret/explain or draw statistical inference; more focused on prediction performance

Courtesy of Dr. Qi Liu

MIDD and AI/ML

MIDD:

Development and application of exposure-based, biological, and statistical models derived from preclinical and clinical data sources to address drug development or regulatory issues*



AI/ML:

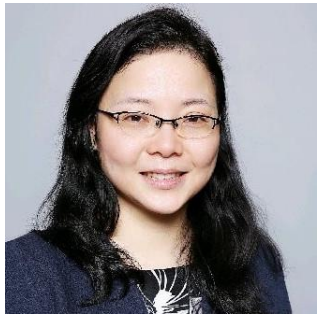
- A powerful tool to understand patient heterogeneity in drug response and improve precision medicine.
  - Links a large quantity of patient characteristics
- Potential Applications:
  - Inform enrichment design
  - Improve patient selection
  - Manage patient risks
  - Optimize dosing
  - Select appropriate treatment

* From PDUFA 6; Includes statistical designs involving complex adaptations, Bayesian methods, or other features requiring computer simulations to determine the operating characteristics of a confirmatory clinical trial.

Huang SM 2019 AAPS

AI/ML Review Function at OCP

OCP/IO



Dr. Qi Liu,
Associate Director for Innovation and Partnership at OCP

Dr. Menglun Wang,
AI/ML reviewer

Dr. Lixia Zhang,
AI/ML reviewer

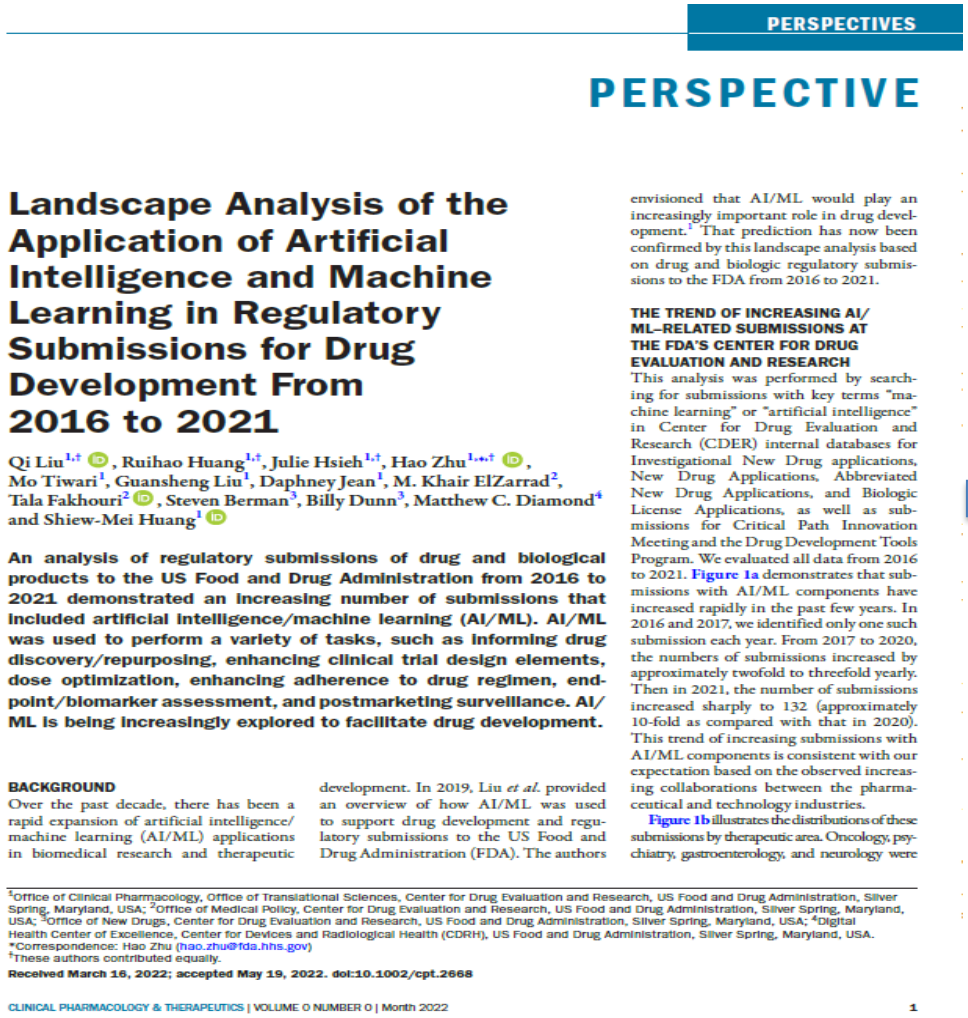
OCP/DPM

Dr. Hao Zhu,
Division Director, DPM

Dr. Ruihao Huang,
AI/ML reviewer

AI/ML Review Staff at OCP

Landscape Analysis for AI/ML Related Submissions



PERSPECTIVES

PERSPECTIVE

Landscape Analysis of the Application of Artificial Intelligence and Machine Learning in Regulatory Submissions for Drug Development From 2016 to 2021

Qi Liu[1,†], Ruihao Huang[1,†], Julie Hsieh[1,†], Hao Zhu[1,*,†], Mo Tiwari[1], Guansheng Liu[1], Daphney Jean[1], M. Khair ElZarrad[2], Tala Fakhouri[2], Steven Berman[3], Billy Dunn[3], Matthew C. Diamond[4] and Shiew-Mei Huang[1]

An analysis of regulatory submissions of drug and biological products to the US Food and Drug Administration from 2016 to 2021 demonstrated an increasing number of submissions that included artificial intelligence/machine learning (AI/ML). AI/ML was used to perform a variety of tasks, such as informing drug discovery/repurposing, enhancing clinical trial design elements, dose optimization, enhancing adherence to drug regimen, endpoint/biomarker assessment, and postmarketing surveillance. AI/ML is being increasingly explored to facilitate drug development.

BACKGROUND

Over the past decade, there has been a rapid expansion of artificial intelligence/machine learning (AI/ML) applications in biomedical research and therapeutic development. In 2019, Liu *et al.* provided an overview of how AI/ML was used to support drug development and regulatory submissions to the US Food and Drug Administration (FDA). The authors envisioned that AI/ML would play an increasingly important role in drug development.[1] That prediction has now been confirmed by this landscape analysis based on drug and biologic regulatory submissions to the FDA from 2016 to 2021.

THE TREND OF INCREASING AI/ML–RELATED SUBMISSIONS AT THE FDA'S CENTER FOR DRUG EVALUATION AND RESEARCH

This analysis was performed by searching for submissions with key terms "machine learning" or "artificial intelligence" in Center for Drug Evaluation and Research (CDER) internal databases for Investigational New Drug applications, New Drug Applications, Abbreviated New Drug Applications, and Biologic License Applications, as well as submissions for Critical Path Innovation Meeting and the Drug Development Tools Program. We evaluated all data from 2016 to 2021. Figure 1a demonstrates that submissions with AI/ML components have increased rapidly in the past few years. In 2016 and 2017, we identified only one such submission each year. From 2017 to 2020, the numbers of submissions increased by approximately twofold to threefold yearly. Then in 2021, the number of submissions increased sharply to 132 (approximately 10-fold as compared with that in 2020). This trend of increasing submissions with AI/ML components is consistent with our expectation based on the observed increasing collaborations between the pharmaceutical and technology industries.

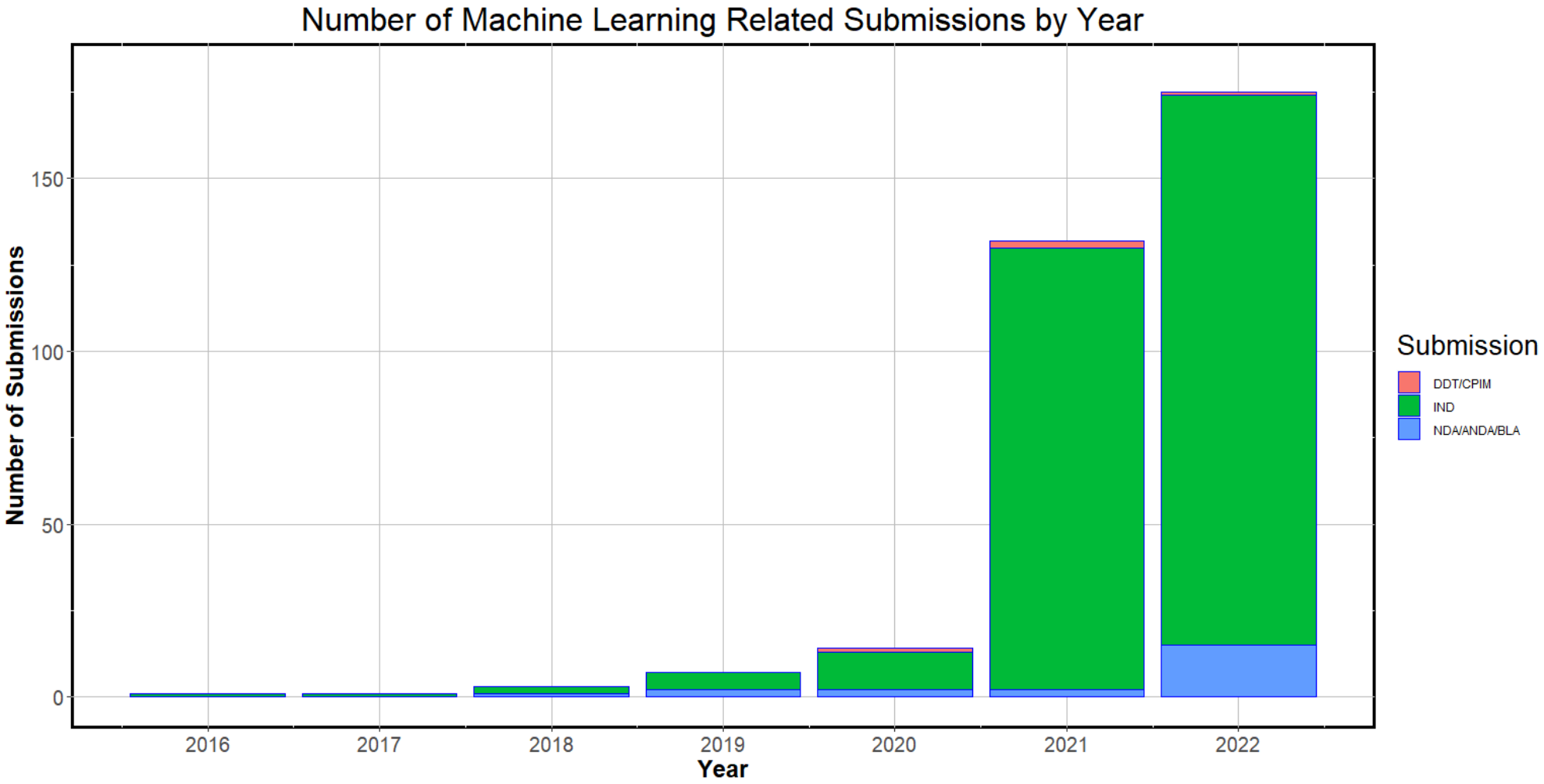
Figure 1b illustrates the distributions of these submissions by therapeutic area. Oncology, psychiatry, gastroenterology, and neurology were

[1]Office of Clinical Pharmacology, Office of Translational Sciences, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, Maryland, USA; [2]Office of Medical Policy, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, Maryland, USA; [3]Office of New Drugs, Center for Drug Evaluation and Research, US Food and Drug Administration, Silver Spring, Maryland, USA; [4]Digital Health Center of Excellence, Center for Devices and Radiological Health (CDRH), US Food and Drug Administration, Silver Spring, Maryland, USA.
*Correspondence: Hao Zhu (hao.zhu@fda.hhs.gov)
†These authors contributed equally.

Received March 16, 2022; accepted May 19, 2022. doi:10.1002/cpt.2668

- Update the analysis by including the submissions in year 2022
- Summarize the submissions by submission type (IND, NDA/BLA, etc.), development stage, and disease areas by year from 2016 to 2022

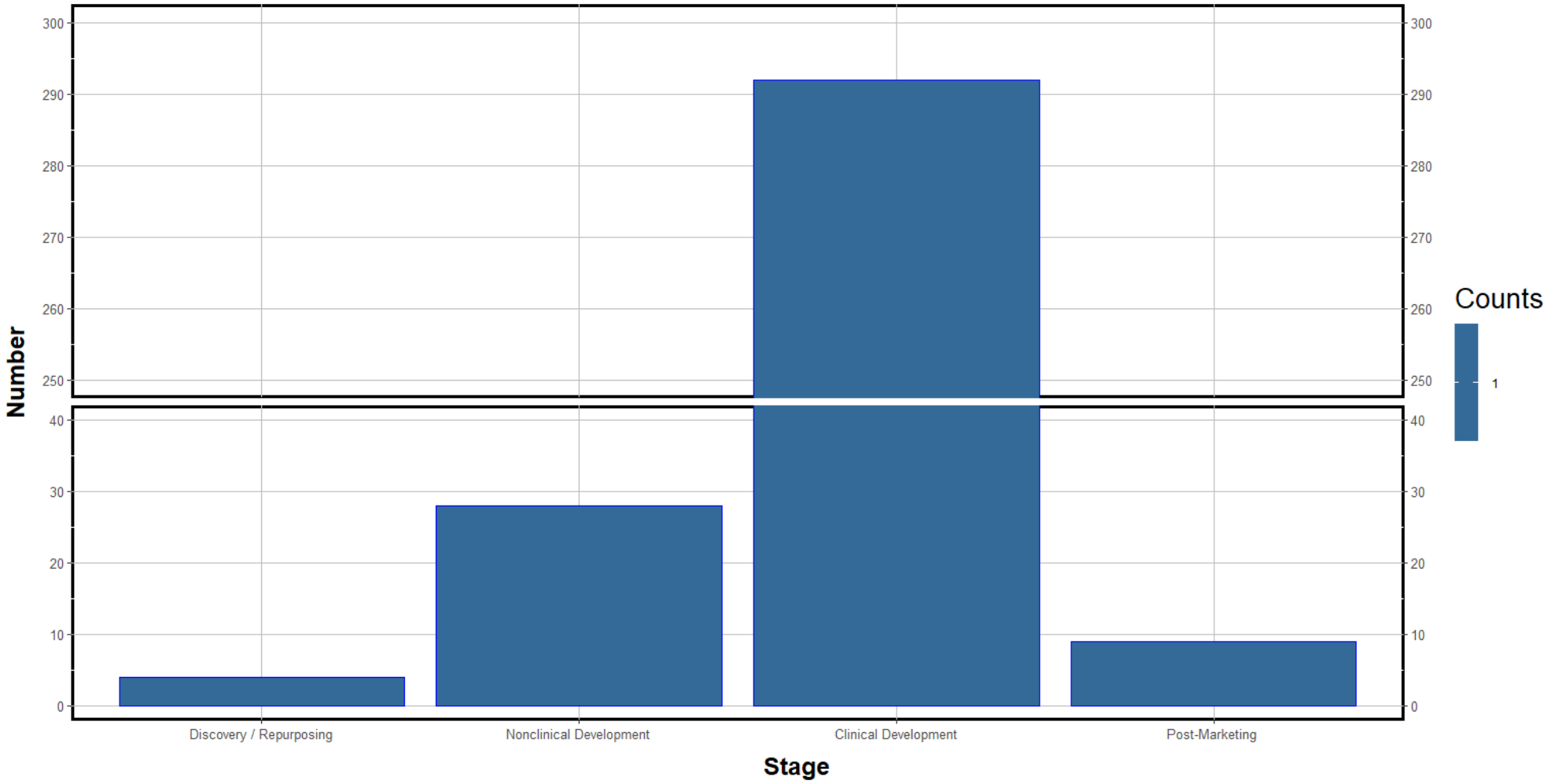
AI/ML Submissions over Time



2016 - 2022

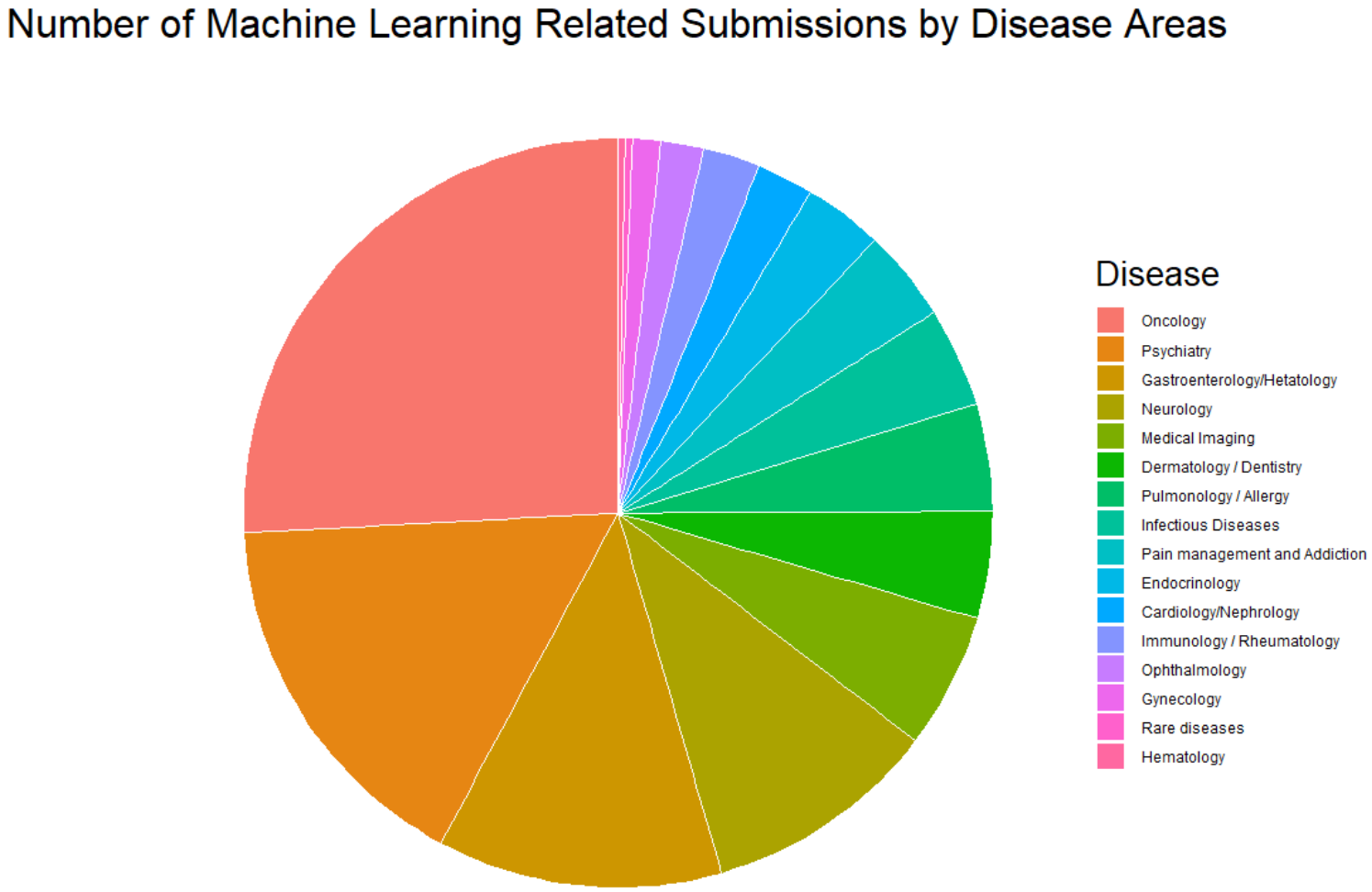
DDT: Drug Development Tool;
CPIM: Critical Path Innovation Meeting

AI/ML Submissions by Development Stage



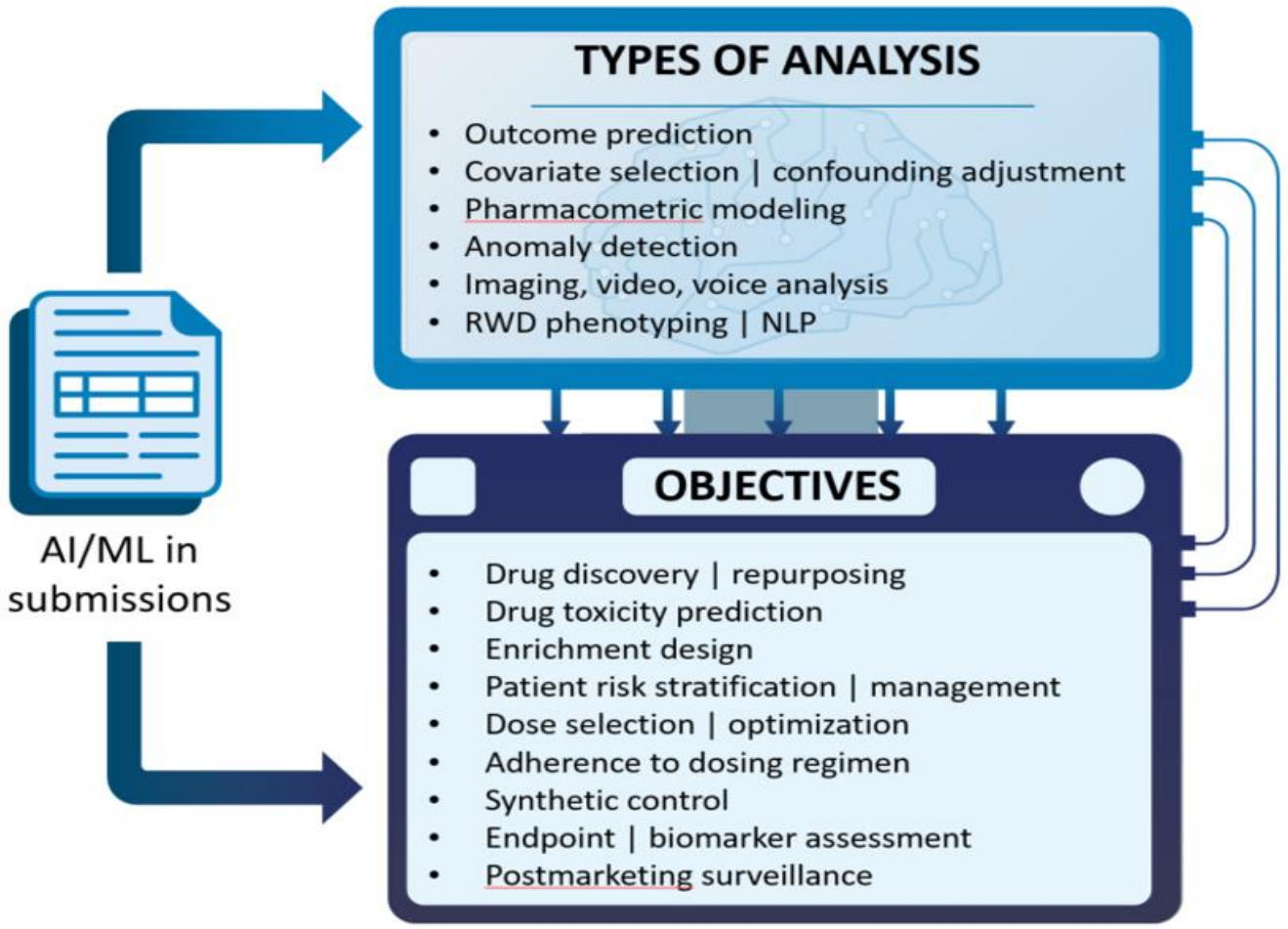
2016 - 2022

AI/ML Submissions by Therapeutic Areas



2016 - 2022

Types of AI/ML Related Analyses & Objectives



Review Case Examples

Examples in AI/ML related Submissions

- **Biomarker / Endpoint Assessment:**
  - AI algorithm is used to evaluate imaging-based biomarker as endpoint
- **Patient Selection:**
  - To use an AI-based diagnostic biomarker in conjunction with clinical assessment to enroll patients who are likely at a defined disease stage.
  - To enroll patients based on a companion diagnostic developed by using an AI algorithm linking EEGs, digital biomarkers detected through wearables, and other clinical symptoms.

----------INDICATIONS AND USAGE----------

KINERET is an interleukin-1 receptor antagonist indicated for:

Rheumatoid Arthritis (RA)

- Reduction in signs and symptoms and slowing the progression of structural damage in moderately to severely active rheumatoid arthritis, in patients 18 years of age or older who have failed 1 or more disease modifying antirheumatic drugs (DMARDs) (1.1)

Cryopyrin-Associated Periodic Syndromes (CAPS)

- Treatment of Neonatal-Onset Multisystem Inflammatory Disease (NOMID) (1.2)

Deficiency of Interleukin-1 Receptor Antagonist (DIRA)

- Treatment of Deficiency of Interleukin-1 Receptor Antagonist (DIRA) (1.3)

Source:* *US Package Insert of Kineret ®

14.1 Clinical Study in COVID-19

SAVE-MORE (NCT04680949) was a randomized, double-blind, placebo-controlled study to evaluate the safety and efficacy of Kineret in adult (≥18 years) patients with COVID-19 pneumonia who were at risk of developing severe respiratory failure (SRF), defined as pO2/FiO2 < 150 mmHg necessitating high flow oxygenation (HFO)/NIV/MV. All patients were hospitalized adults with COVID-19 pneumonia, radiologically confirmed by chest X-ray or CT, but had not progressed to SRF. All enrolled patients in this study were required to have a plasma soluble urokinase plasminogen activator receptor (suPAR) level ≥ 6 ng/mL. The suPAR assay is not commercially available for use in the United States, *[see Emergency Use Authorization (1.1)]*. Key exclusion criteria were pO2/FiO2 < 150 mmHg, requirement for NIV, requirement for MV, requirement for ECMO, and < 1500 neutrophils/mm^3. The mean age of participants was 61.9 years (standard deviation [SD] 12.1 years), and 57.9% were male.

Source:* *Fact Sheet of Kineret ®

Objectives

The objective for this analysis is to search for an alternative score based on commonly measured patient characteristics to identify patients with suPAR ≥ 6 ng/mL.

Source: Anakinra EUA review https://www.fda.gov/media/163546/download P-58

A machine learning classification model with elastic net regularization (ref: Regularization and Variable Selection via the Elastic Net) was used to select additional contributing features. A sensitivity analysis of feature importance was conducted via exploring model

A neural network-based model was applied to independently assess the importance of baseline NLR and its cut-off value. The neural network-based model is designed to simultaneously maximize sensitivity while maintaining positive predictive value (PPV) larger than or equal to 0.95 in model development. PPV is prioritized at the 0.95 level to ensure that patients selected by the determined score are more closely aligned with those in the SAVE-MORE trial. The cut-off value and the feature importance were assessed through the Gumbel-softmax technique (Ref: Categorical Reparameterization with Gumbel-Softmax). The neural network-based model confirmed that baseline NLR is an important feature, and the appropriate cut-off is around 7.

In summary, both neural network and elastic net regularization produced similar results, which suggested that an additional criterion of NLR ≥7 should be added. This is an inflammation and immune related biomarker. From a biological perspective, addition of this NLR based criterion is reasonable, as suPAR is an inflammation and immune related biomarker and none of the components of SURROGATE is clearly related to inflammation and immune system. This also appears to be a logical addition considering the

Source: Anakinra EUA review https://www.fda.gov/media/163546/download P-60

Evaluation Metrics

Based on the analysis objectives, the most critical metrics for evaluating the predictive performance of the selected scores are PPV and specificity. Sensitivity should also be considered as an additional important metric.

PPV, which is prevalence-dependent, is the probability that patients selected by a scoring rule (i.e., score positive) are patients with suPAR≥ 6 ng/mL at baseline. A high PPV is important to ensure that patients selected by a score rule are closely aligned with those enrolled in the SAVE-MORE trial.
Specificity, which is prevalence-independent, is defined as the probability that patients with suPAR < 6 ng/mL at baseline could be identified and rejected by a scoring rule (i.e., score negative). 1- specificity equals to false positive rate. A high specificity (or low false positive rate) ensures that patients with suPAR<6 ng/mL at baseline are less likely to be selected for anakinra treatment by a defined score.
Sensitivity is the probability that patients with suPAR≥6 ng/mL at baseline could be identified by a scoring rule. High sensitivity is preferable, as low sensitivity means some patients with baseline suPAR ≥ 6 ng/mL may not be identified.

Source: Anakinra EUA review https://www.fda.gov/media/163546/download ***P-61***

SCORE 2	suPAR>=6	suPAR<6	Total
positive, n	231 (**PPV=0.95**,**Sensitivity=0.41**)	12 (FPR=0.04)	243
negative, n	338	256 (NPV=0.43, **Specificity=0.96**)	594
Total,n	569	268	837

(B)

Training Data

SCORE 2	suPAR>=6	suPAR<6	Total
postive, n	95 (**PPV=0.94**,**Sensitivity=0.37**)	6 (FPR=0.07)	101
negative, n	159	76 (NPV=0.32, **Specificity=0.93**)	235
Total, n	254	82	336

(B)

Test Data

Source: Anakinra EUA review https://www.fda.gov/media/163546/download , P-62 Table 21 B, P-63 Table 22 B

Anakinra EUA Fact Sheet ***

1.1 Patient Population Identification

KINERET is authorized for emergency use for the treatment of COVID-19 in hospitalized adults with positive results of direct SARS-CoV-2 viral testing with pneumonia requiring supplemental oxygen (low- or high-flow oxygen) who are at risk of progressing to severe respiratory failure and likely to have an elevated plasma suPAR.

In the SAVE-MORE trial used to support the efficacy and safety of KINERET in COVID-19, key exclusion criteria were: pO2/FiO2 ratio < 150 mmHg, requirement for non-invasive ventilation (NIV), requirement for mechanical ventilation (MV), requirement for extra-corporeal membrane oxygenation (ECMO), and < 1500 neutrophils/mm^3.

All enrolled patients were required to have a plasma soluble urokinase plasminogen activator receptor (suPAR) level ≥ 6 ng/mL *[see Clinical Studies (14.1)]*. The suPAR assay is not commercially available in the United States. In order to identify a comparable population as was studied in the SAVE-MORE trial, an alternative patient identification method was developed to select patients most likely to have suPAR ≥ 6 ng/mL based on commonly measured patient characteristics. Patients meeting at least three of the following eight criteria are considered likely to have suPAR ≥ 6 ng/mL at baseline:

likely to have suPAR ≥ 6 ng/mL based on commonly measured patient characteristics. Patients meeting at least three of the following eight criteria are considered likely to have suPAR ≥ 6 ng/mL at baseline:

1. Age ≥ 75 years
2. Severe pneumonia by WHO criteria[1]
3. Current/previous smoking status
4. Sequential Organ Failure Assessment (SOFA)[2] score ≥ 3
5. Neutrophil-to-lymphocyte ratio (NLR) ≥ 7
6. Hemoglobin ≤ 10.5 g/dL
7. Medical history of ischemic stroke
8. Blood urea ≥ 50 mg/dL and/or medical history of renal disease

Source: Section 1.1 Fact Sheet of Kineret ®

Research Projects

- For information on our Machine Learning Precision Medicine fellowship, please visit https://www.zintellect.com/Opportunity/Details/FDA-CDER-2023-1224)

AI/ML Research in OCP

Landscape analyses

- Application of Machine Learning in Drug Development and Regulation: Current Status and Future Potential. (PMID: 31925955)
- Landscape Analysis of the Application of Artificial Intelligence and Machine Learning in Regulatory Submissions for Drug Development From 2016 to 2021. (PMID: 35707940)

Methodology Exploration

- Long short-term memory recurrent neural network for pharmacokinetic-pharmacodynamic modeling. (PMID: 33210994)
- A novel approach for personalized response model: deep learning with individual dropout feature ranking. (PMID: 33104924)
- Application of machine learning based methods in exposure-response analysis. (PMID: 35275315)
- Methods for preventing prediction for out-of-scope data
- Interpretable/Explainable ML

Application for Therapeutic Optimization/Individuation

- Ongoing research: Use ML to predict prognosis or treatment outcome (both efficacy and toxicity)
- Medical imaging data for precision medicine (in collaboration with CDRH, CBER and OCE)

● *Published*
● *Ongoing*

Examples of Ongoing Research Activities (1)

- **To improve clinical trial design**
  - Issue: The existence of placebo responders may attenuate the detectable efficacy signal and therefore decrease the chance for trial success.
  - Approach:
    - To Identify patient baseline characteristics predictive of placebo responders by using AI/ML models.
    - To apply the selected patient characteristics as additional exclusion criteria.
    - To evaluate the treatment effect after additional patients are excluded.
  - The treatment effect is increased after patients with baseline characteristics associated with a large placebo response being removed. This finding may provide additional insights for future clinical trial design.

Examples of Ongoing Research Activities (2)

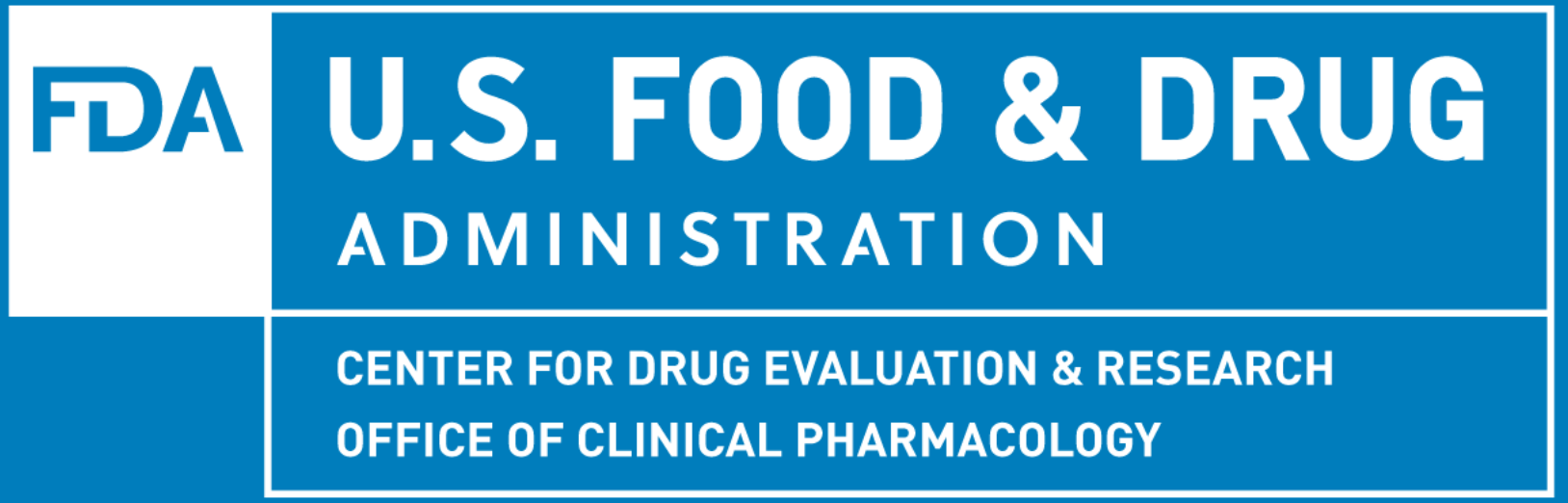
- **To improve precision medicine**
  - Issue: Patients may experience rare but severe adverse events.
  - Approach:
    - To Identify patient baseline characteristics that can be linked to the incidence of the rare adverse events by using AI/ML models.
    - To compare the outcomes between clinical trials and real-world data to cross validate the findings.
  - The identified patient characteristics may inform the appropriate selection of therapy and risk mitigation plan.

Take-Home Messages

- AI/ML, as a new addition to MIDD, is a powerful tool to improve drug development and patient care.
- The application of AI/ML to support drug development is expanding rapidly.
- With the capability to link various patient characteristics (e.g., demographic information, vital signs, and lab measurements), AI/ML tools may provide additional insights on precision medicine.

Acknowledgement

- Dr. Qi Liu
- Dr. Ruihao Huang
- Dr. Menglun Wang
- Dr. Lixia Zhang
- Dr. Jeffery Florian
- Dr. Kimberly Bergman
- Dr. Shiew-Mei Huang
- Dr. Issam Zineh
- Anakinra Review Team Members
- DPM colleagues
- OCP colleagues
- ORISE fellows


FDA
U.S. FOOD & DRUG
ADMINISTRATION
CENTER FOR DRUG EVALUATION & RESEARCH
OFFICE OF CLINICAL PHARMACOLOGY