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

- **Explicit model structure defined with domain knowledge and assumptions**
- Tend to excel at handling data with a small or moderate number of independent variables
- Easier to interpret/explain or draw statistical inference

### The Newer Machine Learning Approach

- **Very general/ flexible model structure, with less reliance on domain knowledge and assumptions**
- 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
- 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*

* 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.

**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

**Potential Models:**
- **PK/PD**
- **Exposure-Response**
- **In Silico**
- **Clinical Trial Simulations**
- **PK/PopPK/PBPK**
- **Disease Models**
- **Clinical Trial Models**
- **Systems Biology**
- **QSP/CiPA**
- **AI / ML**
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

- 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

Number of Machine Learning Related Submissions by Year

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

Number of Machine Learning Related Submissions by Disease Areas

Disease
- Oncology
- Psychiatry
- Gastroenterology/Hepatology
- Neurology
- Medical Imaging
- Dermatology/Dentistry
- Pulmonology/Allergy
- Infectious Diseases
- Pain Management and Addiction
- Endocrinology
- Cardiology/Nephrology
- Immunology/Rheumatology
- Ophthalmology
- Gynecology
- Rare Diseases
- Hematology

2016 - 2022
Types of AI/ML Related Analyses & Objectives

**TYPES OF ANALYSIS**
- Outcome prediction
- Covariate selection | confounding adjustment
- **Pharmacometric** modeling
- Anomaly detection
- Imaging, video, voice analysis
- RWD phenotyping | NLP

**OBJECTIVES**
- Drug discovery | repurposing
- Drug toxicity prediction
- Enrichment design
- Patient risk stratification | management
- Dose selection | optimization
- Adherence to dosing regimen
- Synthetic control
- Endpoint | biomarker assessment
- **Postmarketing** surveillance
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.
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/mm³. 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
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.

<table>
<thead>
<tr>
<th>SCORE 2</th>
<th>suPAR&gt;=6</th>
<th>suPAR&lt;6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>positive, n</td>
<td>231 (PPV=0.95, Sensitivity=0.41)</td>
<td>12 (FPR=0.04)</td>
<td>243</td>
</tr>
<tr>
<td>negative, n</td>
<td>338 (NPV=0.43, Specificity=0.96)</td>
<td>256</td>
<td>594</td>
</tr>
<tr>
<td>Total, n</td>
<td>569</td>
<td>268</td>
<td>837</td>
</tr>
</tbody>
</table>

Training Data

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

<table>
<thead>
<tr>
<th>SCORE 2</th>
<th>suPAR&gt;=6</th>
<th>suPAR&lt;6</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>positive, n</td>
<td>95 (PPV=0.94, Sensitivity=0.37)</td>
<td>6 (FPR=0.07)</td>
<td>101</td>
</tr>
<tr>
<td>negative, n</td>
<td>159 (NPV=0.32, Specificity=0.93)</td>
<td>76</td>
<td>235</td>
</tr>
<tr>
<td>Total, n</td>
<td>254</td>
<td>82</td>
<td>336</td>
</tr>
</tbody>
</table>

Test Data
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-corposoreal membrane oxygenation (ECMO), and < 1500 neutrophils/mm³.

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:

1. Age ≥ 75 years
2. Severe pneumonia by WHO criteria
3. Current/previous smoking status
4. Sequential Organ Failure Assessment (SOFA) 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
- 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