

AI 創薬の支援を目指した ADME データのキュレーションと予測モデル構築
ADME Data Curation and Constructing Prediction Models
for Supporting AI Drug Discovery

江崎 剛史
Tsuyoshi Esaki

滋賀大学データサイエンス・AI イノベーション研究推進センター
Data Science and AI Innovation Research Promotion Center, Shiga University

近年は人工知能 (Artificial Intelligence, AI) を始めとする情報科学手法の開発が進み、様々な分野で AI を活用した課題解決や業務改善への期待が膨らんでいる。創薬の分野も例外ではなく、DeepMind 社の AlphaFold に見られる成果もあり、今後ますます AI の貢献が期待される。

創薬の研究開発では、早期に疾患の原因である標的分子と結合する化合物を効率的に探索する情報科学手法の開発が進められている。しかし、結合すると推測された化合物が、医薬品としての適性を持つか否かを予測する手法の開発はあまり進んでいない。医薬品としての適性を評価するために、薬の体内動態 (Absorption, Distribution, Metabolism, Excretion) である ADME を考慮することが必要となるが、水溶性や代謝安定性、膜透過性など、考慮すべき項目が多いことに加え、各試験で異なる条件が設定されるため、予測モデルの構築にまとめて使えるデータは意外に多くない。機械学習を始めとした AI 技術で ADME を予測するモデルを構築するには、同じ実験条件で取得された大量のデータが必要となる。そのため、公共データベースの情報を十分に精査して、予測モデル構築に適したデータセットを作成する必要がある。

そこで、ADME の一つである *in vitro* のヒト肝ミクロソームによる代謝安定性試験の予測モデルを構築するために、公共データを独自に精査したデータセットを作成し、化合物の構造情報だけを使った予測モデルを構築した¹。その結果、精査して特定の実験条件だけに絞り込んだデータセットを使うことで、代謝安定性の予測精度が向上することを確認することができた。この研究から、AI 創薬を加速させるためには、実験条件の精査を行うことが重要であると痛感し、データの精査を簡略化するためのワークフローの構築を開始した。膜透過性など様々な ADME データに対してもキュレーションを行いながら、現在はこれらのデータセットを用いた ADME を予測するモデルの高精度化を目指すとともに、モデルの構築例も公開している²。

今後は、これらのデータセットを使用した予測モデルの構築を進めると同時に、創薬研究者が自由に ADME 予測モデルを構築できる環境を整備し、AI を活用した創薬の効率化に貢献したい。

【参考文献】

1. T. Esaki, R. Watanabe, H. Kawashima, R. Ohashi, Y. Natsume-Kitatani, C Nagao, K. Mizuguchi, Data Curation can Improve the Prediction Accuracy of Metabolic Intrinsic Clearance, *Mol. Inform.*, 2018, 38(1-2); 1800086
2. T. Esaki, T. Yonezawa, D. Yamazaki, K. Ikeda, Prediction Models for Fraction of Absorption and Membrane Permeability using Mordred Descriptors, *CBI J.*, 2022, 22; 46-54