

日本医療研究開発機構 医薬品等規制調和・評価研究事業 事後評価報告書

公開

I 基本情報

研究開発課題名: (日本語) 医薬品開発段階におけるバリデーション高度化を目指す対照化合物選定システムの開発

(英語) Development of reference compound recommendation system for advanced validation at the development stage of safety assessment test

研究開発実施期間: 令和4年8月1日～令和7年3月31日

研究開発代表者 氏名: (日本語) 水野 忠快
(英語) Tadahaya Mizuno

研究開発代表者 所属機関・部署・役職:
(日本語) 東京大学・大学院薬学系研究科・助教
(英語) Laboratory of Molecular Pharmacokinetics, Graduate School of Pharmaceutical Sciences, The University of Tokyo, Assistant Professor

II 研究開発の概要

研究開発の成果およびその意義等 (和文)

近年、技術の進展や動物実験に関する3Rs (Replacement, Reduction, Refinement) の理念の浸透に伴い、医薬品を含む化学物質の安全性を評価するための多様な手法が開発されてきた。これらの手法が、実際に公衆衛生の維持・向上に資する信頼性の高い評価技術として社会的に受容されるためには、その予測性能および信頼性を客観的に検証すること、すなわちバリデーション研究を通じた科学的な裏付けが不可欠である。

バリデーション研究の実施にあたっては、安全性評価手法の性能を適切に評価するための対照化合物 (reference compounds) の選定が極めて重要な役割を果たす。これらの化合物は、毒性強度、化学構造の多様性、物理化学的性質といった複数の観点を考慮し、専門家の判断によってバランスよく選定されてきた。特に、国際的な分類指針であるGHS (Globally Harmonized System) に準拠したデータベースを参照することが多く、化学空間の代表性と妥当性が確保されるよう設計されている。

しかしながら、研究開発初期、特にアカデミアにおける基礎研究段階では、こうした正式なバリデーション研究で求められる受け入れ基準との乖離が多く見られる。すなわち、評価法の性能検証に用いられる対照化合物の選定基準が十分に確立されておらず、その結果として初期段階で得られた予測性能の評価が信頼性を欠き、バリデーション段階に進んだ際に施設間再現性が低くなるといった課題が生じている。

これらの背景を踏まえ、本研究では、研究開発初期段階からバリデーション研究における受容基準を意識した対照化合物リストを、客観的かつ効率的に構築するためのデータ駆動型手法の開発を目的とした。具体的には、対照化合物の選定を多目的最適化問題として定式化し、オペレーションズリサーチ分野で広く活用されている遺伝的アルゴリズム (Genetic Algorithm: GA) を適用した。遺伝的アルゴリズムは、生物の進化過程を模倣したアルゴリズムであり、自然淘汰、突然変異、交叉などの原理を用いて、複雑な組合せ最適化問題に対して高い探索性能を発揮する。

本研究においては、化学構造、物理化学的性質、および毒性強度といった複数の観点から、化合物の多様性を同時に最大化するよう設計された目的関数を定義し、遺伝的アルゴリズムにより最適化を実施した。最適化された化合物リストについては、既存のバリデーション研究にて用いられた対照化合物リストやランダムに抽出された化合物リストと比較し、その多様性を多面的に評価した。また、*in silico* 毒性予測モデルを用いて、遺伝的アルゴリズムにより構築された化合物リストを適用した場合の予測性能を検証した。その結果、ランダムに生成された化合物リストを用いた場合と比較して、遺伝的アルゴリズムにより選定された化合物リストでは予測精度が有意に低下する傾向が見られた。この結果は、遺伝的アルゴリズムによって得られた化合物リストが、より厳密な検証セットとして機能し、モデルのロバスト性を評価するうえで有用であることを示唆している。

これらの成果を踏まえた本研究の主な貢献は、以下の3点である。

- ✓ 対照化合物選定問題を多目的最適化問題として定式化し、遺伝的アルゴリズムにより解くという新たな枠組みを提示した。
- ✓ 化学構造、物理化学的性質、毒性強度といった多基準を同時に最適化することで、バランスと多様性を兼ね備えた化合物リストを構築可能であることを実証した。
- ✓ 最適化された化合物リストが、モデル予測性能を意図的に低下させる“難易度の高い検証セット”として機能することを確認し、毒性予測モデルのロバスト性検証に有用であることを示した。

以上のように、本研究は、研究開発段階からバリデーション研究への移行を円滑にし、バリデーションの質と成功率を高めるための新たな支援技術として、データ駆動型の対照化合物選定手法を提案するものである。最終的な化合物選定においては依然として専門家の知見が不可欠であるが、本手法はその判断を補完する基盤として機能し得る。また、初期段階においてバリデーション研究に近い品質の検証を実施できることから、新規手法の実用化可能性を高める効果も期待される。

一方、本研究で用いた最適化アルゴリズムにはいくつかの制約が存在する。まず、化合物リストのサイズに関しては、既存のバリデーション研究との比較を容易にするため固定値を用いたが、今後の応用においては、課題の複雑性に応じて柔軟にサイズを変更可能とする必要がある。また、現実のバリデーション研究では、化合物の価格、入手可能性、そして実験の実施容易性といった要素が化合物選定に大きな影響を与える。さらに、各施設において安定して同様の測定結果が得られること、すなわち化合物の取扱いやすさや物理的安定性も重要な要素である。

しかしながら、これらの実務的な制約に関する信頼性の高い数値データは現在十分に整備されておらず、本アルゴリズムへの組み込みは困難であった。今後、これらの因子について十分な定量情報が整備された際には、目的関数あるいは制約条件として本アルゴリズムに統合することで、より現実的かつ実用的な最適化が可能になると考えられる。

本研究はバリデーションにて重要な対照化合物選定をシステム化して、専門家でなくとも活用可能にすることで、研究開発段階でのバリデーションを高度化する。これにより研究開発→バリデーション研究の流れが加速され、最先端技術による評価手法の開発・実用化サイクルが加速される。また本研究は、研究開発→バリデーション研究の流れを考慮し、下流の一部をシステム化して上流に移植することで流れを加速できることを訴える。レギュラトリーサイエンスとデータサイエンスの新たな接点を提唱することで、レギュラトリーサイエンスの強化に貢献し、医薬品等の品質・有効性・安全性評価手法の開発又検証を推進すると期待される。

Results and Significance of Research and Development

Recent advances in technology and the promotion of the 3Rs in animal experimentation have led to the development of numerous methods for assessing the safety of chemical substances, including pharmaceuticals. For these methods to be accepted as reliable tools for protecting public health, it is essential to conduct validation studies that objectively evaluate their predictive performance and reliability.

A crucial aspect of validation studies is the selection of reference compounds used to assess the performance of safety evaluation methods. These compounds are typically selected by experts based on multiple criteria, such as toxicity potency, chemical structures, and physicochemical properties, ensuring a balanced representation of the relevant chemical space. However, optimizing for one specific criterion, such as toxicity distribution, can result in trade-offs with others, such as structural diversity.

Furthermore, in the early stages of research and development—particularly in academia—reference compound selection often diverges from the criteria used in formal validation studies. As a result, performance evaluations conducted during these phases may lack sufficient reliability, and methods that appear promising in-house frequently demonstrate low inter-laboratory reproducibility when subjected to validation studies.

To address these issues, this study aimed to develop a data-driven and systematic approach to selecting well-balanced reference compounds that meet validation-level acceptance criteria. Specifically, we formulated the compound selection problem as a multi-objective optimization task and applied a Genetic Algorithm (GA), a widely used method in operations research. GA utilizes evolutionary principles such as natural selection, mutation, and crossover to efficiently explore complex combinatorial spaces.

The proposed GA-based framework simultaneously maximizes diversity in chemical structure, physicochemical properties, and toxicity profiles. We applied this method to generate optimized compound lists for validation assays and compared them against established reference lists from prior studies to evaluate diversity across multiple objectives.

Our results demonstrated that the GA-optimized compound lists exhibited significantly greater overall diversity than randomly generated lists. Moreover, when *in silico* toxicity prediction models were tested using these GA-optimized lists, predictive accuracy was notably reduced compared to that obtained using random compound selections. This reduction in performance reflects the stringency of GA-optimized lists and suggests their utility in rigorously assessing model robustness.

In summary, the key contributions of this study are as follows:

- ✓ We propose a novel framework that formulates reference compound selection as a multi-objective optimization problem solved using GA, introducing a new perspective in the field.
- ✓ Our method produces balanced and diverse compound lists by simultaneously optimizing multiple criteria.
- ✓ Compound lists closer to the Pareto front, representing high overall diversity, led to reduced predictive accuracy in toxicity models, indicating their suitability for stress-testing predictive model robustness.

This approach enables the generation of reference compound lists that support rigorous validation and may help bridge the gap between early-stage research and formal validation studies. By incorporating this method from the initial stages of method development, the quality and reproducibility of performance evaluations can be improved, potentially enhancing the likelihood of successful implementation in regulatory settings. Although expert input remains indispensable for final compound selection, the proposed framework provides a data-driven foundation to support efficient and objective list construction.

Several practical considerations were not incorporated into the current algorithm.

1. Size of Compound Lists

In this study, the size of compound lists was fixed to match the number of compounds used in existing validation assays, enabling direct comparison. However, allowing variable list sizes could expand the applicability of the method to more complex problems in the future.

2. Cost and Experimental Feasibility

Factors such as compound cost, availability, and experimental feasibility are important in real-world validation studies. For example, it is essential that compounds used in validation assays exhibit consistent measurement results across laboratories, which depends on handling ease and chemical stability. Due to the lack of reliable numerical data on these factors, they were not included in the current optimization framework. If such data become available, these criteria could be incorporated as additional objectives or constraints in future iterations of the algorithm.