

厚生労働行政推進調査事業費補助金（化学物質リスク研究事業）

国内外で開発されOECDで公定化されるNAMを活用した試験法の行政的な
受け入れに対応するための研究

令和6年度分担研究報告書

皮膚感作性*in silico*機械学習モデルのTG 開発

研究分担者 安部 賀央里

名古屋市立大学データサイエンス研究科 准教授

研究要旨

近年、動物実験に依存しない毒性評価として、New Approach Methodology (NAM) の開発が着目されている。また、経済協力開発機構 (OECD) では、有害性発現経路 (Adverse Outcome Pathway: AOP)に基づいて、*in vitro*、*in silico*、*in vivo* を組み合わせて化学物質の有害性を評価する Integrated Approaches to Testing and Assessment (IATA)の開発が進められている。化学物質の皮膚感作性の新規評価手法として、AOP に対応した複数の *in vitro* 試験法や *in silico* 手法を組み合わせた統合的アプローチが推奨されており、皮膚感作性試験の確定方式 (DASS: Defined Approach for Skin Sensitization) ガイドライン 497 が公定化されている。そこで、本分担研究では、皮膚感作性 *in silico* 機械学習モデルの OECD TG497 への追加収載のための基盤構築を目的とし、GHS 区分を判定する回帰モデルの改良を実施した。

研究協力者

足利 太可雄

国立医薬品食品衛生研究所 安全性生物試験研究センター ゲノム安全科学部 室長

徳永 朱莉

名古屋市立大学大学院薬学研究科

レギュラトリーサイエンス分野

ドポイントの一つである。従来、皮膚感作性はマウスを用いた *in vivo* 試験である局所リンパ節試験 (Local Lymph Node Assay: LLNA) 等で評価され、皮膚感作性を成立させる化学物質の閾値濃度である EC₃ 値をもとに定量的な評価が検討されている。しかし、近年、動物実験に依存しない毒性評価として、New Approach Methodology (NAM) の開発が着目されている。また、経済協力開発機構 (OECD) では、有害性発現経路 (Adverse Outcome Pathway: AOP)に基づいて *in vitro*、*in silico*、*in vivo* の情報を

A. 研究目的

皮膚感作性とは、化学物質等により接触性皮膚炎を引き起こす毒性のことであり、化学物質の安全性評価において重要なエン

組み合わせて化学物質の有害性を評価する試験と評価のための統合的アプローチ (integrated approaches to testing and assessment: IATA) および デファインドアプローチ (DA: Defined Approach, 定義済み総合判定方式) の開発が進められている。そのため、皮膚感作性の新規評価手法として、AOP に対応した複数の *in vitro* 試験法や *in silico* 手法を組み合わせた統合的アプローチが推奨されている。そこで、本分担研究では、機械学習を用いて GHS 区分 (1A, 1B) を判定する皮膚感作性 *in silico* モデルを開発し、さらにモデルの透明性や解釈性を考慮し QSAR モデル報告様式 (QSAR Model Reporting Format : QMRF) の作成を通して、OECD 皮膚感作性試験の DA であるガイドライン 497 への追加収載のための基盤構築を目的とする。皮膚感作性に関する新たな *in silico* NAM を日本国内で開発することにより、NAM の開発に大きく貢献し、さらには新規評価手法の公定化が達成されることにより我が国の化学物質の安全性評価への活用が期待される。

B. 研究方法

これまでに、名古屋市立大学と国立医薬品食品衛生研究所の足利太可雄にて開発した、皮膚感作性の強度指標であるマウスの LLNA EC3 値の定量的な予測モデル (Ambe et al., Regul Toxicol Pharmacol 125, 105019 (2021)) を改良した。具体的には、皮膚感作性を対象として機械学習による予測モデルの実用化や規制安全性評

価への受け入れに向けて、適用領域の設定や外部検証を実施した。また、モデルの解釈性・説明性を目指し、モデル作成に用いた説明変数の重要度の指標である SHAP (SHapley Additive exPlanations) 値を算出し予測結果の判断根拠を示した。学習データには皮膚感作性の組み合わせ評価のためのガイドライン (DASS, OECD No.497, 2021) に掲載されている化学物質のデータを用いた。予測モデルは機械学習を用いて決定木系アルゴリズムである XGBoost による回帰モデルを構築し、GHS 区分を行った。また、さらなる予測モデルの性能向上や精緻化を目指し、他の予測モデルとの比較や、学習データの追加に向けデータ準備を実施した。

(倫理面への配慮)

二次データを使用するため該当無し

C. 研究結果

モデル構築および内部検証用データとして皮膚感作性の組み合わせ評価のためのガイドライン (DASS, OECD No.497, 2021) に記載されているリストから154物質、外部検証用データとして化粧品業界団体により公開された文献 (Hoffmann, 2022) から72物質の LLNA 試験結果情報を使用した。LLNA EC3値を目的変数とし、決定木系アルゴリズムである XGBoost による回帰モデルを構築した。説明変数には、皮膚感作性の発症メカニズムに対応する3種類の *in vitro* 試験結果 (DPRA、KeratinoSensTM、h-CLAT)、OECD

QSAR ToolBoxやOPERAから算出した分子記述子、構造アラート情報から21個の変数を使用した。モデルの適用領域は、説明変数情報から、①3種類のin vitro試験結果のうち2種類以上が陽性、②k近傍法によるデータ間の距離をもとに設定した。また、各説明変数のSHAP値を算出し、モデルに対する重要度、および寄与の方向を可視化した。内部検証および外部検証データについて、適用領域内の物質を用いてモデルによる予測値が実測値の1/5倍から5倍の範囲内にある物質の割合を算出したところ、内部検証データでは80%であり、外部検証データでは65%であった。また、算出された各説明変数のSHAP値をもとに、予測モデルへの重要度を順位付けしたところ、使用した3種類全て(DPRA、KeratinoSensTM、h-CLAT)のin vitro試験結果が上位を占めた。さらに、in vitro試験において毒性が強いほど、皮膚感作性強度が強く予測され、正の相関が確認された。SHAPを用いることで、各物質の予測に寄与する変数が可視化できることや、その変数が正に寄与したか、負に寄与したかが確認できるため、モデルの予測結果を評価する際に有益な情報が提供できる。

また、モデルの予測性能を比較するために、外部検証用データに対して、安部らが2021年に構築したCatBoostモデル(Ambe et al., Regul Toxicol Pharmacol 125, 105019 (2021))や、資生堂のANNモデル(Hirota, et al., J. Appl. Toxicol., 35, 1333-1347 (2015))を用いて予測を行った。予測値が実測値の1/5倍から5倍の範囲内

にある物質の割合を算出したところ、本DASSモデルで36.1%、CatBoostモデルで44.4%、ANNモデルで33.3%となった。3つのモデル間で性能に大きな差はなかったことから、DASSモデルは既存の皮膚感作性強度指標を定量的に予測する機械学習モデルと比較し、予測性能が同等であることが示された。

次に、本DASS回帰モデル(XGBoostRegressor)にて算出されたLLNA EC3予測値を用いて、LLNA EC3予測値が2%以下であればGHS 1A、LLNA EC3予測値が2%より大きければGHS 1Bと判定した。内部検証のうち81物質についてGHS区分の一致を確認したところ、一致率79.0%、過大評価13.6%、過小評価7.4%であった。また、陰性については予測に使用した3種類のin vitro試験が3つとも陰性または2つ陰性の場合には陰性と判定した。これら陰性物質もあわせて計139物質の一致を確認したところ、一致率67.6%、過大評価7.9%、過小評価24.5%であった。

学習データの追加に関しては、外部検証に使用したデータや、文献情報から追加可能なデータの精査を実施中である。

D. 考察

SHAP値を用いた予測への重要度や寄与の方向の結果から、予測結果の判断根拠を確認でき、本モデルの解釈性を向上させた。皮膚感作性強度の予測においてAOPに関するin vitro試験情報を使用することは、予測

性能やモデルの解釈性向上に有用であることが示唆された。また、他のモデルとの比較から、本DASSモデルは予測性能も大きな違いがないことが確認された。さらに、適用領域の設定、モデルの汎化性能を評価する外部検証の実施、および予測結果に対する判断根拠を視覚的に表すことのできるSHAPの実装を行っているところが他のモデルに対して優位であると考えられる。これらの点は規制評価での使用を目指した定量的な皮膚感作性強度指標予測モデルとして有用性が高いと考えられる。

また、本モデルの有用性を確認するために、現在 OECD TG497 への追加が議論されている SARA-ICE モデルとの性能比較や、学習データの追加を検討する必要がある。

引き続き OECD Series on Testing and Assessment, No.386 (2023)の QMRF フォーマットを用いて QMRF の作成準備を進める。

健康危険情報

該当なし

E. 結論

皮膚感作性の発症メカニズムに基づく、解釈性の高い、皮膚感作性in silico機械学習モデルが構築できたと考えられる。本モデルは機械学習を活用したin silico NAMとして行政利用に向けた検討が期待できる。

F. 研究発表

F.1. 論文発表

1. [Ambe K.](#), Aoki Y, Murashima M, Wachino C, Deki Y, Ieda M, Kondo M, Hibi Y, Kimura K, Hamano T, Tohkin M. Prediction of cisplatin-induced acute kidney injury using an interpretable machine learning model and electronic medical record information. *Clinical and Translational Science*. 18(1), e70115 (2025).
2. [Ambe K.](#), Nakamori M., Tohno R., Suzuki K., Sasaki T., Tohkin M., Yoshinari K. Machine learning-based in silico prediction of the inhibitory activity of chemical substances against rat and human cytochrome P450s. *Chem. Res. Toxicol.* 37(11), 1843-1850 (2024).
3. Ashikaga T., Hatano K., Iwasa H., Kinoshita K., Nakamura N., [Ambe K.](#), Tohkin M. Next Generation Risk Assessment Case Study: A Skin Sensitization Quantitative Risk Assessment for Bandrowski's Base Existing in Hair Color Formulations. *Journal of Japanese Cosmetic Science Society*. 48(2), 73-77 (2024).
4. Watanabe T., [Ambe K.](#), Tohkin M. Streamlining Considerations for Safety Measures: A Predictive Model for Addition of Clinically Significant Adverse Reactions to Japanese Drug Package Inserts. *Biol Pharm Bull.* 47(3), 611-619 (2024).
5. Watanabe T., [Ambe K.](#), Tohkin M. Predicting the Addition of Information Regarding Clinically Significant Adverse Drug Reactions to Japanese Drug Package Inserts Using a Machine -Learning Model. *Ther Innov Regul Sci.* 58(2), 357-367 (2024).

F.2 学会発表

招待講演

1. Kaori Ambe Prediction of skin sensitization using machine learning JSAAE & ASCCT-ESTIV Joint Webinar (2024.12)
2. 安部賀央里 皮膚感作性評価における機械学習モデルの活用 スポンサーシンポジウム 革新的創薬支援ツールとしての New Approach Methodologies (NAMs) 最前線 CBI 学会 2024 年大会 (2024.10)(東京)
3. 安部賀央里 機械学習を活用したレギュラトリーサイエンス研究 講演データ解析を取り入れたレギュラトリーサイエンス研究の紹介 第 10 回次世代を担う若手のためのレギュラトリーサイエンスフォーラム (2024.9)(滋賀)
4. 安部賀央里 機械学習を用いた副作用予測に関する研究 シンポジウム 100 歳を超える健康設計への挑戦 第 25 回応用薬理シンポジウム (2024.9)(東京)
5. 安部賀央里 New Approach Methodologies (NAMs) を活用した皮膚感作性強度予測に向けた機械学習モデルの開発 フォーラム II New Approach Methodologies (NAMs) による毒性の評価 フォーラム 2024 衛生薬学・環境トキシコロジー (2024.9)(仙台)
6. 安部賀央里 AI/機械学習 in 毒性学～化学物質の毒性予測モデル開発～ワークショップ 2 ミライの毒性学 明日から役立つ最先端技術 第 51 回日本毒性学会学術年会 (2024.7)(博多)
7. 安部賀央里 機械学習を活用した薬

剤性腎障害の予測モデルの開発 ワークショップ 7 ICT は急性期の腎臓診療に変革をもたらすか? 第 67 回日本腎臓学会学術総会 (2024.6)(横浜)

8. 安部賀央里 データベースと機械学習を活用した副作用予測研究シンポジウム 5 医療 DX 時代を先導する「薬学×データサイエンス」の取り組み 第 26 回医薬品情報学会総会・学術大会 (2024.6)(千葉)

一般発表

1. 早川知里、安部賀央里、山田隆志、吉成浩一、頭金正博 機械学習によるヒトアセチルコリンエステラーゼ活性阻害の in silico 予測モデルの開発 日本動物実験代替法学会 第 37 回大会 (2024.11)(宇都宮)
2. 徳永朱莉、木下啓、岩佐帆乃夏、波多野浩太、中村伸昭、足利太可雄、安部賀央里 アミノ酸誘導体結合性試験 ADRA 等を使用した強感作性物質の皮膚感作性強度を予測する機械学習モデルの開発 日本動物実験代替法学会第 37 回大会 (2024.11)(宇都宮)
3. Kouki Maebara, Kyoko Ondo, Tomoaki Tochitani, Toru Usui, Izuru Miyawaki, Kaori Ambe A deep learning model for predicting chemical-induced rat hepatocellular necrosis using transcriptome data CBI 学会 2024 年大会 (2024.10) (東京)
4. 鈴木孝太郎、安部賀央里、有吉純平、神谷由紀子、頭金正博 ギャップマー型アンチセンスによるマウスでの血液中アラニンアミノトランスフェラーゼ上昇を予測するための in silico モデルの開発 第 51 回日本毒性学会学

術年会 (2024.7)(福岡)

5. 徳永朱莉、木下啓、安部賀央里、山田隆志、足利太可雄、頭金正博 機械学習を活用した皮膚感作性強度予測モデルの開発と解釈性向上に向けた検討 第 51 回日本毒性学会学術年会 (2024.7)(福岡)

G. 知的財産権の出願・登録状況

(予定を含む。)

1. 特許取得

該当なし

2. 実用新案登録

該当なし

3. その他

該当なし

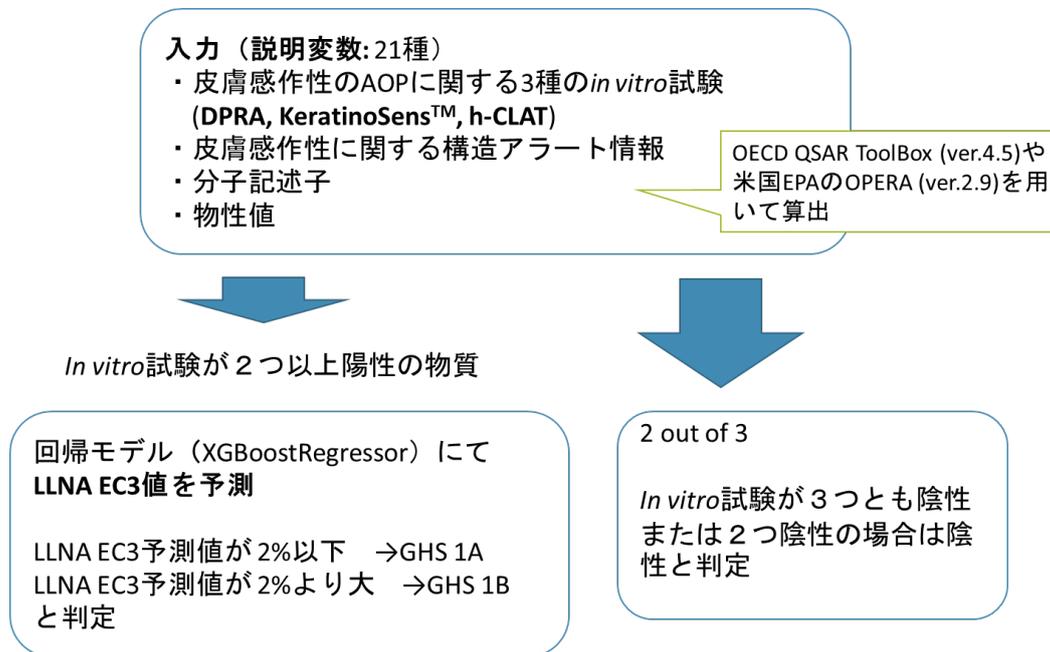


図1 皮膚感作性 *in silico* 機械学習モデルのフローチャート