

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）

*In silico*予測手法の高度化とNew Approach Methodologyの活用に基づく化学物質の
統合的ヒト健康リスク評価系の基盤構築に関する研究

令和3年度 分担研究報告書

機械学習を用いた皮膚感作性試験代替法の開発と
化学物質のリスク評価への活用に関する研究

研究分担者 安部 賀央里

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 講師

研究協力者 足利 太可雄

国立医薬品食品衛生研究所 安全性予測評価部 室長

研究協力者 頭金 正博

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 教授

研究協力者 鈴木 政晴, 伊藤 潤, 木下 啓, 村崎 亘, 片山 早紀, 中森 瑞季, 家田 維哉,
島田 侑裕, 土井 更良

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

研究要旨

動物実験の規制に伴い、皮膚感作性評価における動物実験代替法の開発が急務となっている。2021年には皮膚感作性を評価する際のDefined ApproachesがOECDのガイドラインNo.497に公表され、有害性評価や感作性強度評価において*in chemico*、*in vitro*、*in silico*を組み合わせた考え方が示されている。人工知能技術の一つである機械学習による皮膚感作性の予測モデル構築については、既存の文献情報からマウスの局所リンパ節試験における実験結果（皮膚感作性の強度指標）を予測する高性能な*in silico*モデルを開発した。開発した予測モデルを基に、予測性能の向上、適用範囲の拡大を目指して皮膚感作性に関するデータの追加による検討を行い、規制安全性評価への受け入れを目指した予測モデル開発を進めている。また、既にOECDのテストガイドラインに収載されている*in vitro*試験法について、化学構造情報等から実験結果を予測する*in silico*モデルの開発を行った。*In silico*化を進めることで、*in vitro*実験の効率化や欠損値対応にも貢献できる。さらに、化学物質の構造情報からヒトでの皮膚感作性の有無を判別する高性能な予測モデルの構築にも取り組んだ。

A. 研究目的

皮膚感作性物質は、繰り返しの暴露によって皮膚に局所的な炎症（かぶれ）が生じ

るアレルギー性接触皮膚炎（ACD）を引き起こす化学物質であり、T細胞を介したIV型過敏性免疫反応を誘発すると考えられて

いる。ACD は罹患者の多い疾患であり、生活の質に大きな影響を与えるため、皮膚感作性を引き起こす可能性のある化学物質を適切に評価し、規制することは極めて重要である。これまで、化学物質の皮膚感作性の評価のためのヒトおよび動物を使用する試験法が数多く開発されてきた。特に、マウス局所リンパ節アッセイ (LLNA) は、動物への苦痛が少なく、定量的な評価が可能なことから国際的に推奨されている。しかし、欧州で 2013 年以降、動物実験が実施された成分を含む化粧品の販売が禁止されたため、皮膚感作評価における動物実験代替法の開発が早急に求められている。また、動物実験にかかる費用、時間、動物愛護の観点から、化粧品業界のみならず、世界中の様々な業界において動物を用いない動物実験代替法の開発が加速している。

これまでに、皮膚感作の有害性発現経路 (AOP) の解明が進んでおり、2012 年に OECD ガイダンス文書として公表されている。皮膚感作性の動物実験代替法として、AOP における、各 Key Event に対応した *in chemico* や *in vitro* 手法が開発され、OECD のテストガイドラインに収載されている。一方で、皮膚感作性は免疫反応が関与する複雑なメカニズムであるため、複数の試験法を組み合わせる総合的に評価する Integrated Approach to Testing and Assessment (IATA) が提唱されており、複数の *in vitro* 試験結果や *in silico* の結果を組み合わせる統合的アプローチが検討されている。2021 年 6 月には皮膚感作性を評価する際の Defined Approaches が OECD のガイドライン No.497 に公表され、有害性評価や感作性強度評価において *in chemico*、*in vitro*、*in silico* を組み合わせた考え方が

示されている。

人工知能技術の一つである機械学習は、化学物質の毒性評価における *in silico* 手法の開発において重要な役割を示し、多種多様な毒性関連データを有効活用する手段として大きな期待が集まっている。我々は、機械学習を用いて、皮膚感作性の強度指標である LLNA の Estimated Concentration needed to produce a stimulation index of 3 (EC3) 値を予測する回帰モデルの構築に取り組んできた (Ambe *et al.*, 2021)。

本課題では、皮膚感作性に関する信頼性の高い *in vivo* や *in vitro* の実験結果を収集し、AOP の主要キーイベントを評価する *in vitro* の実験結果や物性値等の化学物質情報から *in vivo* での皮膚感作性を予測する機械学習モデルを構築する。既存の *in silico* モデルとの比較検討を行い、モデルの性能や適用範囲について評価する。また、開発した予測モデルの適用範囲の明確化や拡大に向けて、キーイベントを評価する *in vitro* 実験結果の *in silico* モデル化や、三次元の画像情報等を使用した予測モデルの開発にも取り組む。

さらに、他の毒性エンドポイントへの適用を視野に入れ、開発した予測モデルを化学物質のリスク評価に活用するにあたり、具体的な導入場面を想定したケーススタディを行い、予測結果に求められる信頼性や予測性能を検討し、安全性評価規制への受け入れへ向けた具体的な課題を整理する。

B. 研究方法

B.1 皮膚感作性の強度指標予測モデルの改良

これまでに我々が開発した皮膚感作性の強度指標 (LLNA EC3 値) の予測モデルでは、Hoffmann らの文献の the Cosmetics

Europe database (*Crit Rev Toxicol.* 2018; 48(5): 344-358) を学習データとして使用していた (Ambe *et al.*, 2021)。そこで、本課題では OECD の皮膚感作性に関する Defined Approach のガイドライン (No.497) において検討された化学物質の試験情報を使用してモデルの改良を行った。皮膚感作性の強度指標に使用されている LLNA の EC3 値を目的変数とし、皮膚感作性の AOP に関する *in vitro* 試験結果 (DPRA、KeratinoSens™、h-CLAT)、物性値、OECD QSAR ToolBox から得られる化学情報を説明変数とした。学習データは訓練データと検証データに分割し、勾配ブースティング決定木による回帰モデルを作成した。予測モデルの性能評価は検証データを用いて、決定係数 R^2 値により行った。

B.2 *In vitro* 実験の予測モデル構築

皮膚感作性評価において、複数の *in vitro* 実験結果を使用することは実験にかかる時間やコストの観点から効率的ではない。また、複数の *in vitro* 試験結果や *in silico* の結果を組み合わせ評価を行う統合的アプローチにおいても、*in vitro* 実験値を化学構造情報のみから予測することが可能になれば、欠損値の補完にも使用できるため有益なアプローチとなる。本課題では、皮膚感作性の AOP において MIE (タンパク質との共有結合) を評価する DPRA (OECD TG442C) と Key Event 2 (角化細胞における炎症性応答) を評価する KeratinoSens™ (OECD TG442D) に着目し、DPRA の実験結果から KeratinoSens™ の実験結果を予測する *in silico* モデルの構築を試みた。既存の文献情報から KeratinoSens™ の陽性物質、陰性物質を抽出し、説明変数を DPRA の実験値と

化学物質の構造情報である分子記述子とした。勾配ブースティング決定木系のアルゴリズムを使用して機械学習による判別モデルを構築した。

Hoffmann らの文献の the Cosmetics Europe database (*Crit Rev Toxicol.* 2018; 48(5): 344-358) に含まれる物質をモデル構築用の学習データとした。学習データは陽性物質と陰性物質の割合を保持したまま層化分割し、訓練データと検証データとした。訓練データを用いてパラメータ調整を行い、検証データの結果から最適なモデルを構築した。また、モデル作成に用いたデータソースとは異なる Urbisch のデータセット (*Regul Toxicol Pharmacol.* 2015; 71(2): 337-351) から外部検証データを作成し、構築した予測モデルの汎用性を確認した。説明変数として DPRA の実験値、MOE で計算した分子記述子を使用した。

B.3 ヒトの皮膚感作性評価の予測モデル構築

ヒトの皮膚感作性評価を *in silico* で予測する機械学習モデルの構築を目的とした。Golden らの文献 (*ALTEX.* 2021; 38(1): 33-48) に記載されているヒトでの皮膚感作性に関する情報を学習データとして利用した。目的変数をヒトの皮膚感作性の有無、説明変数を化学構造情報である分子記述子とし、機械学習アルゴリズムとして勾配ブースティング決定木である Catboost を用いて判別モデルを構築した。分子記述子は MOE により計算した。学習データは陽性物質と陰性物質の割合を保持したままの層化分割し、訓練データと検証データとした。訓練データを用いてパラメータ調整を行い、検証データの結果から最適なモデルを構築した。また、

モデル作成に用いたデータソースとは異なる Hoffmann らの文献の the Cosmetics Europe database (*Crit Rev Toxicol.* 2018; 48(5): 344-358) を用いて外部検証データを作成し、構築した予測モデルの汎用性を確認した。また、予測モデルに使用した分子記述子について、変数重要度を算出しモデルに寄与した情報について検討した。

(倫理面への配慮) 本研究は動物を用いた研究を行わないため対象外である。

C. 研究結果

C.1 皮膚感作性の強度指標予測モデルの改良

試験情報等が得られた 171 物質を訓練データと検証データに分割し、構築した回帰モデルにおいて、検証データ 30 物質を用いて評価したところ、決定係数 R^2 値が 0.8 程度となり高性能なモデルが得られた。回帰モデルは Python の XGBoost と CatBoost を使用した。また、本モデルでは、説明変数として OECD QSAR ToolBox のタンパク質結合アラート情報や *in vitro* 試験結果、物性値を用いており、モデルの変数重要度からもこれらの情報が EC3 値の予測に重要であることが分かった。本モデルは OECD のガイドラインにて使用された信頼性の高いデータを用いたことや、テストガイドラインに掲載された *in vitro* 試験法の結果を用いており、予測性能と解釈性を有した実用的な評価手法となり得ることが期待される。

C.2 *In vitro* 実験の予測モデル構築

In vitro 実験値の予測モデル化において Hoffmann(2018)らの文献に含まれ、かつ

SMILESが存在する122物質を学習データセットとした。訓練データと検証データに4:1の割合で分割し、XGBoostを用いて KeratinoSens™の陽性物質と陰性物質を判定するモデルを作成した。また、外部検証データとして Urbisch(2015)らの文献から、学習データである the Cosmetics Europe databaseとの重複物質削除を行った後の103物質を用いて、構築したモデルの汎化性能を確認した。MOEを用いて算出した分子記述子440種と、分子量、LogP、DPRAの実験値2種（システインペプチド減少率とリジンペプチド減少率）の計444種を説明変数とした。さらに変数重要度を算出し、重要度の高い説明変数32種を用いたところ、検証データセットのROC-AUCが0.892、感度0.871、特異度0.788であった。また、外部検証用データでは、ROC-AUCが0.820、感度が0.881、特異度が0.629であり、高い予測性能と汎化性能を示した。判別モデルはPythonのXGBoostClassifierを使用した。

学習データを用いて構築した判別モデルの変数重要度を確認したところ、最も重要な変数はDPRAの実験値であるシステインペプチド減少率であった。

C.3 ヒトの皮膚感作性評価の予測モデル構築

ヒトの皮膚感作性予測モデルにおいて、Golden(2021)らの文献に記載されている465物質から、SMILESを取得できなかった22物質と計算ソフトでSMILESを読み込めなかった6物質を除外し、437物質（陽性298物質、陰性139物質）をモデル構築のための学習用データセットとした。陽性物質と陰性物質の比を保持したまま3:1に層化分割したところ、訓練データが327物質、検

証データが110物質となった。一方で、Hoffmann(2018)らの the Cosmetics Europe databaseに記載されている128物質から、Golden(2021)らの文献に掲載されている物質との重複を除いた29物質を外部検証用データセットとした。外部検証用データセットは、陽性26物質、陰性3物質であった。MOEを用いて算出した435種の分子記述子のうち、変数重要度の高い上位53種を説明変数とし、ヒトの感作性の有無を目的変数とした。PythonのCatBoostClassifierを用いて皮膚感作性の有無を判別するモデルを構築した。また、k近傍法(k=3)による適用範囲を設定し、モデルの性能評価は適用範囲内の物質を用いて行った(図1)。判別モデルの性能評価として、検証用データセットのROC-AUCが0.81、Balanced Accuracyが0.71、感度0.91、特異度0.50であった。外部検証では、ROC-AUCが0.91、Balanced Accuracyが0.76、感度0.84、特異度0.67であり、高い汎化性能も確認できた。

学習データにより構築した判別モデルについて変数重要度を確認したところ、上位10個の中にLUMOに関する変数が含まれていた(表1)。

D. 考察

皮膚感作性の強度指標であるLLNAのEC3値を予測する回帰モデルの改良については、新たなデータとしてOECDの皮膚感作性に関するDefined Approachのガイドライン(No.497)において公開された化学物質の試験情報を使用した。本モデルの変数重要度からタンパク質結合アラート情報や*in vitro*試験結果が重要であることが分かった。本モデルはOECDが公開した信頼性の高いデータを用いたことや、ガイドライン

に掲載された*in vitro*試験法の結果を用いており、性能と解釈性を有した実用的な評価手法となり得ることが期待される。今後は、回帰モデルを使用したケーススタディを行い、モデルの適用範囲についても検討する必要がある。

*In vitro*実験値の予測モデル化において、KeratinoSens™の結果を予測する*in silico*モデルは既存の文献情報を有効活用し、高い予測性能を示した。*In vitro*実験は動物を使用しない代替法として有用だが、化学物質の入手や、*in vitro*実験において問題となる溶解性などの観点から、コストや時間がかかる。*In silico*での高性能な*in vitro*実験値予測モデルの開発は、皮膚感作性の評価において実験値が無い場合の穴埋めや、*in vitro*実験が必要かどうかのスクリーニングにも使用できる。また、予測モデルの変数重要度の結果から、KeratinoSens™の予測にはDPRAのシステインペプチド減少率が大きく関連していることが確認された。これらの情報はAOPの検証やさらなるメカニズムの理解にも有益だと考えられる。

ヒトの皮膚感作性予測モデルにおいて、化学物質の構造情報のみから、感作性の有無を精度良く判別する機械学習モデルが構築できた。動物実験代替法において、最終的な目標であるヒトでの有害事象の予測は重要なアプローチとなる。今後、変数重要度の検討や*in vitro*実験値を加えたモデルの構築が必要だと考えている。

E. 結論

機械学習を用いた*in silico*での皮膚感作性評価について、信頼性の高い学習データを用いた高性能な予測モデルの構築ができた。さらに、適用範囲の拡大やモデルの透明性、結果の説明性についても検討した。今後は

ケーススタディを行うことで、規制安全性評価への受け入れに向けた具体的な課題の整理が可能となる。また、*in vitro*実験についても*in silico*モデルを構築し、効率化に加えて実験値の欠損等の問題点にも取り組んでいる。ヒトでの皮膚感作性の有無を判別するモデル構築も開始し、皮膚感作性評価への統合的なアプローチが期待できる。

F. 研究発表

F.1. 論文発表

1. Ashikaga T, Ambe K, Suzuki M, Kurimoto M, Yamada T, Tohkin M. Establishment of a threshold of toxicological concern concept for skin sensitization by *in vitro/in silico* approaches. *Journal of Japanese Cosmetic Science Society*. **45**(4), 331-335, 2021.
2. Ambe K, Suzuki M, Ashikaga T, Tohkin M. Development of quantitative model of a local lymph node assay for evaluating skin sensitization potency applying machine learning CatBoost. *Regul. Toxicol. Pharmacol.* **125**, 105019, 2021.
3. Ambe K, Ohya K, Takada W, Suzuki M, Tohkin M. *In Silico* Approach to Predict Severe Cutaneous Adverse Reactions Using the Japanese Adverse Drug Event Report Database. *Clin. Transl. Sci.* **4**(2), 756-763, 2021.

F.2. 学会発表

1. 佐々木裕、三輪誠、安部賀央里、頭金正博：組み込み・除外判定を機械読解により実現した系統的レビュー。第 248 回自然言語処理研究会・第 226 回コンピュータビジョンとイメージメデ

ィア合同研究発表会 (2021 年 5 月 Web 開催)

2. 片山早紀、土井更良、立木孝幸、安部賀央里、頭金正博：*In vitro* 試験情報を利用した薬物性肝障害の *in silico* 予測手法の開発。第 48 回日本毒性学会学術年会 (2021 年 7 月神戸・Web 開催)
3. 中森瑞季、東野竜空、安部賀央里、頭金正博、佐々木崇光、吉成浩一：機械学習を用いた薬物代謝酵素阻害活性の *in silico* 予測手法の開発。第 48 回日本毒性学会学術年会 (2021 年 7 月神戸・Web 開催)
4. 吉井優花、秋田彩佑、安部賀央里、頭金正博：臨床試験データを用いた疾患修飾性抗リウマチ薬とメトトレキサート併用時の有効性と安全性に関する研究。医療薬学フォーラム 2021 (2021 年 7 月 Web 開催)
5. 家田維哉、安部賀央里、頭金正博：医薬品副作用自発報告データベースを用いた機械学習による薬剤性急性腎障害の予測手法の開発。医療薬学フォーラム 2021 (2021 年 7 月 Web 開催)
6. Takao Ashikaga, Kaori Ambe, Masaharu Suzuki, Masayuki Kurimoto, Takashi Yamada, Masahiro Tohkin: Establishment of a risk assessment method and threshold of toxicological concern (TTC) concept for skin sensitization by non-animal approaches. 11th World Congress on Alternatives and Animal Use in the Life Sciences 2020 (2021 年 8 月オンライン 1 年延期開催)
7. 安部賀央里：毒性データベースを用いた化学物質のインシリコ毒性予測。シンポジウム S5：計算機科学が拓く薬

- 学の未来、第 65 回日本薬学会関東支部大会 (2021 年 9 月オンライン)
8. Saki Katayama, Sarara Doi, Takayuki Tachiki, Kaori Ambe, Masahiro Tohkin, Toshimasa Jindo : Development of *in silico* prediction method for idiosyncratic drug-induced liver injury. CBI 学会 2021 年大会 (2021 年 10 月オンライン)
 9. 安部賀央里 : 化学物質の安全性評価における *in silico* 予測モデル. フォーカストセッション FS-07 創薬における WET 研究者と計算毒性学とのコラボレーション、および最新の安全性評価研究と Ames/QSAR 国際チャレンジプログラム、CBI 学会 2021 年大会 (2021 年 10 月オンライン)
 10. 秋田彩佑、魏捷、安部賀央里、頭金正博既存の臨床試験データを用いた直接経口抗凝固薬の有効性と安全性に関する民族差の検討. Young Investigator Award 第 31 回日本医療薬学会年会 (2021 年 11 月オンライン)
 11. 東野竜空、中森瑞季、安部賀央里、頭金正博、佐々木崇光、吉成 浩一 : 機械学習によるラット薬物代謝酵素および胆汁酸トランスポーターの阻害活性予測モデルの開発. 日本動物実験代替法学会第 34 回大会 (2021 年 11 月オンライン)
 12. 家田維哉、安部賀央里、頭金正博 : 医薬品副作用自発報告データベースを用いたアンサンブル機械学習モデルによる薬剤性急性腎障害の予測. 第 42 回日本臨床薬理学会学術総会 (2021 年 12 月仙台・ハイブリッド)
 13. 土井更良、安部賀央里、頭金正博 : 機械学習による薬物性肝障害の予測～大規模副作用報告データベースの活用～. 第 42 回日本臨床薬理学会学術総会 (2021 年 12 月仙台・ハイブリッド)
 14. 安部賀央里 : 機械学習法による特異体質性副作用の予測. シンポジウム 臨床薬理学と人工知能、第 42 回日本臨床薬理学会学術総会 (2021 年 12 月仙台・ハイブリッド)
 15. 中森瑞季、東野竜空、安部賀央里、頭金正博、佐々木崇光、吉成浩一 : 機械学習による薬物代謝酵素の阻害活性予測モデルの開発. 大学院生・学部生シンポジウム GS01 データ駆動型アプローチによる医薬品の毒性予測 日本薬学会第 142 年会 (2022 年 3 月オンライン)
 16. 伊藤潤、安部賀央里、足利太可雄、頭金正博 : ヒト皮膚感作性データを用いた機械学習による *in silico* 予測モデルの開発. 日本薬学会第 142 年会 (2022 年 3 月オンライン)
- G. 知的財産権の出願・登録状況
(予定を含む。)
1. 特許取得
なし
 2. 実用新案登録
なし
 3. その他
なし

K近傍法:

近傍k個のデータを用いてデータ密度を推定する方法

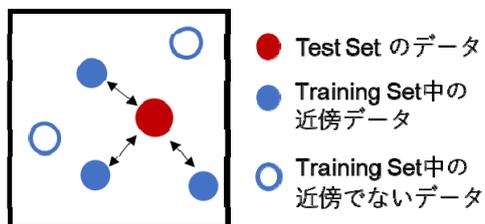


図1 適用範囲の設定方法 (k 近傍法)

表1 ヒトの皮膚感作性評価の予測モデルにおける上位10位の分子記述子

1	a_nN
2	AM1_IP
3	reactive
4	density
5	mutagenic
6	Q_VSA_FPPOS
7	a_ICM
8	dipoleY
9	AM1_LUMO
10	Q_VSA_PPOS