

In vitro/in silico 試験による皮膚感作性の毒性学的懸念の閾値 (TTC) コンセプトの確立

国立医薬品食品衛生研究所安全性予測評価部

足利 太可雄

Recently, several *in vitro* skin sensitization tests have been listed in the OECD guidelines. The threshold of toxicological concern (TTC) is a threshold value for human exposure that does not show any obvious adverse effects at the lower doses for all chemicals. In this study, we aimed to develop a quantitative prediction model using *in vitro/in silico* dataset and establish a TTC concept for skin sensitization.

The EC3 value, which is the endpoint of the LLNA, was used as the objective variable, and data on 120 substances were extracted from the dataset published by Cosmetics Europe. *In vitro* tests (DPRA, KeratinoSens™ and h-CLAT) data and physico-chemical properties were used as the explanatory variables. A quantitative prediction model for EC3 was developed using support vector regression (SVR), which is a machine learning approach. Predicted EC3 values were used to establish a no expected sensitization induction level (NESIL), and acceptable exposure level (AEL) for each chemical was calculated by dividing NESIL by sensitization assessment factor (SAF). Then by fitting gamma distribution of the AELs using a negative log (10) scale, 95 and 99 percentile probability were calculated as the Dermal Sensitization Threshold (DST) value. Finally, conversion of the DST to the threshold concentration of a women's face cream was done as an example of the application of this concept.

This prediction model was validated by 3-fold cross validation, and the accuracy of prediction of potency class in five categories was 45.8%. Assuming 20% of all chemicals to be skin sensitizers, the DST ($\mu\text{g}/\text{cm}^2$) for women's face cream was calculated as 0.129 (99 percentile) and 3.99 (95 percentile), respectively. Furthermore, the threshold concentration of this type of products was calculated as 0.008 % (for DST 99 percentile) and 0.26% (for DST 95percentile). This TTC concept will be useful for the safety evaluation of, for example, impurities in cosmetic ingredients as a non-animal approach.

1. 緒言

毒性学的懸念の閾値 (TTC : Threshold of Toxicological Concern) とは、あらゆる化学物質についてそれ以下の曝露量では明らかな有害影響が現れないとするヒト曝露の閾値として設定される値であり、その手法は過去の毒性試験データの統計学的解析により発展してきたものである。TTC手法は毒性未知の化学物質について毒性試験を行わずにリスクを評価するアプローチであり、対象物質の化学構造の情報とヒト曝露量に関する情報のみが必要とされるため、毒性学的情報は極めて限られているが曝露量が通常極めて低い化学物質群を包括的に評価するために有用であると考えられている。TTC手法はこれまで主に経口曝露の全身毒性について検討されており、香料¹⁾や医薬品の不純物²⁾など一部では公的規制においても受け入れられている。一方、経皮曝露による主要な毒性である皮膚感作性についても、SaffordによりTTCコンセプト (DST : Dermal Sensitization Threshold) が提案されて以来研究が進んでいる^{3,4)}。動物実験の廃止、原料調達が多様化、安全性保

証の厳格化など化粧品開発をとりまく最近の状況において、皮膚感作性におけるTTC手法は、原料中の不純物や天然物中の微量成分などヒト曝露量が非常に低いと考えられる物質に関する、毒性試験を伴わない包括的安全性評価法として強く期待されている。

最近 *in vitro* 皮膚感作性試験が相次いで OECD ガイドラインに収載され、*in silico* 試験法も含めた試験法の組み合わせについても OECD ガイダンスが公開された⁵⁾。また日欧の化粧品工業会は重要課題として *in vitro/in silico* 試験法による皮膚感作性のリスク評価に取り組んでいる^{6,7)}。さらに最近、日本化粧品工業連合会による日本人女性の化粧品使用量の調査⁸⁾があった。

こうした状況から、最近国際的主要化粧品メーカーでも皮膚感作性のTTCコンセプトの実用化が検討されている。そこで本研究では、従来用いられてきた *in vivo* 試験データの代わりに将来増加が見込まれる *in vitro/in silico* 試験データによる皮膚感作性のTTCコンセプトの確立を目指し、検討を行った。

2. 方法

2.1. データセットの確立

マウス皮膚感作性試験LLNAのエンドポイントであるEC3 (%)を精度よく予測可能なモデルを構築するために、Cosmetics Europeにより公表された、*in vivo/in vitro* データセット⁷⁾のうち全ての試験データが揃っている120化合物のデータを用いた。予測モデル作成のため、そのうち



Establishment of a threshold of toxicological concern (TTC) concept for skin sensitization by *in vitro/in silico* approaches

Takao Ashikaga

National Institute of Health Sciences, Division of Risk Assessment

90 化合物を training set、30 化合物を test set としてランダムに分割して使用した。目的変数は同一化学物質で複数の *in vivo* データがある場合、LLNA EC3 ($\mu\text{mol}/\text{cm}^2$) の中央値とした。説明変数は、*in vitro* 試験の DPRA, h-CLAT, KeratinoSensTM それぞれにおいて、Lys および Cys を含むペプチド減少率 (%), CV75 (μM) および MIT (μM)、EC1.5 (μM) および IC50 (μM) を用いた。さらに説明変数として、主に皮膚透過性に関わる被験物質の物理化学定数 (Molecular weight, LogP, LogS, Boiling point, Melting point) を用いた。データの欠損値処理として、LLNA 陰性物質の EC3 の中央値を 150% (後に $\mu\text{mol}/\text{cm}^2$ に単位変換)、DPRA ペプチド減少率が 0 のものを 0.001%、h-CLAT 陰性物質の MIT 値を 10,000 $\mu\text{g}/\text{mL}$ とした。さらに LLNA および *in vitro* 試験データについては、常用対数変換を行った。

2.2. 機械学習による強度 (EC3 値) 予測モデルの開発

説明変数から目的変数を予測するために非線形の機械学習モデルである Support Vector Regression (SVR) を選択した。データセットとして準備した 120 化合物について 3-fold cross validation による性能評価を行った。モデルの評価指標としては R^2 (決定係数) と RMSE (平均平方二乗誤差) を用いた。RMSE は値が小さいほど予測値と実測値の誤差が小さいことを示す。また、 R^2 は予測モデルが適切にデータを表現できるかを示し 1 に近づくほど良いモデルといえる。RMSE と R^2 の計算式は下記に記した。

$$\text{RMSE (root mean squared error)} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - f_i)^2}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - f_i)^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2}$$

y_i : 実験値

\bar{y}_i : 実験値の平均値

f_i : 予測値

2.3. 皮膚感作性 DST 値の設定

得られた予測 EC3 (%) 値を単位面積当たりの曝露量 ($\mu\text{g}/\text{cm}^2$) に変換後、既知の関係式³⁾からヒトにおける無感作誘導濃度 (NESIL: No Expected Sensitization Induction Level) を算出した。さらに、個人間変動、製剤系 (スキン

ケアやシャンプーなど) および使用状況 (髭剃り後など) から設定した感作評価係数 (SAF: Sensitization Assessment Factor)⁹⁾ で除することで許容曝露濃度 (AEL: Acceptable Exposure Level) を算出し、最終的に各化学物質の AEL を常用対数に変換した値について、ヒストグラムを作成した。市販の統計解析ソフトである JMP を用い、最尤法によりこのヒストグラムをガンマ分布にフィッティングした。このようにして得られた AEL の予測モデルと、*in vivo* 試験結果から得られた従来モデルを 2 種のパラメーター (形状母数と尺度母数) を指標に比較した。Negative Log (10) AEL の 99 パーセンタイルと 95 パーセンタイルの値から、DST (99 percentile) と DST (95 percentile) 値を算出した。さらに、代表的な化粧品の商品タイプである女性用フェースクリームを例として、本モデルと従来の動物実験による DST 値の比較を行った。最後に日本化粧品工業連合会の調査結果による製品タイプ別の使用量と曝露面積⁸⁾ から日本人女性におけるフェースクリームの曝露量 (CEL: Consumer Exposure Level) を算出し、CEL=DST 値とした場合の製品中の配合濃度閾値を計算した。

3. 結果

作成した EC3 値予測モデルの予測値と実測値の散布図を図 1 に示す。training set と test set の R^2 値の平均は、それぞれ 0.584, 0.585 であり、RMSE 値はそれぞれ 0.545, 0.653 であった。各 cross validation の R^2 と RMSE を表 1

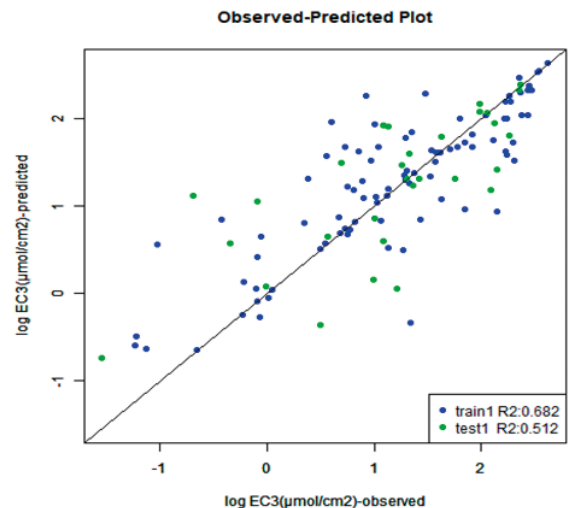


図1 EC3 ($\mu\text{mol}/\text{cm}^2$) の予測値と実測値の対応

表1 3-fold cross validationにおける R^2 と RMSE

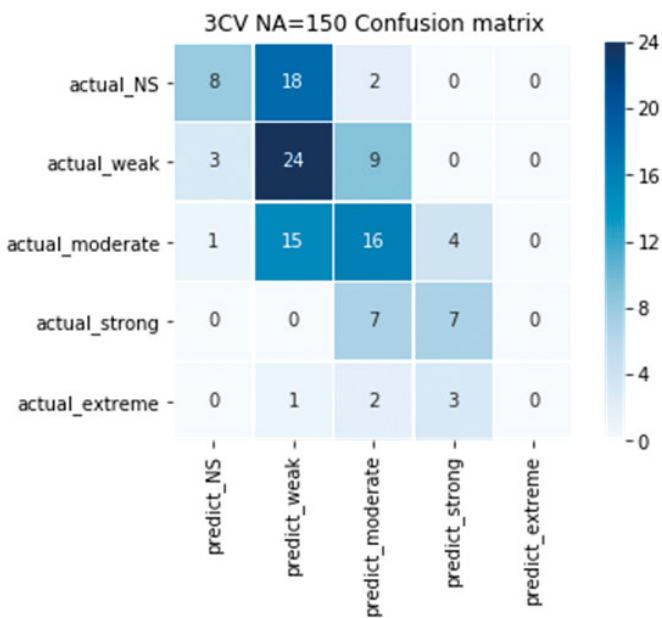
	R^2		RMSE	
	training set	test set	training set	test set
1 st -iteration	0.648	0.463	0.508	0.758
2 nd -iteration	0.561	0.623	0.536	0.632
3 rd -iteration	0.543	0.657	0.587	0.570
average	0.584	0.585	0.544	0.653

に示す。またカテゴリ分類(Extreme, Strong, Moderate, Weak, Non-Sensitization)における予測値と実測値の一致率は45.8%であった。120化学物質中、予測値と実測値のカテゴリが1区分以内のずれに収まったのは114化学物質であり、全体の90%以上であった。予測値と実測値のカテゴリごとの対応を表2に示す。

次に得られたEC3値予測モデルにより、LLNA陽性90物質についてEC3予測値を算出し、各種数値変換後、最終的にNegative Log(10) AELを算出した。なお、女性用フェースクリームのSAFはInternational Fragrance Association (IFRA)の基準に従って100とした⁹⁾。図2に、Negative Log(10) AELのヒストグラムおよびガンマ分布にフィッティングした結果を示す。得られたガンマ分布の形状母数と尺度母数はそれぞれ1.98、0.73であった。

全化学物質の20%が皮膚感作性陽性と仮定し、今回得られたガンマ分布から、女性用フェースクリームのNegative Log(10) AELの各パーセンタイルを算出した。99パーセンタイルと95パーセンタイルの値は、それぞれ0.901、-0.601であった。これらの値をAELに変換し、それぞれをDST値とすると、DST(99percentile)=0.126($\mu\text{g}/\text{cm}^2$)、DST(95percentile)=3.99($\mu\text{g}/\text{cm}^2$)であった。

表2 予測値と実測値のカテゴリごとの対応



予測と実測のカテゴリが一致した物質:
55 substances (45.8%)

予測と実測のカテゴリが1区分ずれた物質:
59 substances (49.1%)

予測と実測のカテゴリが2区分ずれた物質:
5 substances (4.1%)

予測と実測のカテゴリが3区分ずれた物質:
1 substances (4.1%)

さらに女性用フェースクリームの使用量(0.84g/day、日本人女性90percentile)と曝露面積(551cm²、日本人女性頭部面積の半分)から、CELを算出すると、1.52(mg/cm²/day)であった。そこでCEL=DSTとして、化学物質の女性用フェースクリームの配合濃度閾値を算出すると、DST(99percentile)において0.008%、DST(95percentile)において0.262%とそれぞれ算出された。

4. 考察

今回開発に用いた*in vivo*のLLNAデータは、Cosmetics Europeにより公表され、現在開発中であるOECD Defined Approachによる皮膚感作性試験テストガイドラインでも使用されている信頼性の高いものである。予測モデル開発において目的変数の質は大変重要であるため、現時点で最も信頼性の高いと考えられるデータセットを使用した。

開発したSVMによるEC3値予測モデルは、R²が0.585であり、カテゴリ分類の一致率も45.8%であったため、予測可能であると判断された。また、各R²とRMSEについてtraining setとtest setは大きな差がなく、汎化性能に問題ないと考えられた。EC3値予測モデルの予測性が良好と判断し、LLNAにおける陽性物質について予測EC3(%)値をNESILに変換後AELを算出し、常用対数に変換してNegative Log(10) AELを算出し、ヒストグラムを作成したところ、ガンマ分布にフィッティング可能と考えられた。ガンマ分布の2種のパラメーターである形状母数と尺度母数はそれぞれ、1.98、0.73であり、LLNAの実測値から求めたガンマ分布の形状母数と尺度母数はそれぞれ、2.57、0.57であったことから、今回開発したEC3値予測モデルによる予測値の強度分布は実測値のそれと類似していることが示された。

今回得られたDST(99 percentile)=0.126、DST(95 percentile)=3.99は、従来報告されたLLNAによるDST(99

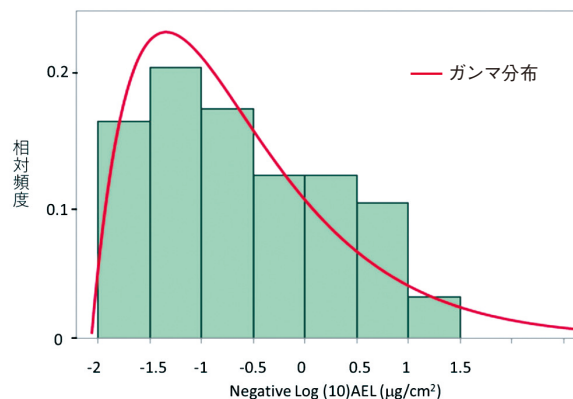


図2 Negative Log(10)スケールにおける皮膚感作性物質のAELの分布

percentile)=0.06、DST (95 percentile)=1.6³⁾と比較して約2倍であった。また皮膚感作のAOP(Adverse Outcome Pathway)において重要なKey Eventであるタンパク質との結合の観点からHigh Potency Category (HPC)に分類される皮膚感作性物質のDST (95 percentile)=1.5とする報告もある¹⁰⁾。今回用いた90品のLLNA陽性物質には、EC3 (%)の予測値がextremeに分類される0.1未満のものがなく、そのことが強度分布に影響を与えた可能性が考えられた。今後強度がextremeに分類される皮膚感作性物質をデータセットに追加するべきと考える。

今回試算したDST値は、Cosmetics Europeにより公表された120化合物のデータを用いたこと、作成したEC3予測モデルにより求めたNegative Log (10) AELがガンマ分布をとること、全化学物質の20%が皮膚感作性陽性であること、CEL=DSTとしたことなど様々な前提条件に基づいており、今後それらの妥当性に関して検証が必要である。さらに、今回acceptable riskのprobabilityは99パーセントイルと95パーセントイルを採用したが、どの値を用いるべきかについても議論が必要である。

また、TTCコンセプトは実務上様々な利点がある一方で、以下に示すデメリットも知られている¹¹⁾。

- ・化学物質の複合曝露の影響を考慮していない。
- ・データセットの量と質の限界
- ・TTCは確率に基づくものであり、すべての化学物質から守るものではない。
- ・適用範囲が限られており、タンパク質などトレーニングセットに入っていない不適格要件の明確化が必要

さらに、本DST値を用いる場合に注意すべき点として、ヒトにおける無感作誘導濃度(NESIL)を用いていることから明らかなように、あくまでも感作未成立状態における閾値であることが挙げられる。したがってすでに感作が成立している個体の場合、今回求めたDST値よりかなり少ない曝露量で皮膚炎が誘発する可能性がある。

以上の課題などから、皮膚感作性におけるTTCコンセプトは、アイデアとして提唱されているものの、未だ行政的規制の受け入れには至っておらず、今後上記の点も含めて幅広い議論が必要になるとと思われる。それらを踏まえた上で、今回得られたDST値は、適切な適用範囲のもとで、皮膚感作性試験を伴わない包括的安全性評価の基準になる可能性があると考えられた。

今回設定したDST値を用いて、製剤(女性用フェースクリーム)中の化学物質の配合濃度閾値を算出した。経皮適用する製剤は、サンスクリーンなど全身に使用するもの、シャンプーなど洗い流すものなど様々であり、今回例として挙げた女性用フェースクリーム以外の製品タイプについても、配合濃度閾値を算出してその妥当性を議論する必要がある。また、製品使用量は、日本化粧品工業連合会の調

査による90パーセントイルの値を用いた。製品使用量や曝露面積は化粧習慣や人種により違いがあると考えられるため、同様の調査結果がある欧州のSCCSガイダンス¹²⁾なども参照し、設定条件による違いを確認すべきと考える。

本研究では、機械学習であるSVRを用いて強度予測モデルを作成した。これまでも機械学習による皮膚感作性の強度予測モデルはいくつか開発されており¹³⁻¹⁵⁾、最近アメリカ環境保護庁(EPA)において、Artificial Neural Network¹⁶⁾による皮膚感作性強度予測モデルによるイソチアゾリノン系防腐剤のリスク評価が行われたことは注目に値する。今後こうした機械学習による皮膚感作性の強度予測モデルによる原料の配合許容濃度予測の実用化が期待されている。

5. 総括

化粧品開発における動物実験が困難になる中、皮膚感作性のTTCコンセプト(DST)の実用化が急務となっている。本コンセプトが確立されれば、閾値未満の曝露量の範囲で安全性試験を行わずに、被験物質(例えば化粧品原料)の原料情報と配合濃度のみで安全性評価が可能になることが期待できるためである。具体的な活用の例として、動物実験以外では特に評価が困難とされる、化粧品原料中の不純物や天然物中の微量成分の安全性評価に効率的で有用な手法となることが挙げられる。皮膚感作性のTTCコンセプトを確立するためには、多様な感作性陽性物質からなる信頼性の高いデータセットが鍵となるが、LLNAなど*in vivo*のデータセットの拡充が期待できない現状を考えると、*in vitro*試験のデータセットの拡充に期待したい。

今回機械学習の手法であるSVRの採用により、予測精度および頑健性の良好な強度予測モデルの開発が可能となった。本モデルにより非動物実験結果による定量的リスク評価(QRA)が可能になり、製品中の配合許容濃度の算出などへの応用が期待できる。

いずれにしても化粧品開発における動物実験が困難になりつつある現在、曝露量が極めて少ない場合に限られるものの、*in vitro*試験さえも不要となる皮膚感作性のTTCコンセプトの確立はニーズが高いと考えられ、本研究がそのきっかけとなれば幸いである。

謝辞

本研究は、名古屋市立大学大学院薬学研究科 医薬品安全性評価学分野および国立医薬品食品衛生研究所 安全性予測評価部第四室との共同研究により行われたものであり、関係者の皆様に深く感謝いたします。

また本研究の遂行にご支援をいただきました公益財団法人コーセー化粧品研究財団に篤く御礼申し上げます。

(引用文献)

- 1) EFSA: Scientific Opinion on Exploring options for providing advice about possible human health risks based on the concept of Threshold of Toxicological Concern (TTC), *EFSA Journal*, 10 (7), 2750 (2012).
- 2) L. Müller, R.J. Mauthe, C.M. Riley, M.M. Andino, D. De Antonis, C. Beels C, A rationale for determining, testing, and controlling specific impurities in pharmaceuticals that possess potential for genotoxicity, *Regulatory Toxicology and Pharmacology*, 44, 198-211 (2006).
- 3) R. J. Safford, he Dermal Sensitization Threshold –A TTC approach for allergic contact dermatitis, *Regulatory Toxicology and Pharmacology*, 51, 195-200 (2008).
- 4) R. J. Safford, A. O. Aptula, N. Gilmour, Refinement of the Dermal Sensitization Threshold (DST) approach using a larger dataset and incorporating mechanistic chemistry domains, *Regulatory Toxicology and Pharmacology*, 60, 218-224 (2011).
- 5) Guidance Document on the Reporting of Defined Approaches and Individual Information Sources to be Used within Integrated Approaches to Testing and Assessment (IATA) for skin sensitisation, https://www.oecd-ilibrary.org/environment/guidance-document-on-the-reporting-of-defined-approaches-and-individual-information-sources-to-be-used-within-integrated-approaches-to-testing-and-assessment-iata-for-skin-sensitisation_9789264279285-en
- 6) O. Takenouchi, S. Fukui, K. Okamoto, S. Kurotani, N. Imai, M. Fujishiro, D. Kyotani, Y. Kato, T. Kasahara, M. Fujita, A. Toyoda, D. Sekiya, S. Watanabe, H. Seto, M. Hirota, T. Ashikaga, M. Miyazawa, Test battery with the human cell line activation test, direct peptide reactivity assay and DEREK based on a 139 chemical data set for predicting skin sensitizing potential and potency of chemicals, *Journal of Applied Toxicology*, 35 (11), 1318-1332 (2015).
- 7) S. Hoffmann, N. Kleinstreuer, N. Alépée, D. Allen, A.M. Api, T. Ashikaga, E. Clouet, M. Cluzel, B. Desprez, N. Gellatly, C. Goebel, P. S. Kern, M. Klaric, J. Kühnl, J. F. Lalko, S. M-Teissier, K. Mewes, M. Miyazawa, R. Parakhia, E. van Vliet, Q. Zang, D. Petersohn, Non-animal methods to predict skin sensitization (I) : the Cosmetics Europe database, *Critical Reviews in Toxicology*, 48 (5), 344-358 (2018).
- 8) M. Yamaguchi, D. Araki, T. Kanamori, Y. Okiyama, H. Seto, M. Uda, M. Usami, Y. Yamamoto, T. Masunaga, H. Sasa, Actual consumption amount of personal care products reflecting Japanese cosmetic habits, *J. Toxicol. Sci.*, 42 (6), 814 (2017).
- 9) A. M. Api, D. A. Basketter, P. A. Cadby, M-F. Cano, G. Ellis, G. F. Gerberick, P. Griem, P. M. McNamee, C. A. Ryan, R. J. Safford, Dermal sensitization quantitative risk assessment (QRA) for fragrance ingredients, *Regulatory Toxicology and Pharmacology*, 52, 3-23 (2008).
- 10) T. Nishijo, A. M. Api, G. F. Gerberick, M. Miyazawa, D. W. Roberts, R. J. Safford, H. Sakaguchi, *Regulatory Toxicology and Pharmacology*, 117, 104732 (2020).
- 11) M.N. Mons, M.B. Heringa, J. van Genderen, L.M. Puijker, W. Brand, C.J. van Leeuwen, Use of the Threshold of Toxicological Concern (TTC) approach for deriving target values for drinking water contaminants. *Water Res.*, 47 (4), 1666-78 (2013).
- 12) THE SCCS'S NOTES OF GUIDANCE FOR THE TESTING OF COSMETIC SUBSTANCES AND THEIR SAFETY EVALUATION 8TH REVISION, http://ec.europa.eu/health/scientific_committees/consumer_safety/docs/sccs_s_006.pdf
- 13) J. Jaworska, A. Natsch, C. Ryan, J. Strickland, T. Ashikaga, M. Miyazawa, Bayesian integrated testing strategy (ITS) for skin sensitization potency assessment: A decision support system for quantitative weight of evidence and adaptive testing strategy, *Archives of toxicology* 89 (12), 2355-2383 (2015).
- 14) M. Hirota, T. Ashikaga, H. Kouzuki, Development of an artificial neural network model for risk assessment of skin sensitization using human cell line activation test, direct peptide reactivity assay, KeratinoSens™ and in silico structure alert parameter, *J. Appl. Toxicol.*, 38 (4), 514-526 (2018).
- 15) H. Li, J. Bai, G. Zhong, H. Lin, C. He, R. Dai, H. Du, L. Huang, Improved defined approaches for predicting skin sensitization hazard and potency in humans, *ALTEX*, 36 (3), 363-372 (2019).
- 16) Hazard Characterization of Isothiazolinones in Support of FIFRA Registration Review, <https://www.linkedin.com/pulse/us-epa-hazard-characterization-isothiazolinones-support-passman>