

平成 29 年度 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）
分担研究報告書

研究課題名：化学物質のヒト健康リスク評価における(定量的)構造活性相関および
カテゴリーアプローチの実用化に関する研究

分担研究課題名：反復投与毒性を指標にした構造活性相関モデルに関する研究

研究分担者 広瀬 明彦 国立医薬品食品衛生研究所 安全性予測評価部 部長
研究協力者 松本 真理子 国立医薬品食品衛生研究所 安全性予測評価部研究員
研究協力者 川村 智子 国立医薬品食品衛生研究所 安全性予測評価部研究員
研究協力者 Sebastien Guesne ラーサ研究所
研究協力者 Lilia Fisk ラーサ研究所

研究要旨

H29 年度は、前年度までに改良してきた肝毒性に関連するプロファイラーを基にして作成した KNIME ワークフローに、ヒトの薬物性肝障害予測システムの確立を目標としたコンソーシアム（MIP-DILI）プロジェクトの成果（Lhasa が収集）を基に構築した肝毒性エンドポイントに関連する MIE/KE のモデル化アプローチを統合した KNIME ワークフローの作成を行った。このプロファイラーでは、関連するエンドポイントに MIE/KE をマッピングすることにより Derek Nexus の既存知識を活用し、リードアクロスモデルには陽性化合物との構造類似性に基づくアプローチを用いている。また Random Forest 機械学習法により予測を行うと共に、反応性代謝物を生成する可能性および CYP450 活性部位への結合性も示すものとなっている。このワークフローに予測対象とする化合物を適用した結果は、別途 Instant JChem を用いて、対象化学物質に関して参照可能なすべての情報が一覧できる可視化するツールも開発した。今回、開発した KNIME ワークフローと可視化ツールを用いることで、肝毒性の誘発性に関する判別結果を単純に表示することはできないが、ヒト細胞株/タンパク質を用いた in vitro データが統合されており、これまで動物試験のデータを用いてきたものより広い化学空間をカバーできるものと考えられる他、専門家が当該化合物の毒性予測を行うのに有用なツールを開発することができた考えられる。

A. 研究目的

本研究に先立つ研究において、げっ歯類反復投与毒性データセットを用いて作成した

肝毒性のラピッドタイプアラートをフルアラートへと発展させる研究を行ってこられ、Derek Nexus における検出感度を、7%

から 40%に上昇させることに成功してきた。しかし、公表論文の調査結果を用いて一つ一つアラートを検討するために相当の時間がかかることが問題点として挙げられてきた。そこで、27年度より肝毒性エンドポイントに関連して予測した key event (KE) と、生体内経路の攪乱に關与する化合物群のプロファイルと組み合わせることにより、予測モデルの構築を試みてきた。H27年度ではこの肝毒性 Key Event プロファイラーは肝毒性反復投与データセットに対して比較的良好な感度を示したが(1000 mg/kg 超でのみ活性を示す化合物を除外すると 60%)、特異度は低かった(44%)。H28年度は、各 Key Event モデルの適用可能範囲の定義と WhichCyp の予測利用、化合物の求電子性炭素原子の部分電荷、KE のリードアクロスモデルの類似性カットオフの較正による改良を行い、27年度に比べて、特異度を上昇(52%)させることで偽陽性を減少させることができた。しかし、低い感度は KE 不足を反映していると考えられた。さらに KNIME ワークフローを使って予測結果を示す可視化ツールのプロトタイプを作成した。そこで、H29年度は、前年度までに改良してきたプロファイラーを基にした KNIME ワークフローの作成と、その結果を表示できる可視化ツールの構築を試みた。また、今年度は、ヒトの薬物性肝障害予測システムの確立を目標としたコンソーシアム(MIP-DILI)プロジェクトの成果を MIE および KE のプロファイラーに追加統合して、KNIME ワークフローの作成を行った。

B. 研究方法

KNIME ワークフローとして、H26 およ

び H27 年度の成果と MIP-DILI プロジェクトの結果を統合した、肝毒性のエンドポイントに関連する Molecular Initiating Event (MIE) および KE のプロファイラーを作成する。また可視化ツールとしては、プロファイラーで得られた結果をユーザーに分かりやすく示すビューを Instant JChem を用いて設計した。

肝臓データベースの構築：MIP-DILI プロジェクトと毒性試験データベースのデータを、1つの SQLite データベースに統合した。このデータベースには、肝毒性のエンドポイントに関連する MIE および KE に加えて、ヒトでの肝毒性 AOP に関する in vivo データが含まれている。各データセットには、標準化した化学構造式および毒性データ、すなわちエンドポイントの名称(BSEP 阻害など)、2値の注釈(活性または不活性)、標準化を行う前のデータおよび構造式を抽出したデータセットの参考文献、動物種(ヒトまたはラットなど)が含まれる。このデータベースには、標準化した 18053 個の化学構造式と 67 のデータセットが集積され、38 種類のエンドポイントをカバーしている。データ抽出元の一覧を表 1 に、データベース内容の要約を表 2 に示した。本データ群は、リードアクロスモデルおよび機械学習モデルのエントリーポイントとした。

C. 研究結果

KNIME ワークフローは 2 種類のアプローチを利用した。1つ目のアプローチでは、Derek Nexus のアラートを肝毒性に関連する MIE および KE でマッピングして用いている。これらのアラートにクエリー化合物がマッチすると、対応する MIE および /

または KE がユーザーに表示される。2 つ目のアプローチは、リードアクロスであり、肝毒性に関する MIE および KE の公開データを収集し、種々の MIE、KE に関して陽性データを有する各化合物と毒性試験データセットから得たクエリー化合物を、Lhasa 社で開発したフラグメントに基づく類似度 (Lhasa Limited 社の Sarah Nexus ソフトウェアにも採用) を計算するツールを用いて比較した。与えられた MIE/KE データセットの少なくとも 1 つの化合物に対して MIE/KE 固有の閾値を上回る類似性スコアを示した場合、そのクエリー化合物に当該 MIE/KE のフラグを付した。さらに、MIP-DILI プロジェクトの成果を統合した肝毒性のエンドポイントに関連する MIE および KE のプロファイラーは、Random Forest アルゴリズムと 2D RDKit 記述子を用いて、機械学習モデルを構築した。統計学習技術に依存し、Random Forest アルゴリズムを用い、各クエリー化合物の信頼性尺度を計算した。最終的に両方のアプローチを、1 つの KNIME ワークフローに統合した。

肝毒性プロファイラー：

肝毒性プロファイラーはモジュール方式で作成されており、次のモジュールを含む。

- 構造式標準化
- 適用可能範囲
- Derek Nexus のアラートにマッピングされた MIE および KE
- リードアクロスモデル
- 機械学習モデル (信頼性尺度を含む)
- 反応性代謝物に対するアラート (Kalgutkar アラート)

● WhichCyp 予測

これらのモジュールは、プロジェクトの一環として作成した肝毒性データベースより取得したデータを用いて、1 つの KNIME ワークフローに統合した。

● 構造式標準化

トレーニングセット内のすべての化合物およびクエリー化合物の構造表記については、各構造式のバリデーション、キュレーションおよび標準化を行った。図 1 に示す通り、構造式を確認、整理し、可能な場合には Lhasa Limited 社の一連の変換ルールを用いて正規化を行った。変換ルールには、芳香族性の認識、互変異性体および共鳴構造の変換、塩や金属の除去などを含む。

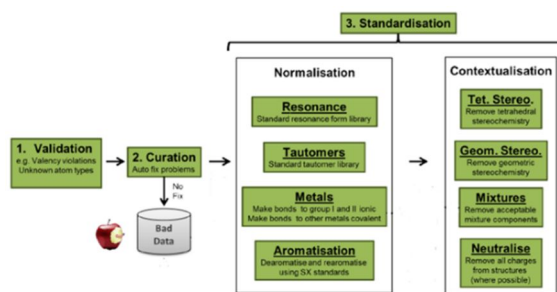


図 3. 化学構造式のバリデーション、キュレーションおよび標準化のプロセス

これらの段階を通過できなかった構造式については、肝毒性プロファイラーのモデル化方法にノイズを与えると考えられるため破棄した。

● 適用可能範囲

図 2 に示す通り、モデルに用いられたデータセットに対応する一連の RDKit 記述子の最大値と最小値を考慮し、各 MIE/KE モデルの適用可能範囲を定義した。KE データセットにおける記述子は、KNIME の RDKit Descriptor calculation ノードで化学構造から計算した。分子量、位相幾何学的

(トポロジカル)表面積、logP、原子数、水素結合供与体数、水素結合受容体数、環の数など 117 の記述子があり [RDKit] KE データセットの大きさを問わず、すべての記述子を使用した。

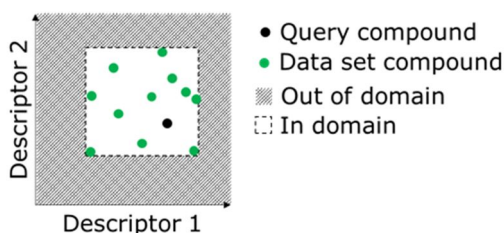


図 2. 適用可能範囲の定義

各 MIE/KE に関して、クエリー化合物の記述子の値が少なくとも 1 つの記述子変数の範囲を超えた場合、当該クエリー化合物はモデルの適用可能範囲外と考え、それ以外は適用可能範囲内とした。

- Derek Nexus のアラートにマッピングされた MIE および KE

肝毒性に関連すると考えられるエンドポイントに対するアラートに KE および / または MIE をマッピングするために、Derek Nexus の知識ベースに既に存在する知識を活用した。肝障害に至るメカニズムとしては、ミトコンドリアの機能障害、トランスポーターの抑制、直接反応性を有する生体異物および反応性代謝物などがある [Roth and Lee, Gijbels and Vinken]。In vivo での肝毒性を記述した肝毒性エンドポイントに関するアラートには、KE および MIE の説明など障害の潜在的メカニズムに関する情報が含まれている。これらの KE および MIE の特定、ならびにアラートへのマッピングは、専門家による肝毒性アラートの分析に基づいて行った。同じ毒性を複数のメ

カニズムで発現しうる化合物クラスに対するアラートは、アラート説明において該当する key event とリンクさせた。また、肝毒性に関連する KE (ミトコンドリア機能障害など) を予測する Derek Nexus のエンドポイントを含めた。このエンドポイントは、分子起始反応 (MIE) ではなく、AOP のより下流の KE であるミトコンドリアへの影響 (ミトコンドリア機能障害) を評価するアッセイに基づいている。

一方、Derek Nexus には、化学物質の固有反応性および反応種への代謝活性化に関する広範な知識も収録されており、これらの知識は変異原性 (化学物質により引き起こされた核 DNA の損傷が原因となる) や皮膚感作 (皮膚タンパク質のハプテン化が特徴である) のエンドポイントに関するアラートで広くカバーされている。肝酵素による化学物質の代謝活性化に起因するハプテンの形成やミトコンドリア DNA の損傷は、肝毒性のメカニズムとしてよく知られている [Yuan and Kaplowitz]。したがって、変異原性および皮膚感作のエンドポイントに関するアラートを分析し、専門家が特定したメカニズムを踏まえて、関連する反応性に基づいた MIE エンドポイント (それぞれ、DNA 反応性求電子剤、タンパク質反応性求電子剤) にリンクさせた。代謝活性化の必要条件も考慮し、可能な場合には、カスタム知識ベースの 2 種のアラート、すなわち直接反応性を有する化合物に対するカスタマイズアラートと生体内活性化を要する化合物に対するカスタマイズアラートにアラートのパターンを分別した。

他の関連する key event に関する変異原性・皮膚感作アラートについても検討し、該

当する key event エンドポイント(酸化剤、核酸塩基模倣物など)を割り当てた。MIE/KE と各エンドポイントに対応するアラートの数の詳細な一覧を表 3 に示す。肝毒性プロファイラーにおいてクエリー化合物がこのようなアラートにマッチすると、肝毒性に至る潜在的メカニズムとして、対応する MIE および/または KE がユーザーに示される。

- リードアクロスモデル

肝毒性データベースから key event データを抽出し、複数のデータセットに存在する化合物に関しては保守的な判定を適用した。これにより key event ごとに 1 つのデータセットが得られ、次に不活性と標識された化合物を key event データセットから除外した。その後、構造フラグメントベースの fingerprint を利用して、各クエリー化合物を各 key event データセット中の化合物と比較した。これらの fingerprint を生成するアルゴリズムは Lhasa Limited 社で開発したもので、Sarah Nexus ソフトウェアで採用されている。2 つのフィンガープリントの類似性は Tanimoto 係数で評価した。Key event データセットの少なくとも 1 つの化合物に対して Tanimoto 係数が 0.7 以上である場合、そのクエリー化合物に当該 key event のフラグを付した。最も類似性の高い化合物(最大 3 個)を、それぞれ該当する Tanimoto 係数とともにアウトプットした。各 MIE/KE に関するリードアクロスモデルの予測に対し、3.2 項「適用可能範囲」に定義した適用可能範囲に関する情報を追加した。

- 機械学習モデル

肝毒性データベースの個々のデータセッ

トに基づき、Random Forest アルゴリズム [Breiman] と 2D RDKit 記述子を用いて、機械学習モデルを構築した。分散が 0 の記述子、およびピアソンの相関係数 $|R|$ が 0.95 を超える記述子は除外した。Random Forest アルゴリズムに用いた記述子の数はデータセットに依存したが、モデルの過学習を防ぐには記述子当たり 10 個以上の化合物が必要である。記述子の数がこの要件(トレーニング化合物 10 個に対して 1 個)を超過した場合には、活性に関する注釈(活性または不活性)との相関性が最も高い記述子を選択する方法をとった。Random Forest モデルのパラメータは次の通り: 1000 trees、equal size with replacement、split criterion - Gini index。

本モデルでクエリー化合物を処理すると、予測とその信頼性(注釈を予測した決定木の数をすべての決定木の数で除して求める)が得られる。各モデルから得られた信頼性は、個々の MIE/KE データセットに関して 5 層交差バリデーション法で求めた balanced accuracy (BA) 平均値により補正した。

1 つの MIE や KE に対して複数のデータセットが存在する場合、個々のモデルの信頼性の加重平均を計算することにより総合的な信頼性を求めた。総合的な信頼性の算出に用いたアルゴリズムと式は図 3 に示す。個々の信頼性は、不活性予測の場合にはマイナス、活性予測の場合にはプラスの値となる。

総合的な信頼性の符号によって、対応する予測(マイナスは不活性、プラスは活性)の注釈をつけた。最後に、総合的な信頼性の絶対値を計算する。総合的な信頼性の範囲は 0(信

信頼性が低い) ~ 1 (信頼性が高い) であり、た。
可視化ツールではパーセンテージに変換し

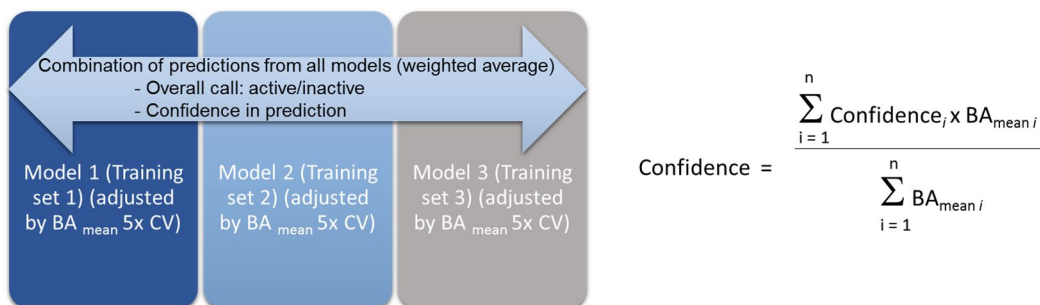


図 3. 交差バリデーション法

元のデータセットを 5 分割し、4 つをトレーニングセットとして用いて Random Forest モデルを作成、残りの 1 つをテストセットとして用いて balanced accuracy (BA) を導出し、5 つの BA の平均値を算出する。

適用可能範囲外 (「●適用可能範囲」の定義に従う) のクエリー化合物に関する予測は信頼性がないとみなし、信頼性スコアの計算対象から除外した。特定のエンドポイントに関するすべてのデータセットの適用可能範囲外にある化合物については、MIE に関する予測は得られなかった。

● 反応性代謝物に対するアラート (Kalgutkar アラート)

反応性代謝物に対するモデルは、文献で毒性作用への関与が報告されている既知の toxiphore に基づき、KNIME で作成した [Stepan et al.] 計 59 種類のアラートを作成し、反応種を生成する可能性があると思われる化学的特徴を SMARTS パターンとして収録した。これらのアラートは Kalgutkar アラートとし、SMARTS パターンとその名称の一覧を表 4 示した。各化合物に関して表示される結果には、反応性代謝物アラートの有無 (yes/no) およびヒットしたアラートの一覧が含まれる。

● WhichCyp 予測

反応性代謝物を生成しうる化学的特徴が存在したとしても、その化合物が必ずしも生体内活性化を受けるわけではないことは既知の通りである。したがって、主な CYP アイソフォーム 5 種 (1A2、2C9、2C19、2D6、3A4) に対する結合を予測する KNIME の WhichCYP ノード [Rostkowski et al.] を用い、化合物の CYP 結合能に関する追加要件を導入した。

反応性代謝物を生成する可能性のある化学物質を確認するためのワークフローを、図 4 に要約する。

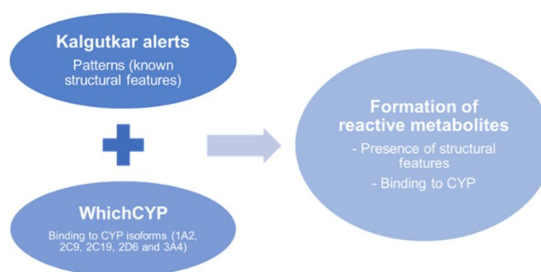


図 4. 反応性代謝物を生成する可能性の

ある化学物質を確認するワークフロー図

各クエリー化合物に関して WhichCYP ノードから得られる結果には、化学物質の CYP 活性部位への結合性の有無(yes/no)および結合する CYP アイソフォームの一覧が含まれる。

D. 考察

今回 MIP-DILI の結果を統合した肝毒性プロファイラーはヒト細胞株/タンパク質で得た *in vitro* データが用いられている。これは製薬会社 MIP-DILI のプロジェクトの目的としてのヒトでの医薬品による臨床毒性を予測することにあるため、ヒト細胞を用いた *in vitro* 試験系/モデルに重点が置かれているという現状がある。ヒト細胞株/タンパク質に関しては他の動物種よりも多くのデータが得られており、結果として、このようなデータを用いて作成したモデルでは、これまで動物試験のデータを用いてきたものより広い化学空間をカバーできるものと考えられる。

一方、昨年度まで予測検証用に使用してきた毒性試験データセットはげっ歯類反復投与毒性試験の結果を収集したもので、その肝毒性に関する注釈は同毒性について報告された最小作用量に基づいている。今回のモデルに使用された薬剤性肝障害のデータと齧歯類の肝毒性物質のデータの間に関係を示さないことが知られており、昨年度まで検証用に使用してきた毒性試験データセットに関して Cooper statistics の観点からプロファイラーの予測精度を分析することは適切ではないと考えられた。

その代わりに、各化合物に関してモデル化法で得られるすべての情報をユーザーに分

かりやすく表示するための可視化ツールをデザインした。例として、毒性試験データセットに含まれる化合物の1つを肝毒性プロファイラーで処理した際のアウトプットを図5に示す。各化合物について、固有のデータセット ID、構造情報、分子量、毒性試験データセットの反復投与試験に基づく注釈、そして Derek Nexus、リードアクロスモデル、機械学習モデルを含むすべてのアプローチをカバーした MIE/KE モデルから得られた結果が1つのビューに表示され、予測対象とする化学物質の様々なプロファイルに関する情報を一覧することができ、専門家による毒性予測を手助けすることが可能なツールを作成することができた。

参考文献

- Breiman, L. Random Forests. *Machine Learning*, 2001, 45, 5-32. doi: 10.1023/A:1010933404324
- Gijbels E, and Vinken M. An Update on adverse outcome pathways leading to liver injury. *Applied In Vitro Toxicology*, 2017, 3, 283-285. doi: 10.1089/aivt.2017.0027.
- Kalgutkar AS, Didiuk MT. Structural alerts, reactive metabolites, and protein covalent binding: how reliable are these attributes as predictors of drug toxicity? *Chem Biodivers*. 2009, 6, 2115-2137. doi: 10.1002/cbdv.200900055.
- Rostkowski, M., Spjuth, O. & Rydberg, P. WhichCyp: Prediction of cytochromes P450 inhibition.

Bioinformatics 29, 2051–2052 (2013).
doi: 10.1093/bioinformatics/btt325

RDKit: Open-Source Cheminformatics Software. RDKit Nodes for KNIME (trusted extension). Available at: <https://tech.knime.org/community/rdkit>. (Accessed: 25th January 2017)

Roth AD, and Lee M-Y. Idiosyncratic drug-induced liver injury (IDILI): potential mechanisms and predictive assays. *BioMed Research International*, 2017, Article ID 9176937. doi: 10.1155/2017/9176937.

Stepan AF, Walker DP, Bauman J, Price DA, Baillie TA, Kalgutkar AS, and Aleo MD. Structural alert/reactive metabolite concept as applied in medicinal chemistry to mitigate the risk of idiosyncratic drug toxicity: a perspective based on the critical examination of trends in the top 200 drugs marketed in the United States. *Chemical Research in Toxicology*, 2011, 24, 1345-1410. doi: 10.1021/tx200168d.

Yuan L, Kaplowitz N. Mechanisms of Drug Induced Liver Injury. *Clinics in Liver Disease*, 2013, 17, 507–518. doi: 10.1016/j.cld.2013.07.002.

E. 結論

H29年度は、前年度までに改良してきた肝毒性に関連するプロファイラーを基にして作成した KNIME ワークフローに、ヒトの薬物性肝障害予測システムの確立を目標としたコンソーシアム (MIP-DILI) プロジ

エクトの成果 (Lhasa が収集) を基に構築した肝毒性エンドポイントに関連する MIE/KE のモデル化アプローチを統合した KNIME ワークフローの作成を行った。このプロファイラーでは、関連するエンドポイントに MIE/KE をマッピングすることにより Derek Nexus の既存知識を活用し、リードアクロモデルには陽性化合物との構造類似性に基づくアプローチを用いている。また Random Forest 機械学習法により予測を行うと共に、反応性代謝物を生成する可能性および CYP450 活性部位への結合性も示すものとなっている。このワークフローに予測対象とする化合物を適用した結果は、別途 Instant JChem を用いて、対象化学物質に関して参照可能なすべての情報が一覧できる可視化するツールも開発した。今回、開発した KNIME ワークフローと可視化ツールを用いることで、肝毒性の誘発性に関する判別結果を単純に表示することはできないが、専門家が当該化合物の毒性予測を行うのに有用なツールを開発することができた考えられる。

F. 研究発表

1. 論文発表

Tanabe, S., Kobayashi, K., Matsumoto, M., Serizawa, H., Igarashi, T., Yamada, T., Hirose, A. Toxicity of repeated 28-day oral administration of acenaphthylene in rats, (2017) *Fundam. Toxicol. Sci.*, 4, 247-259.

Matsumoto, M., Iso, T., Yamaguchi, H., Igarashi, T., Yamada, T., Hirose, A. Summary information of human health hazard assessment of existing

chemical substances (III), *Bull. Natl. Inst. Health Sci.*, 135, 39-44, 2017.

Matsumoto, M., Furukawa, M., Kobayashi, K., Iso, T., Igarashi, T., Yamada, T., Hirose, A., A 28-day repeated oral-dose toxicity study of insecticide synergist N-(2-ethylhexyl)-1-isopropyl-4-methylbicyclo [2.2.2] oct-5-ene-2,3-dicarboximide in rats, (2018) *Fundam. Toxicol. Sci.*, 5, 1-11.

Yamada, T., Tanaka, Y., Hasegawa, R., Igarashi, T., Hirose, A. Male-specific prolongation of prothombin time by industrial chemicals, (2018) *Fundam. Toxicol. Sci.*, 5, 75-82.

2. 学会発表

Yamada, T., Hirose, A., Honma, M. Our recent experiences for development of read-across approach for chemical safety assessment, 10th World Congress of Alternatives and Animal Use in the Life Sciences (August 2017, Seattle, Washington, USA).

Yamada, T., Hirose, A., Yang, C., Rathman, J. Development of a new database of non-cancer toxicity endpoints of industrial chemicals for improving TTC approach. 53rd European Congress of the European Societies of Toxicology, (September 2017, Bratislava, Slovak).

Matsumoto, M., Miura, M., Kawamura, T., Yamada, T., Kobayashi, N., Suzuki, T., Nishimura, T., Hirose, A. Derivation

of health advisory values for sub-acute exposure of contaminants in drinking water. 53rd European Congress of the European Societies of Toxicology, (September 2017, Bratislava, Slovak).

Yamada, T., Hirose, A., Yang, C., Rathman, J. Evaluation of Genotoxicity and Carcinogenicity Profile for the TTC Database Used to Develop Safety Thresholds of Industrial Chemicals. Society of Toxicology Society, 57th Annual Meeting and ToxExpo (March 2018, San Antonio, USA).

Hirose, A., Kawamura, T., Igarashi, T., Yamada, T., Inoue, K., Nishimura, T., Matsumoto, M. Expansion of the sub-acute reference doses for non-quality-standard chemicals in drinking water. Society of Toxicology Society, 57th Annual Meeting and ToxExpo (March 2018, San Antonio, USA).

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. 実用新案特許
なし
3. その他
なし

図 5 可視化ツールのイメージ

The screenshot shows a web-based interface for chemical analysis. It includes several panels:

- STRUCTURE INFORMATION:** Displays a chemical structure and basic data like Name (9506001A), Formula (C19H18ClN3O), and Mol Weight (291.78).
- READ-ACROSS MODELS:** A table listing key event names and their applicability domains.

Key Event Name	In Applicability Domain
1 MitochondrialDysfunction_MMP_Disruption	yes
2 PXR_Agonist	yes
- Active examples:** A table showing similar structures with their similarity and distance scores.

Structure	Similarity Score (Tanimoto similarity based on Fingerprint)	Distance Score (Euclidian distance based on Lipinski's descriptors)
1	0.72	0.02
2	0.71	0.01
- MACHINE LEARNING MODELS: Random Forest:** A table showing predictions for key events.

Key Event Name	Prediction	Confidence (%)
1 BSEP_Inhibition	Inactive	57.3
2 PXR_Agonist	Inactive	51.7
3 NRP4_Inhibition	Inactive	95.0
- STRUCTURAL ALERT MODELS: Derek Nexus Alert Mapped to Key Event:** A table mapping alerts to key events.

Key Event Name	Bioactivation Required	Alert Number	Alert Name
1 Constitutive androstane receptor (CAR) binding	No	547	Phenylalkyltriazole or analog
2 CYP inhibition	No	547	Phenylalkyltriazole or analog
- REACTIVE METABOLITE:** A table showing potential reactive metabolites and their names.

Potential Reactive Metabolite	Alert Names
1 yes	Aryl-OH and Aryl-CX
2 yes	Chlorobenzene
- WhichCyp Prediction:** A table showing binding to CYP active sites and CYP450 isoforms.

Binding	CYP450 Isoform
1 yes	1A2
2 yes	2C19

Callouts provide additional context for these sections:

- List of the Read-across KE models:** Name of KE endpoints, Applicability domain (yes/no).
- Machine Learning Models:** Key Event name, Prediction, Confidence in prediction (%).
- Prediction for Reactive Metabolites:** Potential for formation of Reactive Metabolites, Name of (Kalgutkar) alert(s).
- WhichCyp Prediction:** Binding to CYP active site, CYP450 isoforms.
- List of KE names (mapped to Derek Nexus alerts):** Name of KE endpoints, Requirement for bioactivation (yes/no), Alert number (DX knowledgebase), Name of DX alert.
- Supporting Active examples for Read-across models:** Structure of example, Similarity Score (Tanimoto similarity on fingerprints), Distance score (Euclidian distance on Lipinski's descriptors).

表 1. 肝毒性データベースに含まれるデータソース

Dataset ID	Dataset description	Link
Aleo	Hepatology_2014_60_1015-1022	https://doi.org/10.1002/hep.27206
Aleo_2017	Chem Res Toxicol_2017_30_1219-1229	https://doi.org/10.1021/acs.chemrestox.7b00048
Broccatelli	J Med Chem_2011_54_1740-1751	https://doi.org/10.1021/jm101421d
ChEMBL_TargetId=CHEM BL1870	ChEMBL data on agonists of Retinoid X receptor beta	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL1870
ChEMBL_TargetId=CHEM BL2003	ChEMBL data on agonists of Retinoic acid receptor gamma	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL2003
ChEMBL_TargetId=CHEM BL2004	ChEMBL data on agonists of Retinoid X receptor gamma	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL2004
ChEMBL_TargetId=CHEM BL2008	ChEMBL data on agonists of Retinoic acid receptor beta	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL2008
ChEMBL_TargetId=CHEM BL2034	ChEMBL data on agonists of Glucocorticoid receptor	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL2034
ChEMBL_TargetId=CHEM BL2047	ChEMBL data on agonists of Bile acid receptor FXR	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL2047
ChEMBL_TargetId=CHEM BL2055	ChEMBL data on agonists of Retinoic acid receptor alpha	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL2055

Dataset ID	Dataset description	Link
ChEMBL_TargetId=CHEM BL206	ChEMBL data on agonists of Estrogen receptor alpha	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL206
ChEMBL_TargetId=CHEM BL2061	ChEMBL data on agonists of Retinoid X receptor alpha	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL2061
ChEMBL_TargetId=CHEM BL235	ChEMBL data on agonists of Peroxisome proliferator-activated receptor gamma	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL235
ChEMBL_TargetId=CHEM BL239	ChEMBL data on agonists of Peroxisome proliferator-activated receptor alpha	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL239
ChEMBL_TargetId=CHEM BL242	ChEMBL data on agonists of Estrogen receptor beta	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL242
ChEMBL_TargetId=CHEM BL2808	ChEMBL data on agonists of LXR-alpha	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL2808
ChEMBL_TargetId=CHEM BL3201	ChEMBL data on agonists of Aryl hydrocarbon receptor	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL3201
ChEMBL_TargetId=CHEM BL3401	ChEMBL data on agonists of Pregnane X receptor	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL3401
ChEMBL_TargetId=CHEM BL3979	ChEMBL data on agonists of Peroxisome proliferator-activated receptor delta	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL3979
ChEMBL_TargetId=CHEM BL4093	ChEMBL data on agonists of LXR-beta	https://www.ebi.ac.uk/chembl/target/inspect/CHEM BL4093
Chen	Drug Discovery Today_2016_21_648-653	https://doi.org/10.1016/j.drudis.2016.02.015
De Bruyn	Mol Pharmacol_2013_83_1257-1267	https://doi.org/10.1124/mol.112.084152
Greene	Chem Res Toxicol_2010_23_1215-1222	https://doi.org/10.1021/tx1000865
Karlgren	J Med Chem_2012_55_4740-4763	https://doi.org/10.1021/jm300212s
Kock	Drug Metab Dispos_2014_42_665-674	https://doi.org/10.1124/dmd.113.054304
Kruhlak	Toxicology_2014_321_62-72	https://doi.org/10.1016/j.tox.2014.03.009
Matsson	J Pharmacol Exp Ther_2007_323_19-30	https://doi.org/10.1124/jpet.107.124768
Morgan	Toxicol Sci_2013_136_216-241	https://doi.org/10.1093/toxici/kft176
Mulliner	Chem Res Toxicol_2016_29_757-767	https://doi.org/10.1021/acs.chemrestox.5b00465
Pedersen	Toxicol Sci_2013_136_328-343	https://doi.org/10.1093/toxici/kft197
PubChem_AID=1159531	qHTS assay to identify small molecule agonists of the RXR signaling	https://pubchem.ncbi.nlm.nih.gov/bioassay/1159531

Dataset ID	Dataset description	Link
	pathway: Summary	
PubChem_AID=1159553	qHTS assay to identify small molecule agonists of the retinoic acid receptor (RAR) signaling pathway	https://pubchem.ncbi.nlm.nih.gov/bioassay/1159553
PubChem_AID=1451	Single point confirmation for ABC transporter inhibitors: specifically ABCB1 screen, ABCG2 counter-screen	https://pubchem.ncbi.nlm.nih.gov/bioassay/1451
PubChem_AID=1453	Single point confirmation screening for ABC transporter inhibitors: specifically ABCG2 screen, ABCB1 counter-screen	https://pubchem.ncbi.nlm.nih.gov/bioassay/1453
PubChem_AID=1689	Dose Response for ABC transporter inhibitors: specifically ABCB1 screen, ABCG2 counter-screen	https://pubchem.ncbi.nlm.nih.gov/bioassay/1689
PubChem_AID=1690	Dose Response for ABC transporter inhibitors: specifically ABCG2 screen, ABCB1 counter-screen	https://pubchem.ncbi.nlm.nih.gov/bioassay/1690
PubChem_AID=651754	qHTS assay for small molecule disruptors of the mitochondrial membrane potential - 1 hr assay	https://pubchem.ncbi.nlm.nih.gov/bioassay/651754
PubChem_AID=651755	qHTS assay for small molecule disruptors of the mitochondrial membrane potential - 5 hr assay	https://pubchem.ncbi.nlm.nih.gov/bioassay/651755
PubChem_AID=720637	qHTS assay for small molecule disruptors of the mitochondrial membrane potential: Summary	https://pubchem.ncbi.nlm.nih.gov/bioassay/720637
PubChem_AID=720659	qHTS assay for small molecule activators of the human pregnane X receptor (PXR) signaling pathway	https://pubchem.ncbi.nlm.nih.gov/bioassay/720659
PubChem_AID=720719	qHTS assay to identify small molecule agonists of the glucocorticoid receptor (GR) signaling pathway: Summary	https://pubchem.ncbi.nlm.nih.gov/bioassay/720719
PubChem_AID=743077	qHTS assay to identify small molecule agonists of the estrogen receptor alpha (ER-alpha) signaling pathway: Summary	https://pubchem.ncbi.nlm.nih.gov/bioassay/743077
PubChem_AID=743122	qHTS assay to identify small molecule that activate the aryl hydrocarbon receptor (AhR) signaling pathway: Summary	https://pubchem.ncbi.nlm.nih.gov/bioassay/743122

Dataset ID	Dataset description	Link
PubChem_AID=743140	qHTS assay to identify small molecule agonists of the peroxisome proliferator-activated receptor gamma (PPAR γ) signaling pathway: Summary	https://pubchem.ncbi.nlm.nih.gov/bioassay/743140
PubChem_AID=743227	qHTS assay to identify small molecule agonists of the peroxisome proliferator-activated receptor delta (PPAR δ) signaling pathway: Summary	https://pubchem.ncbi.nlm.nih.gov/bioassay/743227
PubChem_AID=743239	qHTS assay to identify small molecule agonists of the farnesoid-X-receptor (FXR) signaling pathway: Summary	https://pubchem.ncbi.nlm.nih.gov/bioassay/743239
Sakatis	Chem Res Toxicol_2012_25_2067-2082	https://doi.org/10.1021/tx300075j
Schadt	Toxicology In Vitro_2015_30_429-437	https://doi.org/10.1016/j.tiv.2015.09.019
Warner	Drug Metab Dispos_2012_40_2332-2341	https://doi.org/10.1124/dmd.112.047068
Xu	Toxicol Sci_2008_105_97-105	https://doi.org/10.1093/toxicology/kfn109

表 2 肝毒性データベースの項目の概要

Endpoint Name	DatasetId	Species	Number of Active Compounds	Number of Inactive Compounds
agonist	ChEMBL_TargetId=ChEMBL3201	Human	134	18
agonist	PubChem_AID=743122	Human	755	5709
Inhibition	Matsson	Human	42	77
Inhibition	PubChem_AID=1453	Human	16	255
Inhibition	PubChem_AID=1690	Human	16	24
Inhibition	Aleo_2017	Human	33	87
Inhibition	Morgan	Human	135	430
Inhibition	Pedersen	Human	81	150
Inhibition	Warner	Human	321	289
2_MechanismDependentInhibitor	Sakatis	Human	15	157
19_MechanismDependentInhibitor	Sakatis	Human	34	138
9_MechanismDependentInhibitor	Sakatis	Human	14	158
6_MechanismDependentInhibitor	Sakatis	Human	15	156
4-7BQ_MechanismDependentInhibitor	Sakatis	Human	21	151











Endpoint Name	DatasetId	Species	Number of Active Compounds	Number of Inactive Compounds
4-DEF_MechanismDependentInhibitor	Sakatis	Human	36	134
4_TimeDependentInhibitor	Schadt	Human	14	102
alpha_Agonist	ChEMBL_TargetId=ChEMBL206	Human	395	506
alpha_Agonist	PubChem_AID=743077	Human	322	6570
alpha_Agonist	ChEMBL_TargetId=ChEMBL242	Human	317	601
alpha_Agonist	ChEMBL_TargetId=ChEMBL2047	Human	436	206
alpha_Agonist	PubChem_AID=743239	Human	88	6327
alpha_Agonist	ChEMBL_TargetId=ChEMBL2034	Human	422	181
alpha_Agonist	PubChem_AID=720719	Human	147	6549
alpha_Agonist	Sakatis	Human	82	102
alpha_Agonist	Schadt	Human	35	84
alpha_Agonist	Chen	Human	420	228
alpha_Agonist	Greene	Human	303	132
alpha_Agonist	Kruhlak	Human	176	102
alpha_Agonist	Mulliner	Human	1140	452
alpha_Agonist	Sakatis	Human	109	103
alpha_Agonist	Schadt	Human	54	65
alpha_Agonist	Xu	Human	147	135
alpha_Agonist	ChEMBL_TargetId=ChEMBL2808	Human	313	174
alpha_Agonist	ChEMBL_TargetId=ChEMBL4093	Human	457	811
alpha_Agonist	Aleo_2017	Human	59	61
alpha_Agonist	Morgan	Human	23	502
alpha_Agonist	Kock	Human	46	36
alpha_Agonist	Morgan	Human	39	519
alpha_Agonist	Kock	Human	48	35
alpha_Agonist	Morgan	Human	93	463
alpha_Agonist	Aleo	Rat	26	42
alpha_Agonist	PubChem_AID=651754	Human	123	1020
alpha_Agonist	PubChem_AID=651755	Human	110	1082
alpha_Agonist	PubChem_AID=720637	Human	871	4916
alpha_Agonist	De Bruyn	Human	186	1625
alpha_Agonist	Karlgren	Human	76	145
alpha_Agonist	De Bruyn	Human	120	1690
alpha_Agonist	Karlgren	Human	45	177
alpha_Agonist	Karlgren	Human	45	177
alpha_Agonist	Broccatelli	Not specified	621	587
alpha_Agonist	PubChem_AID=1451	Human	18	253
alpha_Agonist	PubChem_AID=1689	Human	9	31
alpha_Agonist	ChEMBL_TargetId=ChEMBL239	Human	1426	536











Endpoint Name	DatasetId	Species	Number of Active Compounds	Number of Inactive Compounds
delta_Agonist	ChEMBL_TargetId=ChEMBL3979	Human	899	499
delta_Agonist	PubChem_AID=743227	Human	67	6077
gamma_Agonist	ChEMBL_TargetId=ChEMBL235	Human	2037	584
gamma_Agonist	PubChem_AID=743140	Human	178	6256
gonist	ChEMBL_TargetId=ChEMBL3401	Human	137	271
gonist	PubChem_AID=720659	Human	207	1664
rho_Agonist	ChEMBL_TargetId=ChEMBL2055	Human	47	91
rho_Agonist	PubChem_AID=1159553	Human	325	5471
tau_Agonist	ChEMBL_TargetId=ChEMBL2008	Human	85	46
mu_Agonist	ChEMBL_TargetId=ChEMBL2003	Human	57	42
rho_Agonist	ChEMBL_TargetId=ChEMBL2061	Human	161	115
rho_Agonist	PubChem_AID=1159531	Human	150	5384
tau_Agonist	ChEMBL_TargetId=ChEMBL1870	Human	40	62
mu_Agonist	ChEMBL_TargetId=ChEMBL2004	Human	40	59











表3 DEREK から活用したアラート一覧










Endpoint	Number of alerts
3-Hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase inhibition	1
Acetyl CoA depletion	3
Androgen receptor binder	3
Angiotensin converting enzyme (ACE) inhibitor	1
Aryl hydrocarbon receptor binder	2
beta-Oxidation competitive substrate	2
beta-Oxidation inhibition	2
Bile transport competitive substrate (M)	1
Binding to tubulin	1
Carnitine depletion	1
Carnitine palmitoyltransferase I inhibition	1
Chelation of metal ions	1
Constitutive androstane receptor (CAR) binding	1
CYP inhibition (M)	2
CYP inhibition	1
DNA reactive electrophile	47
DNA reactive electrophile (M)	64
Electron Transport Chain Inhibition	2
Electron Transport Chain Inhibition - succinate-coenzyme Q reductase (SQR)	1
Estrogen receptor binder	17
Glutathione (GSH) depletion	1
Glutathione (GSH) depletion (M)	14
Inhibition of RNA polymerase (M)	1
Kreb cycle inhibition - alpha-ketoglutarate dehydrogenase complex inhibition	1
Low density lipoprotein (Lys-adduct) reactive electrophile (M)	1
Microsomal Triglyceride Transfer Protein (MTTP) antagonist	2
Mitochondrial dysfunction	24
Nucleobase Mimic	4
Nucleophilic thiol exchange compound	1
Nucleophilic thiol exchange compound (M)	2
Oxidising agent	19
Oxidising agent (M)	20
PPAR binder	7
Protein reactive electrophile	56
Protein reactive electrophile (M)	63
Protein reactive nucleophile	1
Redox Cycler (Metal)	1
RNA reactive electrophile (M)	1
Uncoupler of oxidative phosphorylation	4

表 4 . Kalgutkar アラートの SMARTS パターン

SMARTS pattern	Pattern Name
	N-Substituted Piperidines
	Aryl-CH and Aryl-CX
	Mono- and di-alkyl sulfate or alkyl sulfonate
	Chlorobenzene
	1,3-Benzodioxole
	2-Alkylimidazole
	2-Alkylpyrrole
	2-Aminothiazole
	Alkylaldehyde
	Alkylhydrazine

SMARTS pattern	Pattern Name
	Alkylimidazole
	Alkylpyrrole
	Alkyne
	Anilide
	Aniline
	Arylacetic acid
	Arylhydrazine
	Benzodiazepine
	Benzothiophene
	Bromobenzene

SMARTS pattern	Pattern Name
	Catechol
	Cyclopropyl amine
	Dialkoxyaromatic
	Dialkyltriazine
	Dihaloalkane
	Electrophilic ester
	Epoxide
	Furan
	Heteroarylacetic acid
	Hydrazide

SMARTS pattern	Pattern Name
	o-Alkylphenol
	o-Hydroxyaniline
	p-Alkoxyaniline
	p-Alkylanilide
	p-Alkylaniline
	p-Alkylaromatic ether
	p-Alkylphenol
	p-Hydroxyanilide
	p-Hydroxyaniline

	SMARTS pattern	Pattern Name
41		Thiazolidinedione
42		Thioamide
43		Thioester
44		Thiophene
45		Thiourea
46		beta-Lactam
47		o-Alkoxyaniline
48		o-Alkylanilide
49		o-Alkylaniline
50		o-Alkylaromatic ether

	SMARTS pattern	Pattern Name
51		o-Alkylphenol
52		o-Hydroxyaniline
53		p-Alkoxyaniline
54		p-Alkylanilide
55		p-Alkylaniline
56		p-Alkylaromatic ether
57		p-Alkylphenol
58		p-Hydroxyanilide
59		p-Hydroxyaniline