

化学構造式から受容体親和性を予測する機械学習
モデルの構築とヒト有害事象の予測への応用

2021

酒井 幸

目次

緒言	2
略語	3
第1章 グラフ畳み込みニューラルネットワークを用いた化学構造式からの親和性予測	5
1.1 序論	5
1.2 方法	7
1.3 結果と考察	14
1.4 まとめと展望	32
第2章 受容体親和性予測値を用いた有害事象および原因標的タンパク質の予測	33
2.1 序論	33
2.2 方法	35
2.3 結果と考察	42
2.4 まとめと展望	62
総括	65
謝辞	66
発表論文目録	67
参考文献	68

緒言

有害事象 (Adverse Events; AE) とは、医薬品が投与された患者に生じたあらゆる好ましくない医療上の出来事とされている。AE が原因で入院あるいは入院期間の延長が引き起こされ、ときに障害や死亡にもつながることがあるなど、その社会的および経済的損失は大きい。さらに AE は、治療効果不足、商業的理由に続く、非臨床試験や臨床試験の撤退の主原因でもある。そのため AE を予防する方法や、創薬の早い段階で効率的に AE を発見する手法に対する必要性は高い。

医薬品はその治療標的となる生体分子に強く作用するが、程度の差こそあれ他の生体分子にも作用する。ゆえに医薬品の開発段階において治療標的 (オンターゲット) とそれ以外 (オフターゲット) に対する結合選択性を高めるよう最適化される。さらに非臨床試験・臨床試験を通じて治療効果および安全性について精査される。それにも関わらず、市販後に想定外の生体分子などへの作用によるヒトでの有害事象がしばしば発生する。

多くの医薬品は簡単な化学構造式で表現可能な化合物であり、その薬理活性の発現の第一段階は標的となる生体分子への結合である。生体分子への親和性と薬理作用は密接に関係する現象であるが、どんなに経験を積んだ研究者であっても化学構造式をみただけで両者を事前に予測することは通常難しい。化合物の遺伝子発現に対する影響は比較的容易に、かつ網羅的にデータを取得できるのに対し、生体分子に対する結合親和性データを網羅的に取得することは時間的にも費用的にも現実的ではない。しかし、薬理作用の程度や様式を決定するのはその化学構造であるから、機械学習などで必要な情報を適切に抽出できれば、原理的には化学構造から薬理作用が予測可能と考えられる。

近年、化合物の生体分子に対する親和性実測値や市販後の医薬品による有害事象自発報告などが集積されたデータが公開され、創薬探索研究や医薬品の安全性監視活動に利用されている。かかる背景のもと、本研究ではこれらのビッグデータを活用することにより、化学構造式のみからヒトでの薬効や有害事象を予測する手法の確立を試み、次の新知見を得たので 2 章にわけて報告する。

第 1 章では、化学構造式を入力として、グラフ畳み込みニューラルネットワークを用いた生体分子への親和性予測モデルを構築し、受容体などの多様な標的タンパク質に対する親和性を高精度かつ定量的に予測できることを示した。これは、化学構造式そのものが標的タンパク質に対する親和性を決定づける要因として十分な情報を含んでいるとの見方に立つものである。

第 2 章では、第 1 章で構築した親和性予測モデルの出力する種々の標的タンパク質に対する親和性プロファイルを用いることで、ヒトでの有害事象を予測できることを示した。また、有害事象予測モデルに寄与する因子の解析を通じて、有害事象への関与が疑われる標的タンパク質を導出しえることも示した。

これらの研究成果について、以下に論述する。

略語

ADHD	attention-deficit hyperactivity disorder	注意欠陥・多動性障害
AE	adverse events	有害事象
ANOVA	analysis of variance	分散分析
AUC	area under the ROC curve	受信者操作特性曲線下面積
CNN	convolutional neural network	畳み込みニューラルネットワーク
COX	cyclooxygenase	シクロオキシゲナーゼ
DAT	dopamine transporter	ドパミントランスポータ
ECFP	extended-connectivity circular fingerprint	化学構造の数値表現のひとつ
eNOS	endothelial NOS	内皮型一酸化窒素合成酵素
FAERS	FDA adverse event reporting system	FDA 有害事象自発報告システム
FNN	fully-connected neural network	全結合型ディープニューラルネット ネットワーク (計算手法)
FDA	food and drug administration	米国食品医薬品局
FDR	false discovery rate	偽発見率
NOS	nitric oxide synthase	一酸化窒素合成酵素
GABA	gamma-aminobutyric acid	γ -アミノ酪酸
GCN	graph convolutional neural network	グラフ畳み込みニューラルネット ワーク
GPCR	G protein-coupled receptor	G タンパク質共役受容体
H	H	シャノンエントロピー (本稿ではス キャフォールド多様性の指標)
HBSS	Hanks' balanced salt solution	ハンクス平衡塩液
HDAC	histone deacetylase	ヒストン脱アセチル化酵素
hERG	human ether-a-go-go related gene	カリウムチャンネル遺伝子
IJC	Instant JChem	化学構造データベースシステムの ひとつ
JADER	Japanese adverse drug event report database	PMDA の医薬品副作用データ ベース
KLD	Kullback-Leibler divergence	カルバック・ライブラー情報量
LogP	logP	オクタノール/水分配係数の対数
MAE	mean absolute error	平均絶対誤差
MedDRA	medical dictionary for regulatory activities	ICH 国際医薬用語集
PMDA	pharmaceuticals and medical devices agency	医薬品医療機器総合機構

PR	precision-recall	適合率-再現率
PR-AUC	area under the PR curve	適合率-再現率曲線下面積
PRR	proportional reporting ratio	比例報告比（シグナル検出法）
PT	preferred term	MedDRA における基本語
RF	random forest	ランダムフォレスト（計算手法）
ROC	receiver operating characteristic	受信者操作特性
ROR	reporting odds ratio	報告オッズ比（シグナル検出法）
RMSE	root mean square error	二乗平均平方誤差
SERT	serotonin transporter	セロトニントランスポータ
SHAP	Shapley additive explanations	機械学習モデルの解釈手法
SIDER	side effect resource	市販医薬品の副作用データベース
SMILES	simplified molecular input line entry	化学構造記法のひとつ
SMQ	standardised MedDRA queries	MedDRA 標準検索式
SSRI	selective serotonin reuptake inhibitor	選択的セロトニン再取り込み阻害薬
TdP	torsades de pointes	トルサード ド ポアント（不整脈の一型）
TRP	transient receptor potential cation channel	一過性受容体電位カチオンチャネル
VEGFR	vascular endothelial growth factor receptor	血管内皮増殖因子受容体

第 1 章 グラフ畳み込みニューラルネットワークを用いた化学構造式からの親和性予測

1.1 序論

薬物の薬理作用は、特定の標的タンパク質との結合親和性に依存する。個々の化合物が標的タンパク質にどの程度強く作用するかを構造だけを見て事前に予測することは、通常経験豊富な研究者であっても不可能である。活性予測の問題はケモインフォマティクス分野で長年研究されてきており、最近では機械学習などの急速な進歩により現在の創薬研究の中心的な要素となっている。機械学習に関連する技術のなかでも、深層学習および畳み込み技術はコンピュータビジョンをはじめとする革新的な能力の中核をなすものであり、注目度はますます高まっている¹。これらの技術を化合物に適用することで、化合物構造情報を特徴ベクトルという機械処理の可能な数値の形に変換し、化合物と薬理活性の関係を説明することが可能になる。

ニューラルフィンガープリントとも呼ばれることのある特徴ベクトルと、全結合層を組み合わせたグラフ畳み込みニューラルネットワーク (GCN) モデルは、化合物表現の標準的な手法の 1 つである extended-connectivity circular fingerprint (ECFP) を用いたモデルと比較して、溶解度予測や活性予測などのタスクで予測性能が向上することが示されている^{2,3}。画像認識タスクで使用されるプーリング層と同様の、新しい層を追加定義した GCN モデルが Altae-Tran らによって 2017 年に報告され、この非線形層を含む GCN の機能が DeepChem アプリケーションとしてオープンソースライセンスの下で研究用に活用可能である^{4,5}。拡張性の高い DeepChem を活用することにより、MoleculeNet、PubChemBioAssay および ChEMBL に代表される化合物活性の実測データベースに含まれる多様なアッセイデータについて、化合物の活性の分類（活性の有無の予測）に成功したという報告は数多く存在する^{3,6-10}。

活性分類における目的変数は 1 種類または複数種類のバイナリ値である。活性あり（または活性なし）を定義するために必要な閾値は標的タンパク質により異なるはずだが、いまだに確立されたルールは存在しない¹⁰⁻¹²。閾値の設定問題に加えて、分類タスクにおいては標的タンパク質への結合の「程度」に関する重要な情報が失われる問題もある。例えば IC₅₀ 値が 1 μM の化合物と 1 nM の化合物は、同じ実験条件下では後者の方が明らかに結合力が強いにもかかわらず、分類タスクでは閾値設定によってはどちらも「活性あり」として同等に扱われる。

創薬研究の初期段階において、ハイスループットスクリーニングは重要な情報源であり、また、最適化すべき化合物を選択する上で、定量的な結果は単純な定性的データよりも価値がある。同様に、例えば大規模な仮想化合物ライブラリから限られた数の化合物を購入する場合、定量的な活性予測ができれば優先順位付けが容易になる。さらに、薬理作用を解明するためのツール化合物を同定する場合にも、定性的な予測よりも定量

的な予測の方が有用である。このため、化合物-標的タンパク質複合体の3次元情報¹³、アミノ酸配列情報¹⁴⁻¹⁶、標的タンパク質のアッセイ情報^{17,18}、化合物の結合部位近傍のアミノ酸の原子情報^{19,20}などの情報と組み合わせた化学表現を使った、様々な「定量的な親和性予測モデル（回帰モデル）」が報告されている。しかし全てのタンパク質立体構造が回帰モデル構築に適するほどの解像度で解かれているわけではなく、立体構造が判明している場合であっても、化合物の結合場所、特にアロステリックサイトを正確に推定するのは極めて困難である。

これらとは別に、化合物由来のデータのみを用いた深層学習やGCNによる回帰モデルも報告されている。Gタンパク質共役受容体（GPCR）リガンドの活性を予測するために102,400ビットという非常に長いECFPを特徴ベクトルとして用いたり²¹、化合物構造由来の2種類の特徴ベクトルを連結した複合特徴ベクトルを用いた活性予測の報告も存在する²²。しかし前者はGPCRリガンドについてのみの回帰モデルを報告しており、後者は溶解度予測や化合物-標的タンパク質複合体の立体構造の解かれたものだけを対象としているため、深層学習やGCNによる網羅的な回帰モデルという点では未だ検討の余地がある。

定量的な親和性予測は、無限に存在するともいえる目的変数を予測しようとするものである。活性分類のタスクでも、隠れ層のサイズの大きいGCNアーキテクチャのモデルのほうがそうでないものと比較して予測性能が良いと報告されているため^{12,23}、定量予測ではより大きい隠れ層が必要となることが推測される。

多くの医薬品は単純な化学構造で記述される化合物であり、化学構造はその活性を決定する重要な要素である。もちろん自由度に応じて様々な立体配座をとることができるが、多くの場合その好ましい立体配座は化学構造に固有のものであり、通常特定の立体配座のみがその薬理作用に関与していると考えられる。薬物は吸収され作用部位に到達する必要があるが、薬物の吸収と分布の背後にある物理化学的特性もまたその化学構造に由来する特徴である。

本章では、標的タンパク質への結合親和性を定量的に予測するモデルが、2次元化学構造から自動的に抽出されえる特徴量を用いて構築可能と仮定し、検討を行った。具体的には、ChEMBLから抽出した多様な標的タンパク質に対して回帰モデルを構築することを検討し、化学構造の多様性が予測性能にどのように影響するかを調査するために新たに情報理論指標を導入した。また、既知の回帰モデルと性能を比較し、さらに、セロトニントランスポーター（SERT）のバーチャルスクリーニングにより、回帰モデルの特定した化合物の*in vitro*および*in vivo*アッセイにおける薬理作用を調べた。

1.2 方法

データセット

Bosc らのプロトコル¹¹を参考に ChEMBL からデータを抽出した。まず confidence score が 6 以上、assay type が "B"、standard unit が "nM" のデータを選択した。Confidence score は標的タンパク質のアサインの信頼度を示し、"B" は *in vitro* の結合実験であることを示している。データ抽出に際し activity_comment が "Inconclusive" あるいは "Not determined" ではなく、potential duplicates は 0、data_validity_comment は "Potential author error" ではないという条件も付けた。本研究では、 IC_{50} 、 EC_{50} 、 K_i 、 K_d 、 K_m を v とした場合、それを $-\log_{10}(v)$ の値に変換したものを採用することとし、 pIC_{50} と呼ぶことにした。化合物の構造は ChEMBL から SMILES 形式 (simplified molecular input line entry system) のものを抽出し、Instant JChem 19.8.0 (IJC)²⁴ にて電荷の中和、溶媒分子と塩の除去、一部の官能基の記述の標準化を行った後、最後に canonical SMILES に変換した。この際、重原子数が 1,000 以下の SMILES のみを採用することにした。互変異性体が複数ある化合物については、最も妥当なものが ChEMBL に登録されていると仮定しそのまま使用した。化合物-標的タンパク質のペアで pIC_{50} 値が複数登録されている場合は最大値 (最も活性の高い値) を採用した。なお、本稿においては、標的タンパク質の記載は ChEMBL の提供するもの (英語表記)、あるいは、それに準じたもの (日本語表記) を用いることにする。

データ分割

各標的タンパク質についてデータセットを training-validation set (90%) と test set (10%) の 2 つのサブセットに無作為に分割した。その後 training-validation set をさらに training set (88.8%) と validation set (11.2%) に分割した。これにより分割後の 3 つのサブセットのサイズの比率を約 80:10:10 とした。Validation set をハイパーパラメータの最適化に使用し、test set は回帰モデルの予測性を評価するために使用した。

グラフ畳み込みニューラルネットワークアーキテクチャ

DeepChem に実装された RDKit²⁵ により、図 1 にあるように canonical SMILES を 1 原子あたり 75 次元のバイナリベクトルに変換した (Atom featurization)。この初期ベクトルの構成要素を表 1 に示した。本操作により異なる立体化学をもつ異性体 (シス体、トランス体など) が同一のベクトル表現となる場合は、1 つのベクトル表現が複数の立体異性体の混合物を意味するものと解釈した。初期ベクトルを入力として、グラフ畳み込み層で近傍原子の情報を足し合わせ、プーリング層で近傍原子の最大値にて原子の情報を更新した (Convolution、Pooling)。この操作を繰り返した後にデンス層に変

換した (Dense)。デンス層の同次元の要素をギャザー層で足し合わせて化合物の特徴ベクトルを生成し (Gather)、最後にギャザー層を 1 つの出力層に接続することで (Fully-connected layer) pIC₅₀ を出力する形とした。オブティマイザーには Adam、活性化関数には Rectified Linear Unit (グラフ畳み込み層) および tanh (ギャザー層) を使い、過学習防止と学習効率向上のためにバッチ正規化を用いた。

表 1 初期ベクトルの構成要素

特徴量	説明	次元
Atom type	type of atom (e.g., C, N, O, S) and "Unknown" (one-hot)	44
Degree	number of bonds the atom is connected to (one-hot)	11
Valence	number of implicit valence (one-hot)	7
Formal charge	integer electronic charge	1
Radical electrons	number of radical electrons	1
Hybridization	sp, sp ² , sp ³ , sp ³ d, or sp ³ d ² (one-hot)	5
Aromaticity	whether the atom is part of an aromatic system	1
Hydrogens	number of bonded hydrogen atoms (one-hot)	5

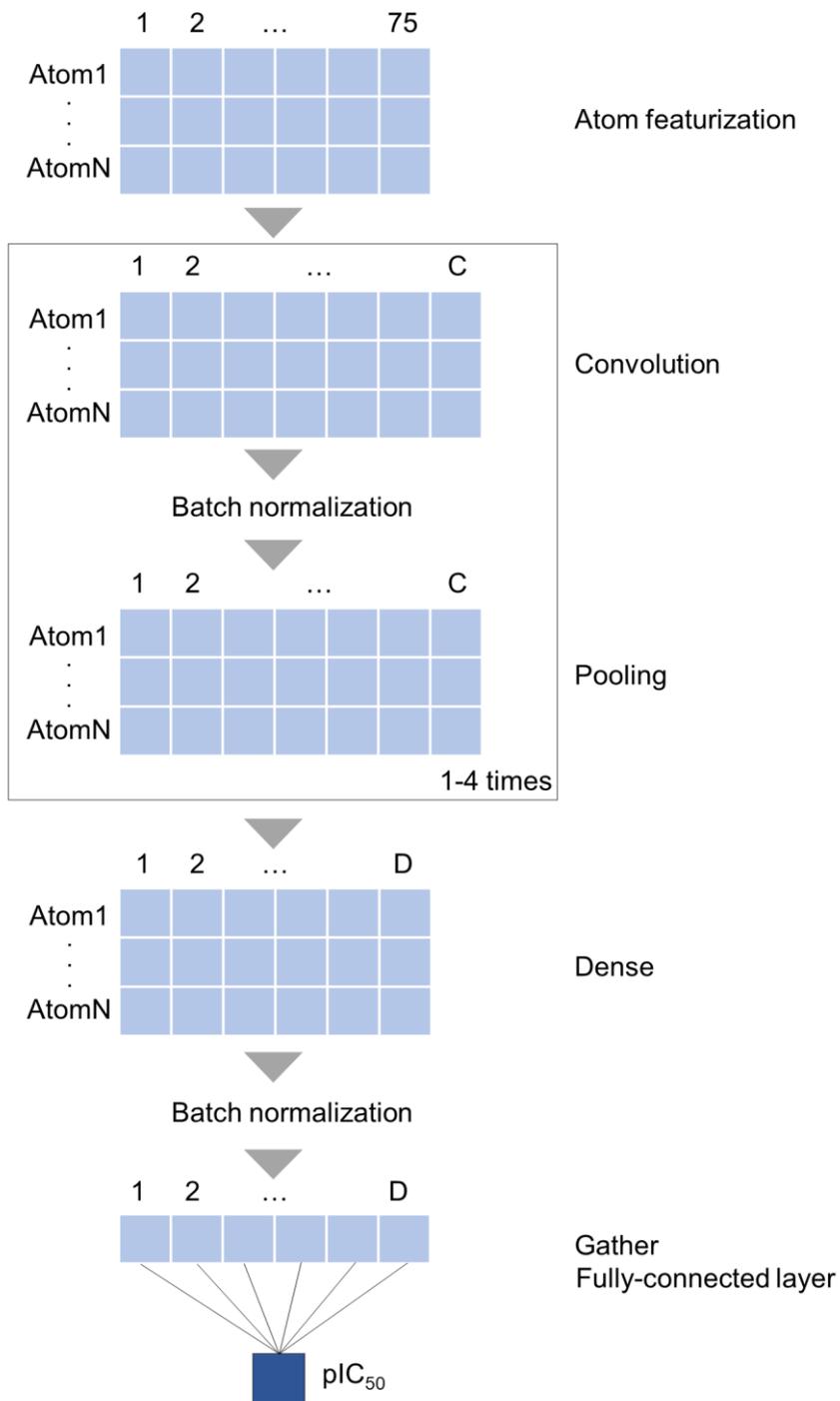


図1 グラフ畳み込みニューラルネットワークのアーキテクチャ

ハイパーパラメータの最適化と回帰モデルの学習

ハイパーパラメータを最適化するために、pyGPGO パッケージ²⁶を使用してガウス過程によるベイズ最適化を実施した。既報の GCN アーキテクチャでは主に 2-4 層のグラフ畳み込み層が使用されているが⁵⁻¹⁰、予備検討に基づき、グラフ畳み込み層が 1 層、2 層、3-4 層のアーキテクチャについて、ハイパーパラメータを個別に検討した。共分散関数に Matérn カーネル、獲得関数に「期待改善度」を用い、ベイズ最適化探索を 100 回実施した。この計算を、ランダムなシード値で初期化された異なる重みベクトルを出発点とし 4 回繰り返した。なお、データセットサイズの小さい標的タンパク質においてはパラメータ探索範囲を制限した。

回帰モデルにおいては平均絶対誤差 (MAE)、二乗平均平方誤差 (RMSE) がモデルの予測性能を表す指標として広く用いられており、説明変数が目的変数をどのくらい説明しているかを示す決定係数 (R^2) を参考にすることもある (式 (1) - (5))。 y_i は化合物の実測値、 f_i は i 番目の化合物の予測値、 y は y_i の平均値、 n は化合物の数である。

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |f_i - y_i| \quad (1)$$

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

$$R^2 = 1 - (\text{RSS}/\text{TSS}) \quad (3)$$

$$\text{TSS (total sum of squares)} = \sum_{i=1}^n (y_i - y)^2 \quad (4)$$

$$\text{RSS (residual sum of squares)} = \sum_{i=1}^n (y_i - f_i)^2 \quad (5)$$

本研究では MAE と式 (6) で定義される新しい指標 (2R2_MAE) を設定し、ハイパーパラメータを評価した。

$$2R2_MAE = (R^2 - \text{MAE}) + R^2 \quad (6)$$

2R2_MAE は以下の考え方に基づいて作成したもので、大きいほど良いことを示す。

- i) 同じ MAE を与えるパラメータ設定であれば、 R^2 の値は大きい方が良い (第 1 項が該当)。
- ii) 第 1 項が同じ値であれば、 R^2 値が大きい方が良い (第 2 項が該当)。

100 個のハイパーパラメータセットから、ホールドアウトした validation set に対して 2R2_MAE 値が最大となるハイパーパラメータセットを 1 つ選択する過程を 4 回繰り返すことで、最終的に各アーキテクチャに対して 4 つのハイパーパラメータセットを得た。グラフ畳み込み層が 1 層のみのアーキテクチャにおいては、全ての R^2 値が 0.45

より低い場合に限りハイパーパラメータの探索計算をやり直した。グラフ畳み込み層が2層のアーキテクチャでは、 R^2 値が0.40以上のハイパーパラメータセットを保持した。これらよりも深いアーキテクチャでは、グラフ畳み込み層が1層あるいは2層のハイパーパラメータセットの R^2 値のいずれよりも高い場合に、そのハイパーパラメータセットを保持した。

最終的な回帰モデルを構築する段階では、最適なハイパーパラメータセット（ただし学習エポック数を除く）を用いてまず100エポック計算した。それに引き続く100エポックにおいてホールドアウトした validation set の最小 MAE が減少しない場合は学習を終わらせた。MAE 値が減少した場合にはさらに100エポックの学習を行い、追加の100エポックで前回の最小 MAE を下回らなく（改善しなく）なるまで、総エポック数の上限を設けずに同じ手順を繰り返した。学習後、各エポックについて 2R2_MAE 値を算出し、2R2_MAE 値が最大となる回帰モデルを最終モデルとして選定した。最終モデルは DeepChem1.3.0 を用いて構築した。ハイパーパラメータ探索時には DeepChem 2.1.0 を用いたが、グラフ畳み込みアルゴリズム自体は両バージョン間で同じものである。

アンサンブルモデル

アンサンブルモデルは機械学習において一般的な手法であり、複数のモデルを構築し組み合わせることを意味する。アンサンブルモデルは、個々のモデルに比べて予測精度を向上させることが多く報告されている^{23,27-29}。本研究でもこの手法を応用し、個々の出力に重み付けをせず単純に平均化することとした。本稿中では特に断りのない限り pIC_{50} 予測値はアンサンブルモデルの出力を使用した。

化合物骨格構造（スキヤフォールド）の多様性

データセット中の化合物構造多様性は、回帰モデルの予測性能と汎化性に影響する要因の1つであると考え、化合物の側鎖を全て取り除き、重原子を全て炭素に置き換えた Murcko スキヤフォールド³⁰の分布を評価した。この評価指標として、情報理論で用いられるシャノンエントロピーの定義を応用し、式(7)、(8)に示す定量的なスキヤフォールド多様性指標 (H) を新たに導入した。 p_i はあるスキヤフォールドを含む化合物の数 (c_i) と化合物の総数 (c) との比である。

$$H = -\sum p_i \log_2 p_i \quad (7)$$

$$p_i = c_i/c \quad (8)$$

H 値が小さいほどデータセットが特定のスキヤフォールドに偏っていることを意味し、一様な分布の場合に H 値は最大となる。各データセットについてスキヤフォールドサ

イズ（スキヤフォールドを構成する炭素の数）を昇順ソートし、1ビンあたり 10,000 スキヤフォールドを含むヒストグラムに変換し、各ビン内の化合物の数をデータセット内の化合物の総数で割って確率分布に変換した（ただし 15 ビン目は 5,515 スキヤフォールドのみ）。ビンの数としては既報^{31,32}を参考に 15 ビンを設定した。15 ビンの場合、H の最大値は 3.91 である（ $H_{\max} = \log_2(15) = 3.91$ ）。

H に加え、無作為分割前後のデータセット間のスキヤフォールド分布の違いを定量化する指標として、Kullback-Leibler divergence (KLD) を採用した。

$$\text{KLD} = -\sum p_i \log_2(p_i/q_i) \quad (9)$$

q_i は未分割データセットにおけるスキヤフォールドの確率分布、 p_i は training set、validation set、test set の確率分布である。KLD は常に非負であり、 $q_i = p_i$ のとき最小値の 0 をとる。この KLD の計算には H の計算に用いたヒストグラムと同じものを使用した。

材料

シタロプラムおよび CHEMBL1377753 (5-chloro-2-(piperidin-4-yl)-1,3-benzothiazole hydrochloride、**1**) はナミキ商事（東京）から購入した。*In vivo* 試験では **1** を使用直前に生理食塩水に溶解した。*In vitro* 試験では、シタロプラムと **1** をハンクス平衡塩溶液 (HBSS; Thermo Fisher Scientific, Waltham) に溶解し、使用するまで -20°C で保管した。

HEK 細胞における SERT 基質の取り込みアッセイ

Neurotransmitter Transporter Uptake Assay Kit (R8173, Molecular Devices, San Jose, CA, USA) を用い、その説明書および以前の報告³³に従って IC₅₀ を測定した。HEK293 細胞を 96 穴黒クリアボトムプレート (#655090, Greiner, Kremsmünster, Austria) に 3.85×10^4 cells/well の密度で播種した。この細胞に、Lipofectamine 2000 (Thermo Fisher Scientific) を用いてプラスミド DNA (hSERT-pcDNA3 (Addgene #15483³⁴) または pcDNA3; 200 ng/well) をトランスフェクトした。28-30 時間培養後、細胞を直接 IC₅₀ の測定に使用した。培地を HBSS に置換したのち、薬物を含む HBSS および色素を含む HBSS を順次添加した。60 分間の静置後、Wallac 1420 ARVOsx マルチラベルカウンター (Perkin Elmer, Waltham, MA, USA) で蛍光を測定した。適用した薬物に起因する蛍光の影響を最小化するために、バックグラウンドを各濃度の薬物を含む pcDNA3 をトランスフェクトした細胞の蛍光と定義した。特異的取り込みは、各 hSERT をトランスフェクトした細胞の蛍光を対応するバックグラウンドで差し引いたものとして定義した。特異的取り込みは薬物の非存在下での取り込みを

100%として計算した。IC₅₀値は Prism 8 (GraphPad Software, San Diego, CA, USA; <https://www.graphpad.com/scientific-software/prism/>) を用いて算出した。

使用動物

全ての動物飼育および実験手順は、京都大学動物実験委員会（承認番号 19-41）の承認を受け、同委員会の倫理指針に従って実施された。成熟雄の C57BL/6J マウス（8-16 週齢、体重 22-28 g、日本 SLC、静岡）を、餌と水を自由に摂取できる状態で（各ケージに 6 匹以下）、一定の室温（24±1°C）と湿度（55±10%）で、12 時間明暗周期で飼育した。動物は各実験群に無作為に割り当てられた。全ての行動試験はその日の明周期中に実施した。

行動試験

全ての行動試験において、実験者は注入する薬物を知らない状態で試験を実施し、結果を分析した。尾懸垂試験は既報に従い実施した³⁵。マウスを馴化後、試験箱（40×40×40 cm）の天井に固定したトランスデューサー（PowerLab 2/26、AD Instruments、Dunedin、New Zealand）に尾をテープで固定してフック（試験箱の床から 35 cm）にぶら下げ無動時間を 6 分間記録した。各薬物の投与は試験の 15 分前に行った。試験中マウスの行動を記録し、尾懸垂試験中に後肢を保持したり、前肢で尾を登ったりしたマウスは解析から除外した。尾懸垂試験の 2 日以上後にオープンフィールド試験を実施した。白色アクリルボックス（50×50×50 cm）からなるオープンフィールドアリーナを使用した。各薬物の投与は試験の 15 分前に行った。各動物の行動を 10 分間カメラで記録し、記録データはビデオトラッキングシステム（ANYmaze version 4.99、Stoelting、Wood Dale、IL、USA）を用いて自動的に解析し、各セッションの総移動距離を測定した。統計学的検定はすべて Prism 8 (GraphPad Software) を用いて行った。特に断りのない限り群間比較には一元配置分散分析 (one-way ANOVA)、次いで Dunnett の多重比較検定を使用し、 $P < 0.05$ で有意に差があると判定した。

1.3 結果と考察

データセット

8つのタンパク質ファミリーに属する542の標的タンパク質からなるデータセットを前節の手順でChEMBLから選択した。さらにそこから、データセットサイズの大きい上位120種、および小規模な7種、合わせて127種の標的タンパク質からなるベンチマークデータセットを作成し、まずこのベンチマークデータセットを用いて検討を進めた。7種の標的タンパク質のデータセットサイズは1,000未満(461-739)で、残り120種のデータセットサイズは1,000以上(1,408-11,632)だった(表2)。

表2 タンパク質ファミリーと標的タンパク質

タンパク質ファミリー	標的タンパク質数	
	データセットサイズ $\geq 1,000$	データセットサイズ $< 1,000$
G protein-coupled receptor (GPCR)	33	0
Enzyme ^a	28	0
Kinase	24	0
Protease	15	0
Ion channel	7	3
Nuclear receptor	5	2
Transporter	3	2
Others	5	0

(a) ChEMBL の分類では kinase と protease は enzyme のサブカテゴリに属しているが、本研究ではこの2つのクラスは別々にカウントされ enzyme には含まれていない。

分類モデルにおいて不活性化化合物を適切に取り込むことで予測精度が向上することが示されており^{6,36}、回帰モデルを構築する際にもそれと同様に、幅広い活性値の化合物をデータセットに含むことが望ましいと考えられた。そこでアッセイの検出限界以上および以下を示す定性的実測値、例えば $IC_{50} > 100,000$ nM も、オフセットせずにそのまま使用した。pIC₅₀ 値の分布範囲は式(3)で示されるように R² に直接影響を与えるものであり、最小値と最大値はアセチル CoA カルボキシラーゼ 2 および α 1A アドレナリン受容体で、pIC₅₀ の分布範囲はそれぞれ 5.15 および 30.0 であった。30.0 という大きな値は $\log K_i = 19$ の化合物によるもので、ChEMBL に誤って登録された可能性がある(実際、その根拠となっている原著論文では 1 μ M で 19%の阻害と記載されており、

上記の可能性は高いものと推定される³⁷⁾。上記のような極端な外れ値は予測性に悪影響を及ぼす可能性があるが、本研究においては validation set の R^2 値が前節で説明した閾値より大きい場合はそのままの値を採用した。

ハイパーパラメータの最適化とモデル構築

GCN は他の機械学習手法と同様ハイパーパラメータの選択に非常に敏感である³⁸⁾。表 3 は探索したパラメータとその探索範囲を示しており、グラフ畳み込み層のサイズの上限は分類タスク⁶⁻¹⁰⁾で報告された値の 9 倍から 32 倍となった。表に記載されていないパラメータについては、DeepChem のデフォルト値を使用した。なお、小規模なデータセットでは過学習や学習不足の問題を緩和するために探索範囲を限定した。

表 3 ハイパーパラメータと探索範囲

ハイパーパラメータ	探索範囲	探索範囲 小規模データセット用
Size of the graph convolutional layers	[32-2,048]	[16-512]
Size of the dense layer	[16-2,048]	[16-512]
Number of graph convolutional layers	1, 2, 3-4	1, 2, 3-4
Learning rate	[0.00,010-0.0,020]	[0.00,010-0.0,020]
Dropout	[0.0-0.50]	[0.0-0.50]
Epoch	[20-200]	[20-200]
Batch size	[10-100]	[10-100]

R^2 、MAE および RMSE の値は、回帰モデルの性能を評価するためによく使用される指標である。 R^2 値は 1 であれば完璧な予測、1 より小さくなるにつれて予測精度が低下することを示し、直感的にモデルの性能を判断しやすい。しかし、 R^2 は式 (4) に示すように、使用するデータセットの活性値幅に影響されるため、異なるデータセットを用いて構築したモデルの予測性能の比較には利用しにくい。 R^2 とは異なり、MAE や RMSE は値が低いほど良いことを示し、両者の関係は式 (10) で表されるが、どちらの指標にもメリット、デメリットが存在する^{39,40)}。

$$\text{MAE} \leq \text{RMSE} \leq \sqrt{n} \text{MAE} \quad (10)$$

RMSE の上限は MAE にデータセットサイズ n の平方根を乗じたものに等しい。これは、データセットサイズが大きくなると RMSE が大きくなる傾向があり、異なるデー

タセットサイズ間でのモデル性能の評価が困難であることを意味している。さらに、ハイパーパラメータ探索によって得られた様々なパラメータセットの MAE、RMSE、 R^2 を調べたところ、 R^2 値が優れていても（例えば 0.54 対 0.67）、その MAE 値が最小 MAE 値よりも悪い（例えば 0.84 対 0.86）ハイパーパラメータセットも存在することが分かった。これらの理由から、本研究では MAE と 2R2_MAE を使用してハイパーパラメータセットを評価した。 R^2 値は通常[0-1]の値をとる。MAE は[0- ∞]の値をとるため R^2 値とは範囲が異なる。しかしベンチマークデータセットの MAE 値はほぼ[0-1]の範囲であったため、式 (6) にあるように R^2 や MAE を同列に適用しても、ハイパーパラメータセットの現実的な評価という観点からは、その影響は極めて限定的であると推定された。

式 (6) の適応にあたっての懸念点としては、2R2_MAE は R^2 と MAE のバランスに基づくものであるため、 R^2 が高く（望ましい）、MAE が高く（望ましくない）、結果、2R2_MAE が高い（望ましいという判定）となる場合が考えられることだった。この問題を調べるために、i) 最大 2R2_MAE、ii) 最小 MAE、という基準に基づいて選択されたハイパーパラメータの組み合わせによって、validation set の MAE と R^2 の値がどのように影響されるかを比較分析した。その結果、2R2_MAE 最大値に基づいて選択したハイパーパラメータセットでは、最小 MAE 値基準に基づくものよりも MAE が若干悪く（平均増加量 0.0046、最大増加量は 0.092）、 R^2 は良くなる傾向があった（平均増加量は 0.0082、最大増加量は 0.14）。全体的には 2R2_MAE の基準に基づくハイパーパラメータの選択によりベンチマークデータセットにおいて妥当なモデルを選択できたと推察された（図 2a、b）。

グラフ畳み込み層とデンス層のサイズは、標的タンパク質によって異なるものの、ハイパーパラメータ探索範囲の上限に近づく傾向があった。この結果は、隠れ層のサイズが大きい分類モデルが良好な分類性能を示した既報^{12,23}と一致する傾向であった。

最終モデルを得るため、2R2_MAE 最大値を与える最適なハイパーパラメータセット（学習エポック数を除く）を用いて固定ランダムシードで training set を再学習した。一部のモデルでは、例えばハイパーパラメータ最適化の過程で $R^2 = 0.53$ を示したハイパーパラメータセットが再学習後には $R^2 = 0.16$ になるなど、予測性能が再現しないものがあつた。全体の約 1.2% のモデルに再現性の欠如が見られたが、アンサンブル学習時にそのようなモデルは除外し、1 つの標的タンパク質あたり 6-9 個の個別モデルを得た。

一般に、隠れ層の数が多いアーキテクチャほど複雑で高度な特徴の抽出が可能であり、予測性能も良くなる。一方、本研究で良好な性能を示したモデルの多くはグラフ畳み込み層が 1 層であり、グラフ畳み込み層が 4 層のモデルは、ハイパーパラメータ探索過程でどの標的タンパク質においてもグラフ畳み込み層が 3 層のモデルを上回らなかった。この矛盾しているように見える現象の説明として考えられるのは、プーリング層にて化合物の特徴を抽出するだけでなく、その情報を必要以上にそぎ落とした可能性である。GCN は本来ラプラシアン平滑化の一種であり、ラプラシアン平滑化を繰り返すと

化合物の局所的化学環境が識別できなくなる可能性が指摘されており⁴¹、これが今回の結果を説明できる可能性がある。一方で、層の数が増えるにつれてより遠い原子の情報を取り込むことができるグラフ畳み込みの特徴を生かすためには、現在のアーキテクチャに改良の余地が残されていると考えられる。

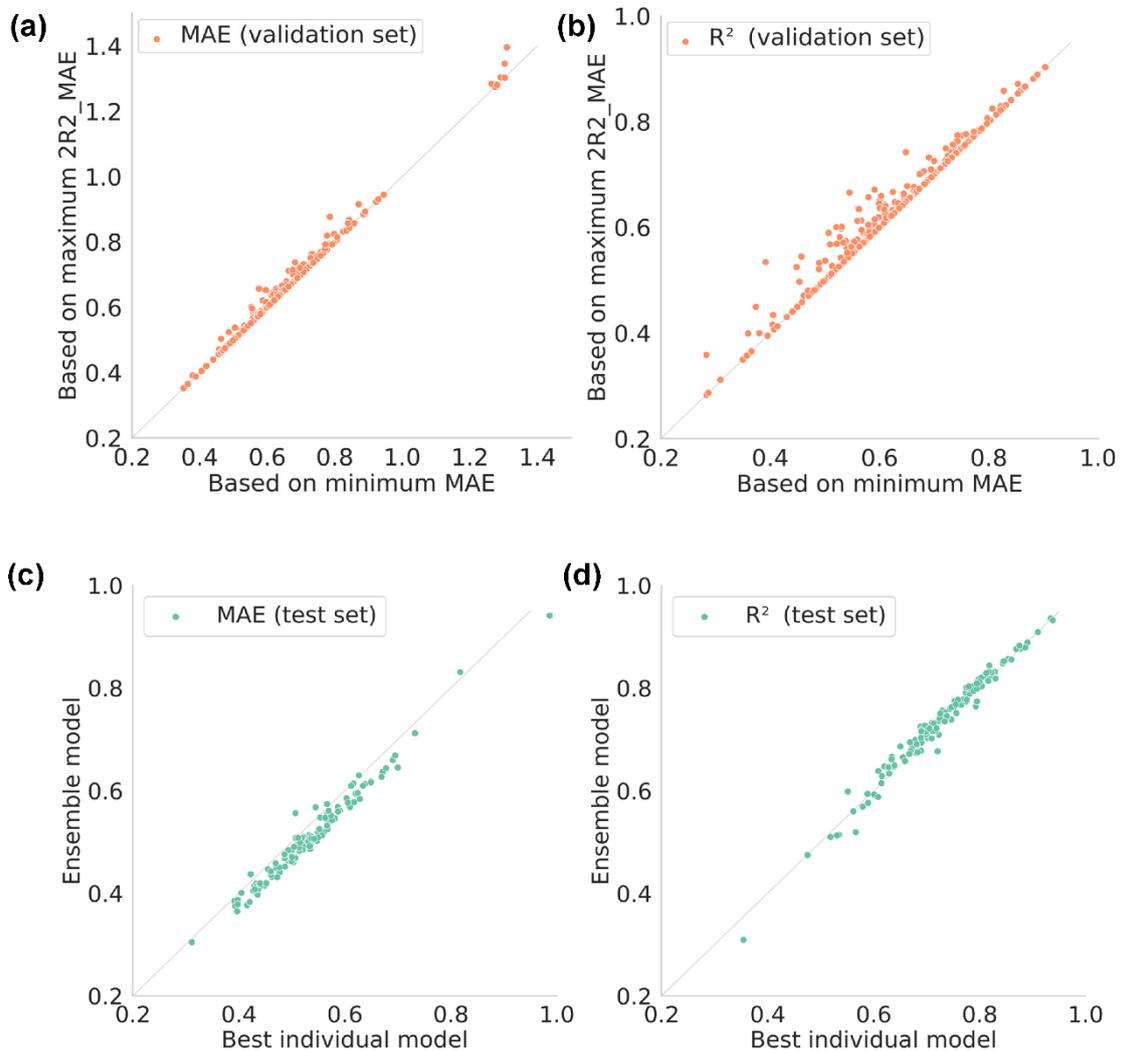


図2 2R2_MAE 指標に基づくモデル選択とアンサンブルモデルの予測性能への影響

(a, b) 最小 MAE および最大 2R2_MAE の基準に従って選択されたハイパーパラメータセットによる、validation set での MAE (a) と R^2 (b) を比較したもの。対角線上の点は、両方の基準で同じハイパーパラメータセットが選択されたことを表す。どちらの基準でも MAE 値に顕著な差は観察されなかった。最大 2R2_MAE 基準でハイパーパラメータセットを選択した場合、 R^2 値が向上する傾向が見られた。

(c, d) アンサンブルモデルと最良個別モデルの test set に対する MAE (c) と R^2 (d) の比較。アンサンブルモデルでは MAE 値が減少し、 R^2 値が有意に増加した。

アンサンブルモデル

1つの標的タンパク質あたり 6-9 個の個別モデルを得、その個々の回帰モデルの予測値を重み付けせずに平均化し、アンサンブル予測値を生成した。図 2c、d は test set における MAE と R^2 を比較したものである。図 2c において対角線より下の領域はアンサンブルモデルの性能が良いことを示しており、120 個の標的タンパク質がこの領域に入った。図 2d では、対角線より上の領域においてアンサンブル予測が最良の個別モデルよりも良い結果であることを示し、94 個の標的タンパク質がこの領域に入った。最良の個別モデルとアンサンブルモデルの、MAE と R^2 分布の平均の違いについて統計的有意性を片側 Wilcoxon 符号順位検定で検定した。帰無仮説は $P = 5.51 \times 10^{-20}$ と 1.02×10^{-8} で棄却され、アンサンブルモデルがより低い平均 MAE とより高い平均 R^2 を与えることが示された。アンサンブルモデルによる性能向上は、他の研究^{23,27-29} で得られた結果とも一致するものであり、この結果は、準最適なハイパーパラメータの組み合わせが多数存在し、類似の組み合わせであっても化合物の異なる特性を捉えている可能性を示唆している。

経験的に $MAE < 0.6$ または $R^2 > 0.6$ のいずれかを満たすモデルを良いモデルと見なすことができるが、本研究では 86% (111 標的タンパク質) および 91% (116 標的タンパク質) のモデルがそれぞれ $MAE < 0.6$ および $R^2 > 0.6$ の基準を満たした。表 4 にアンサンブルモデルの MAE と R^2 を示した。

図 3 は 8 つのタンパク質ファミリーのアンサンブルモデルの予測性能を示したものである。全てのタンパク質ファミリーで 75 パーセントにおける MAE は 0.6 以下だった。0.8 を超えたのはニコチン性アセチルコリン受容体 $\alpha 4\beta 2$ とヒト免疫不全ウイルス 1 型プロテアーゼの 2 種のみであり、原因として約 2% の化合物が著しく低い予測 pIC_{50} 値を示したため MAE が増大したことが考えられた。Validation set と test set の MAE 値は training set の MAE 値よりも大きくなる傾向があり、ある程度の過学習が起こったことを示唆しているが、多くの MAE 値は本研究で基準とした $MAE < 0.6$ を満たしていた。なお、ChEMBL では kinase と protease は enzyme に含まれるが、本研究では個別に集計することとした。

表 4 標的タンパク質と回帰モデルの詳細

種類 ^a	標的タンパク質	サイズ ^b	H ^c	畳み込み層 のサイズ	デンス層 のサイズ	MAE	R ²
G	Orexin receptor 1	2,852	3.13	52-1,489	688-2,028	0.36	0.79
G	Serotonin 7 (5-HT7) receptor	2,395	3.36	99-1,978	420-2,031	0.42	0.74
G	Orexin receptor 2	3,079	3.11	248-1,708	621-2,026	0.45	0.71
G	Cannabinoid CB1 receptor	6,966	2.90	99-1,460	979-2,031	0.46	0.76
G	G protein-coupled receptor 44	2,655	2.52	64-1,238	740-1,889	0.47	0.65
G	Dopamine D1 receptor	2,774	2.82	232-1,014	999-2,041	0.47	0.71
G	Melanin-concentrating hormone receptor 1	3,124	3.48	97-1,909	896-2,038	0.47	0.60
G	Serotonin 6 (5-HT6) receptor	3,106	2.78	50-2,027	227-2,045	0.47	0.72
G	Corticotropin releasing factor receptor 1	2,471	1.87	295-1,561	829-1,998	0.48	0.63
G	Adenosine A2a receptor	6,206	2.91	317-1,862	788-1,974	0.49	0.69
G	Histamine H3 receptor	4,021	2.88	114-1,245	964-1,977	0.50	0.65
G	Dopamine D2 receptor	11,632	3.35	129-1,148	1,038-1,929	0.50	0.67
G	Dopamine D3 receptor	4,983	3.45	206-1,982	1,297-1,943	0.50	0.72
G	Serotonin 2a (5-HT2a) receptor	6,189	3.25	87-1,758	882-1,992	0.51	0.69
G	Serotonin 2c (5-HT2c) receptor	3,807	2.81	91-1,703	558-1,994	0.51	0.72
G	Adenosine A3 receptor	4,502	3.00	99-1,279	757-2,031	0.51	0.72
G	Delta opioid receptor	6,946	3.74	157-1,764	411-1,911	0.52	0.78
G	Cannabinoid CB2 receptor	6,413	2.71	150-1,577	371-1,980	0.52	0.70
G	Kappa opioid receptor	6,645	3.60	211-2,044	1,166-1,953	0.52	0.75
G	Serotonin 1a (5-HT1a) receptor	8,261	3.31	97-1,860	697-2,021	0.52	0.69
G	Metabotropic glutamate receptor 5	3,535	2.10	175-1,604	637-1,827	0.53	0.67
G	Adenosine A1 receptor	7,650	2.78	291-2,036	1,365-2,028	0.54	0.65
G	Cholecystokinin B receptor	2,459	3.74	163-1,253	722-1,994	0.55	0.75
G	Dopamine D4 receptor	2,622	3.11	125-1,999	1,414-1,980	0.57	0.51
G	Mu opioid receptor	8,904	3.71	75-1,282	897-1,808	0.58	0.73
G	Melanocortin receptor 4	2,902	3.72	107-1,917	251-1,932	0.60	0.61
G	Neurokinin 1 receptor	2,888	3.71	45-963	919-2,010	0.61	0.70
G	Endothelin receptor ET-A	2,486	3.58	274-1,979	572-1,928	0.61	0.77
G	Muscarinic acetylcholine receptor M2	2,135	3.22	45-1,457	588-1,927	0.62	0.83
G	Gonadotropin-releasing hormone receptor	1,726	3.72	256-1,486	1,065-2,006	0.62	0.94
G	Cholecystokinin A receptor	2,423	3.70	50-1,719	583-1,905	0.64	0.51
G	Muscarinic acetylcholine receptor M1	2,465	2.78	88-1,308	562-2,014	0.65	0.80
G	Alpha-1a adrenergic receptor	2,193	3.43	108-1,061	562-2,002	0.66	0.67
E	Acetyl-CoA carboxylase 2	3,136	2.56	99-1,867	849-2,035	0.30	0.68
E	Poly [ADP-ribose] polymerase-1	3,101	2.91	165-1,206	679-2,001	0.38	0.82
E	Cholinesterase	3,011	3.50	222-1,843	920-2,022	0.39	0.82
E	Nicotinamide phosphoribosyltransferase	2,342	3.12	166-1,592	679-1,945	0.41	0.68
E	PI3-kinase p110-alpha subunit	5,699	3.38	161-1,725	724-1,968	0.43	0.78
E	11-beta-hydroxysteroid dehydrogenase 1	3,410	2.54	154-1,067	862-1,917	0.43	0.79
E	Arachidonate 5-lipoxygenase	2,953	2.09	94-1,574	649-1,837	0.45	0.64
E	Carbonic anhydrase XII	2,664	1.50	120-1,436	817-1,904	0.46	0.69
E	Phosphodiesterase 10A	4,506	3.29	231-1,502	854-2,033	0.49	0.75

(続き)							
E	Cyclooxygenase-1	4,635	1.69	58-1,330	553-1,897	0.49	0.47
E	Protein-tyrosine phosphatase 1B	3,263	3.04	264-1,073	577-1,979	0.50	0.59
E	Carbonic anhydrase I	5,305	1.71	366-1,904	1,150-1,814	0.50	0.74
E	PI3-kinase p110-delta subunit	2,647	3.31	46-1,365	640-1,991	0.50	0.71
E	Carbonic anhydrase IX	3,431	1.59	119-1,642	781-1,966	0.50	0.70
E	Acetylcholinesterase	9,737	3.26	103-1,299	687-2,044	0.51	0.71
E	Integrase	2,545	2.37	177-1,833	756-1,742	0.51	0.82
E	Carbonic anhydrase II	5,808	1.65	43-1,735	798-1,913	0.53	0.76
E	Cytochrome P450 19A1	2,196	1.62	272-1,768	681-1,967	0.54	0.73
E	Butyrylcholinesterase	3,451	3.25	67-1,645	938-2,014	0.55	0.73
E	Human immunodeficiency virus type 1 reverse transcriptase	2,310	1.70	179-1,433	825-1,968	0.55	0.68
E	Protein farnesyltransferase	2,653	3.34	190-1,638	333-1,947	0.56	0.77
E	PI3-kinase p110-gamma subunit	2,596	3.23	165-2,019	556-1,959	0.56	0.59
E	Monoamine oxidase A	3,666	1.61	269-1,281	741-1,985	0.58	0.63
E	Gamma-secretase	1,932	3.16	32-1,634	1,130-1,980	0.59	0.67
E	Anandamide amidohydrolase	3,242	2.72	254-1,871	1,339-2,005	0.59	0.86
E	Cyclooxygenase-2	5,085	1.72	72-1,441	562-1,915	0.61	0.68
E	Monoamine oxidase B	4,283	1.62	73-1,996	721-2,043	0.63	0.66
E	Dihydrofolate reductase	2,309	1.74	103-1,721	1,035-1,998	0.71	0.57
I	HERG	9,198	3.46	156-1,455	1,038-1,888	0.38	0.66
I	Voltage-gated potassium channel subunit Kv1.5	739	3.27	94-430	151-495	0.39	0.53
I	Sodium channel protein type IX alpha subunit	5,677	3.13	125-1,331	887-1,907	0.42	0.72
I	Vanilloid receptor	2,856	2.56	45-1,898	342-1,728	0.46	0.78
I	Transient receptor potential cation channel subfamily M member 8	721	2.74	160-510	67-489	0.51	0.76
I	Apoptosis regulator Bcl-2	2,542	2.31	156-1,437	1,177-2,001	0.51	0.88
I	Neuronal acetylcholine receptor protein alpha-7 subunit	2,195	2.12	114-1,796	989-1,984	0.51	0.75
I	Transient receptor potential cation channel subfamily A member 1	621	1.81	41-399	269-467	0.56	0.68
I	P2X purinoceptor 7	2,247	1.79	68-664	591-1,719	0.63	0.31
I	Neuronal acetylcholine receptor; alpha4/beta2	2,625	1.56	36-1,926	690-2,000	0.83	0.51
K	Nerve growth factor receptor Trk-A	2,587	3.36	196-1,694	856-1,938	0.37	0.71
K	Insulin-like growth factor I receptor	3,019	3.76	98-1,731	495-1,642	0.40	0.85
K	Tyrosine-protein kinase JAK1	4,345	3.19	233-1,827	934-1,989	0.41	0.81
K	Serine/threonine-protein kinase mTOR	4,414	3.60	420-1,710	652-1,838	0.41	0.81
K	Tyrosine-protein kinase JAK2	5,933	3.29	158-1,676	1,132-1,963	0.42	0.80
K	Serine/threonine-protein kinase B-raf	3,234	3.54	68-1,500	586-1,994	0.44	0.80
K	Hepatocyte growth factor receptor	3,276	3.66	126-2,011	687-1,954	0.44	0.80
K	Tyrosine-protein kinase SYK	3,368	3.39	52-1,737	1,146-1,939	0.44	0.80
K	Fibroblast growth factor receptor 1	2,473	3.31	87-1,726	572-1,909	0.45	0.83
K	Serine/threonine-protein kinase PIM1	4,022	2.93	193-1,494	1,144-1,996	0.45	0.88
K	Fibroblast growth factor receptor 3	1,408	3.04	153-2,000	519-1,916	0.46	0.59
K	Tyrosine-protein kinase SRC	4,183	3.39	89-2,031	550-1,975	0.47	0.83
K	Tyrosine-protein kinase receptor FLT3	2,274	3.50	101-1,617	833-1,988	0.48	0.74
K	Epidermal growth factor receptor erbB1	7,122	3.24	159-1,878	791-1,957	0.48	0.80

(続き)

K	Serine/threonine-protein kinase AKT	3,114	3.59	97-1,912	824-1,887	0.48	0.77
K	MAP kinase p38 alpha	4,518	3.37	91-2,034	463-1,837	0.48	0.73
K	Serine/threonine-protein kinase PIM2	2,406	2.93	171-1,790	590-1,910	0.48	0.72
K	Tyrosine-protein kinase JAK3	3,225	3.15	227-1,430	781-1,926	0.49	0.77
K	Vascular endothelial growth factor receptor 2	8,936	3.40	133-2,044	1,275-1,998	0.49	0.69
K	MAP kinase ERK2	3,712	3.28	589-1,566	920-1,947	0.50	0.80
K	Cyclin-dependent kinase 2	2,352	2.87	74-1,508	1,025-1,893	0.51	0.74
K	Serine/threonine-protein kinase Aurora-A	2,949	3.39	486-1,687	668-2,009	0.56	0.75
K	Tyrosine-protein kinase ABL	2,249	3.17	201-1,741	522-2,047	0.56	0.83
K	Glycogen synthase kinase-3 beta	2,702	2.76	64-1,715	894-1,802	0.61	0.58
N	Thyroid hormone receptor alpha	461	1.56	109-441	277-508	0.38	0.82
N	Glucocorticoid receptor	2,293	3.26	273-1,982	728-2,034	0.48	0.78
N	Peroxisome proliferator-activated receptor gamma	3,018	3.32	192-1,412	951-2,015	0.51	0.72
N	Vitamin D receptor	546	1.91	39-366	234-464	0.51	0.88
N	Androgen Receptor	2,597	1.68	193-1,865	1,295-1,942	0.56	0.72
N	Estrogen receptor alpha	2,878	2.83	61-1,188	746-1,494	0.57	0.76
N	Estrogen receptor beta	2,229	2.59	358-1,503	584-1,666	0.67	0.68
P	Cathepsin D	2,568	3.29	59-1,946	952-1,968	0.39	0.85
P	Matrix metalloproteinase-1	3,746	2.49	79-2,043	915-1,868	0.42	0.81
P	ADAM17	2,410	3.15	151-1,434	703-1,976	0.42	0.89
P	Cathepsin S	2,309	3.44	104-1,386	1,152-2,025	0.46	0.79
P	Thrombin	6,703	3.50	220-1,779	1,399-1,894	0.49	0.84
P	Beta-secretase 1	7,554	3.17	295-1,934	737-1,971	0.50	0.74
P	Dipeptidyl peptidase IV	4,517	2.36	108-1,793	762-2,015	0.53	0.77
P	Matrix metalloproteinase 13	2,844	2.83	161-976	746-1,820	0.54	0.74
P	Trypsin I	2,501	3.34	202-1,979	630-1,934	0.57	0.91
P	Leukocyte elastase	2,808	2.64	239-1,702	923-1,984	0.57	0.93
P	Matrix metalloproteinase-2	4,085	2.76	133-1,732	434-2,039	0.58	0.81
P	Renin	3,010	3.64	82-1,840	536-1,916	0.58	0.74
P	Matrix metalloproteinase 9	3,149	2.65	240-1,503	210-1,824	0.59	0.70
P	Coagulation factor X	6,307	3.49	310-1,191	535-1,952	0.64	0.80
P	Human immunodeficiency virus type 1 protease	4,571	3.65	147-1,032	901-1,997	0.94	0.65
T	Potassium-transporting ATPase	532	1.58	207-501	221-496	0.40	0.52
T	GABA transporter 1	576	2.76	22-334	179-509	0.44	0.86
T	Dopamine transporter	5,908	2.20	318-1,930	691-2,033	0.48	0.76
T	Norepinephrine transporter	4,342	1.91	166-1,713	1,048-1,954	0.50	0.70
T	Serotonin transporter	7,886	2.73	178-1,129	677-1,806	0.51	0.75
O	Histone deacetylase 1	4,239	2.49	198-2,022	1,533-2,008	0.41	0.74
O	Bromodomain-containing protein 4	2,208	2.87	424-1,484	477-1,470	0.41	0.82
O	Histone deacetylase 6	2,725	1.97	119-1,920	605-1,926	0.42	0.82
O	p53-binding protein Mdm-2	2,346	3.51	33-1,936	1,277-1,950	0.42	0.88
O	Sigma opioid receptor	3,094	2.53	48-1,366	502-1,814	0.53	0.56

(a) G; GPCR. E; Enzyme. I; Ion channel. K; Kinase. N; Nuclear receptor. P; Protease. T; Transporter. O; Other.

(b) データセットサイズ (化合物数) (c) 式 (7) で定義されるスキャフォールド多様性の指標

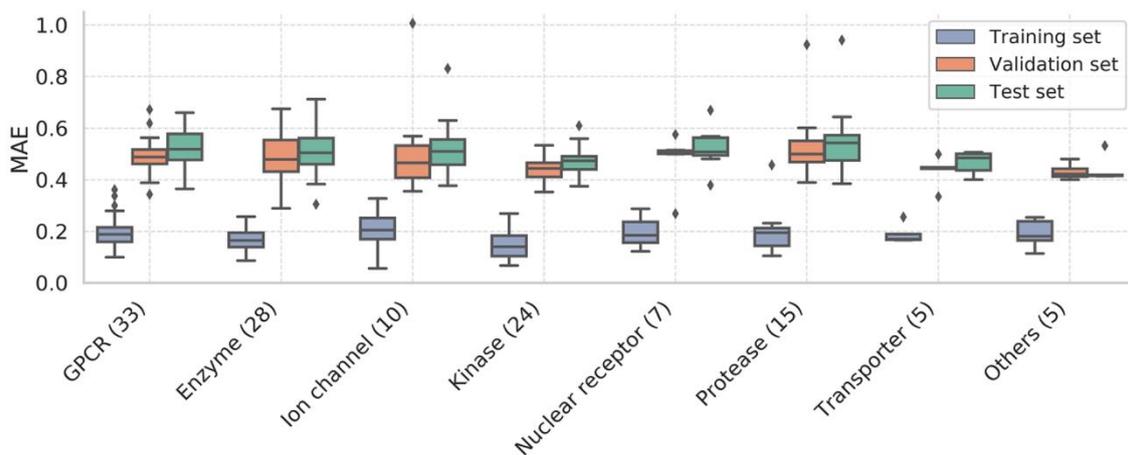


図3 タンパク質ファミリー別のアンサンブルモデルの MAE

箱の中の横線は中央値、ひげの端は MAE の最大値と最小値、箱の底辺と上辺は 25 パーセントイルと 75 パーセントイル、ひげの外にある点は外れ値であることを示している。X 軸の名称の後の数字は各ファミリーに含まれる標的タンパク質の数を示している。本章を通じてデータサブセットに対して同じカラーコードを使用している。

畳み込みニューラルネットワーク (CNN) モデルとの比較

化合物の2次元構造をあたかも化合物の写真であるかのようにみだててデータ処理を行い、毒性や薬理活性の定性的・定量的予測に利用する CNN モデルが報告されている。入力画像としては2次元の化学構造^{42,43}や、3次元画像として描かれた化合物のスナップショット⁴⁴が用いられている。各化合物の特徴ベクトルは画像データに対して畳み込み演算を行うことで抽出される⁴²⁻⁴⁵。

Cortés-Ciriano らは画像処理に広く用いられているアーキテクチャの ResNet-52 と VGG-19 を用い、ChEMBL (release 23) から抽出した 25 種類の標的タンパク質に対しスクリーニングされた化合物の2次元化合物画像データを入力として、pIC₅₀を予測する CNN モデル (KekuleScope) を報告した⁴³。そこでこの KekuleScope のデータセットに対して本研究の GCN 手法を適応した回帰モデルを構築し、RMSE 値を比較した。結果、KekuleScope のそれと同等の予測性能を示し、また、間接的ではあるが、Cortés-Ciriano らが同時に報告したランダムフォレスト (RF) モデルや全結合型ディープニューラルネットワーク (FNN) モデルの予測性能にも近かった (表 5)。また、ChEMBL (リリース 25) を用いて構築した 22 種の回帰モデルの RMSE 値を比較したところ、全ての RMSE 値が KekuleScope、RF および FNN の値と同等かそれ以下であり、予測性能が同等もしくは勝っていることがわかった。

表 5 KekuleScope との比較

	モデル	RMSE ^a	RMSE ^b
	CNN	0.76±0.078	
kekuleScope	RF	0.68±0.070	
	FNN	0.71±0.076	
本研究	GCN	0.74±0.091	0.49±0.11

(a) KekuleScope データセット

(b) ChEMBL (リリース 25)

化合物骨格構造 (スキヤフォールド) の多様性とデータセットサイズの影響

データセット (特に training set) に含まれる化合物構造の多様性は、モデルの適用領域を決定づける要因の1つとして考えられる。構造の多様性の定義として広く受け入れられているのが Murcko の化合物骨格構造 (スキヤフォールド) であり³⁰、データセットの構造多様性を評価するためにこのスキヤフォールドを適用したり^{23,46,47}、目的

の活性を発現するために有利なスキヤフォールドを化合物デザインに取り入れたという報告が多く存在する⁴⁸。

本研究に用いた ChEMBL のデータセットを調べたところ、インスリン様成長因子 I 受容体 (707 スキヤフォールド) から炭酸脱水酵素 XII (356 スキヤフォールド) まで、145,515 種のユニークなスキヤフォールドの存在が確認された。そこで、各標的タンパク質にこのスキヤフォールドが何種類含まれるか可視化することで、スキヤフォールド多様性と予測性能の関係を解析することを考えたが、標的タンパク質 A と B が同じスキヤフォールドを有する場合であっても、そのスキヤフォールドの分布が等しいかどうかは別問題であることに気づいた。なぜなら、例えば標的タンパク質 A はほぼ小さなスキヤフォールドの化合物で構成され、標的タンパク質 B はほぼ大きなスキヤフォールドの化合物から構成されるといった場合が考えられるからである。そこで、スキヤフォールド多様性とスキヤフォールド分布の関係を分析するために、様々な連続・不連続データ分布を定量的な情報量に変換可能なシャノンエントロピー (H) を、スキヤフォールド多様性の指標として設定した (式 7)。

スキヤフォールド出現数をヒストグラムで表現する場合、同じビン数、同じ範囲に分割すれば、H 値はビン間隔の大きさに依存しない。本研究のデータセットにおいては 145,515 種のスキヤフォールドを炭素数の少ないものから昇順ソートし 15 ビンに分割しているが (1 ビンあたり 10,000 スキヤフォールド、15 番目のビンのみ 5,515 スキヤフォールド)、15 ビンのヒストグラムの H は 0 から 3.91 の値をとる。ChEMBL 自身の H は 2.82 であり、一様分布 (H = 3.91) から逸脱していることがわかった。また、ヒストグラムを変換した確率分布と合わせて考えることにより、この偏りは小さいスキヤフォールドへの偏りと関連していることがわかった (図 4a)。SERT は ChEMBL と同様の確率分布を示し、H の値も類似していた (2.73)。一方、SERT と同じくセロトニンを認識するセロトニン 1A 受容体 (5-HT1A) は、2 番目から 7 番目までのビンで均一分布を示し、H 値は 3.31 (図 4b) であった。全ての H 値は表 4 に示した通りであり、図 4c のバイオリンプロットは各タンパク質ファミリーの H 値の分布を可視化したものである。水平方向の破線は 2.82 (ChEMBL の H 値) を示している。GPCR とキナーゼの H 値分布の第 1 四分位は 2.82 より大きく、多くの標的タンパク質で多様なスキヤフォールドを万遍なく有していることが示された。酵素、イオンチャネルおよび核内受容体は、多様性の高い標的タンパク質から低いものまで幅広く存在していた。

創薬研究の初期スクリーニングにおいては、様々な化合物を選択することが重要であると指摘されており^{49,50}、training set のスキヤフォールド多様性が高いほど、回帰モデルの適用領域は大きくなる⁶。この観点から、本研究のデータセットについて H と MAE の関係を評価したところ、図 4d に示すように気になる相関性は見られなかった。また、スピアマンの順位相関係数は training set で -0.11、validation set で -0.023、test set で -0.062 であった。これらの結果は、スキヤフォールドの多様性に関係なく、どのようなデータセットに対しても本回帰モデルが高精度の予測を行える可能性を示唆している。

分割後のデータセット (training set、validation set および test set) 間でのスキヤフォールド分布の違いも回帰モデルの性能に影響を与える可能性がある。というのも、分割後のデータセット間でスキヤフォールドが重複しないように調整したモデル構築アプローチにおいて、予測精度が低下する傾向が報告されているからである^{10,28}。また、無作為にデータセット分割を行った場合でも、特に小規模なデータセットでは、データセット間のスキヤフォールド分布が意図せず不均一になる可能性がある。そこで、データセット間のスキヤフォールド分布の違いを定量化するために、KLD (式 (9)) を導入した。比較対象の分割データセット間でスキヤフォールド分布が同じであるとき KLD は最小値の 0 となる。H 値を用いる場合、異なるスキヤフォールド分布であっても同じ H 値となることがあり得るが、KLD 値は分布が異なれば値も異なるため分布の違いを評価しやすい利点がある。

図 4e は KLD と MAE の関係を示したものである。プラス記号は 1,000 化合物以下の小規模なデータセットであることを意味する。予想通り、training set (青い点) は全ての標的タンパク質で KLD 値が非常に小さく、分割前と分割後のスキヤフォールド分布がほぼ同じであることがわかった。1,000 化合物以上存在するデータセットから分割された validation set と test set の KLD 値の多くは 0.07 以下の類似した KLD 分布で、ランダム分割関数が期待通りに機能したことが伺えた。小規模なデータセットでは KLD 値が 0.07 より大きいものが多く、スキヤフォールド分布が意図せず偏っていることが示唆された。また、今回調査した 127 の標的タンパク質では、KLD と MAE に相関は見られず、スピアマン順位相関係数は training set、validation set、test set でそれぞれ 0.094、0.16、0.068 であった。興味深いことに、小規模データセットの標的タンパク質において、KLD 値が大きいにもかかわらず MAE は 0.1-0.6 の範囲に収まっていた。このことは、本研究で扱った範囲内のスキヤフォールド分布の違いであれば回帰モデル性能に明確な影響を与えないことを示唆するものである。

図 4f は training set のサイズと test set の MAE を比較したものである。ここでもデータセットサイズと MAE の間に明確な相関は見られず、特定のタンパク質ファミリーへの偏りも見られなかった。しかし、別の報告で指摘されているように⁵¹、小規模なデータセットでは過学習や過少学習、相対的なノイズの影響といった固有の問題があるため、回帰モデルの予測性能の評価には一層の慎重さを要すると考えられる。なお、データの無作為分割と予測性能への影響をみるため、training-validation set の無作為分割を 2 回繰り返し、本研究で主に用いたデータセットと併せて合計 3 つの分割データセットを作成し回帰モデルを構築したところ、これら 3 セットの予測性能はほぼ同程度であったため、無作為分割の予測性能への影響は限定的であると推察された。

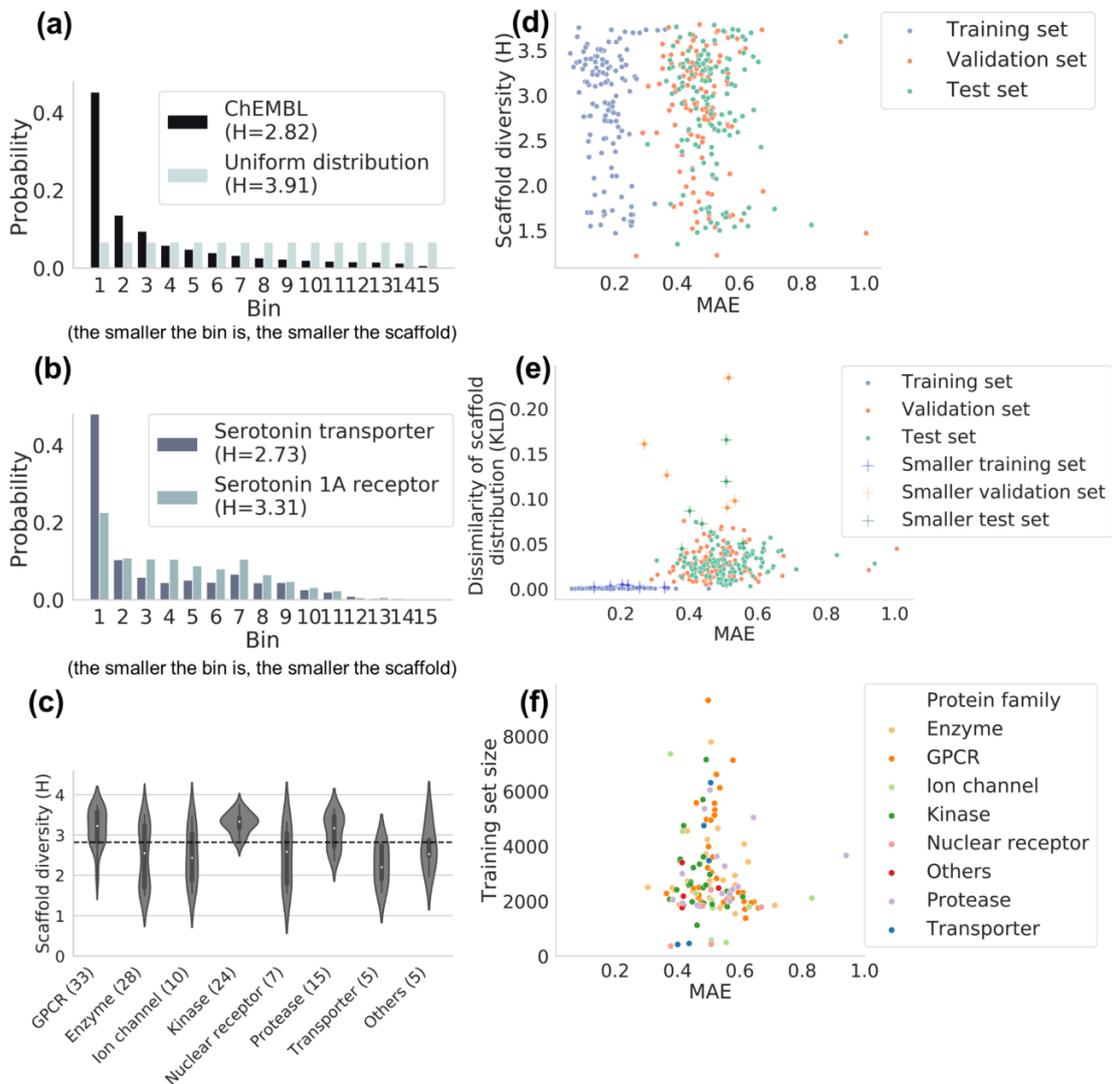


図 4 データセットのスキヤフォールド多様性が予測性能に与える影響

(a) ChEMBL のスキヤフォールド確率分布。一様分布と比較して ChEMBL はより小さなスキヤフォールドに偏り H 値が小さくなっている。

(b) セロトニントランスポーター (SERT) とセロトニン 1A 受容体 (5-HT1A) のデータセットに含まれるスキヤフォールドの確率分布。H 値は 5-HT1A の方が大きく、SERT よりも分布が幅広い。

(c) タンパク質ファミリー別の H 値分布。X 軸の各名称の後の数字は各ファミリーの標的タンパク質の数を示す。

(d-f) スキヤフォールド多様性 (d)、スキヤフォールド分布の非類似度 (e) および training set のデータサイズ (f) の MAE への影響。

バーチャルスクリーニング

構築した回帰モデルの性能をさらに評価するために、1.2節の方法の節で述べた手順の assay_type = B フィルターを除いた条件にて ChEMBL の 1,777,353 種類の化合物について SERT 活性を計算した。市販の選択的セロトニン再取り込み阻害剤 (SSRI) の オクタノール/水分配係数 (logP) 値は 2.29-5.15 (IJC で計算) の範囲であるため、logP フィルター (logP > 2) を用いて化合物を絞り込んでいった。さらに、SERT の予測 $pIC_{50} \geq 7.5$ 、5-HT1A 受容体の予測 $pIC_{50} \leq 6.0$ 、モノアミン関連タンパク質 (SERT、その他のセロトニン受容体、ドパミン受容体とトランスポーター、およびアドレナリン受容体) およびオピオイド受容体に対する親和性実測報告のない化合物の中から、目視検査の後、合成品が利用可能であった化合物 **1** (図 5a) を薬理活性試験に供した。化合物 **1** には SERT に対する薬理試験報告は存在しなかったが、比較的親和性実測値の高いものとして transient receptor potential canonical 4 ($pIC_{50} = 8.10$) および nuclear factor erythroid 2-related factor 2 ($pIC_{50} = 6.19$) が報告されていた。

In vitro アッセイ

予測 IC_{50} が約 10 nM ($pIC_{50} = 7.97$) の化合物 **1** の阻害活性を、SERT を一過性に発現している HEK293 細胞で測定したところ、SERT による特異的な基質取り込みは SSRI であるシタロプラムと同様に **1** によって用量依存的に阻害された (図 5b)。非線形回帰分析の結果、**1** の IC_{50} 値は 6.24 nM と予測値に近似し、シタロプラムの IC_{50} 値は 2.13 nM であった。SERT に対する実測データが存在する 10,270 個の ChEMBL 化合物について **1** との構造類似性を IJC で計算したところ、最も類似した化合物 (tanimoto 係数 = 0.78; 値が大きいほど類似性が高い) の実測 pIC_{50} は 5.93 と弱いものだった。さらに、最も実測値の高い化合物 ($pIC_{50} = 11.70$; カルシウム感知受容体拮抗剤として研究されていた化合物) との構造類似度はわずか 0.26 であった。回帰モデルを構築する際に同じスキヤフォールドを有する化合物を training set、validation set および test set で同時に使用しない方が良いとの報告があるが^{28,51}、今回の結果は、GCN モデルが局所的な化学環境と化合物の活性値の関係を学習するにあたり、厳密なスキヤフォールド分布の制御は必要でない可能性も示唆するものと考えられた。

行動試験

SSRI は抗うつ薬として広く使用されているため⁵²、**1** が抗うつ薬様作用を有するかどうかをマウスで検討した。化合物 **1** (1、10 mg/kg、腹腔内投与) を投与したところ、うつ様状態を表すとされている尾懸垂試験における無動時間が有意に減少した一方、オープンフィールド試験における運動量には有意な影響が観察されなかった (図 5c、d)。これらの結果は、**1** が抗うつ薬様作用を有することを示唆している。化合物 **1** の計

算 logP 値が 2.94 であることから判断すると、**1** が中枢神経系に分布した結果、抗うつ作用を示した可能性は高いと考えられる。一方で、**1** には transient receptor potential canonical 4 ($pIC_{50} = 8.10$) および nuclear factor erythroid 2-related factor 2 ($pIC_{50} = 6.19$) に対する活性が報告されていることから、抗うつ作用はこれら 2 つ、あるいはその他の未知の標的タンパク質に作用して生じた可能性も考えられ、今後のさらなる解析が必要である。

回帰モデルの標的タンパク質網羅性の拡充

回帰モデルの性能、予測妥当性については満足のいく結果を得られた一方で、データセットサイズが小さいという理由により予測モデル構築の対象外とした標的タンパク質も存在していた。対象外となった標的タンパク質の中には、ヒスタミン H2 受容体やアドレナリン β 2 受容体など、ヒトでの薬理作用を考慮する上で重要性の高い標的タンパク質も含まれていた。そこで、回帰モデルの構築対象を拡充し、1.2 方法の節に述べた方法にてさらに 415 種類の標的タンパク質に対する回帰モデルを追加し、最終的には 542 種類のモデルを構築した。図 6 に全 542 標的タンパク質に対する MAE を示した。全般に予測性能は良好であり、395 種 (73%) において test set における MAE は 0.6 以下という精度を達成した。

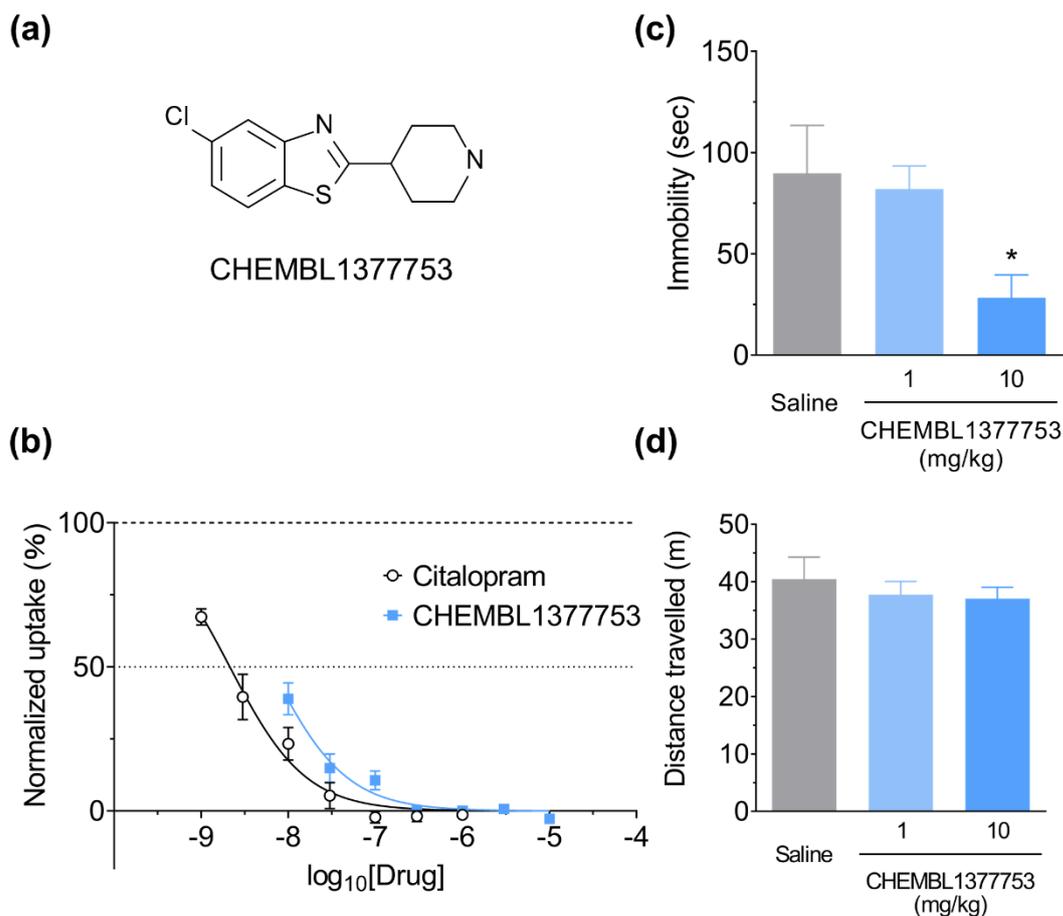


図5 セロトニントランスポーター (SERT) 予測モデルに対する実験的検証

(a) CHEMBL1377753 (1)。

(b) 化合物 1 は SERT の基質取り込みを阻害した。ヒト SERT を発現する細胞において、リガンド非存在または存在下で SERT により取り込まれる蛍光基質の量を測定した。特異的取り込み量はリガンド非存在下での値で正規化した。データは $n = 4$ の 2 つの独立した実験における平均 \pm s.e.m. を示している。

(c) - (d) 化合物 1 の投与により、マウスにおいて抗うつ薬様作用が誘発された。化合物 1 (1、10 mg/kg) を腹腔内投与した後、尾懸垂試験における無動時間 (c) またはオープンフィールド試験における移動距離 (d) を測定した。(c) 化合物 1 は無動時間を有意に減少させた。データは平均値 \pm s.e.m. を表す。One-way ANOVA の結果は $F(2,19) = 3.64$ 、 $P = 0.046$ だった。Dunnett の多重比較検定では生理食塩水群に対し $*P < 0.05$ だった。 $n = 7-8$ マウス/群 (d) 化合物 1 は移動距離に有意な影響を与えなかった。データは平均値 \pm s.e.m. を表す。One-way ANOVA では $F(2,15) = 0.41$ 、 $P = 0.67$ だった。 $n = 6$ マウス/群。

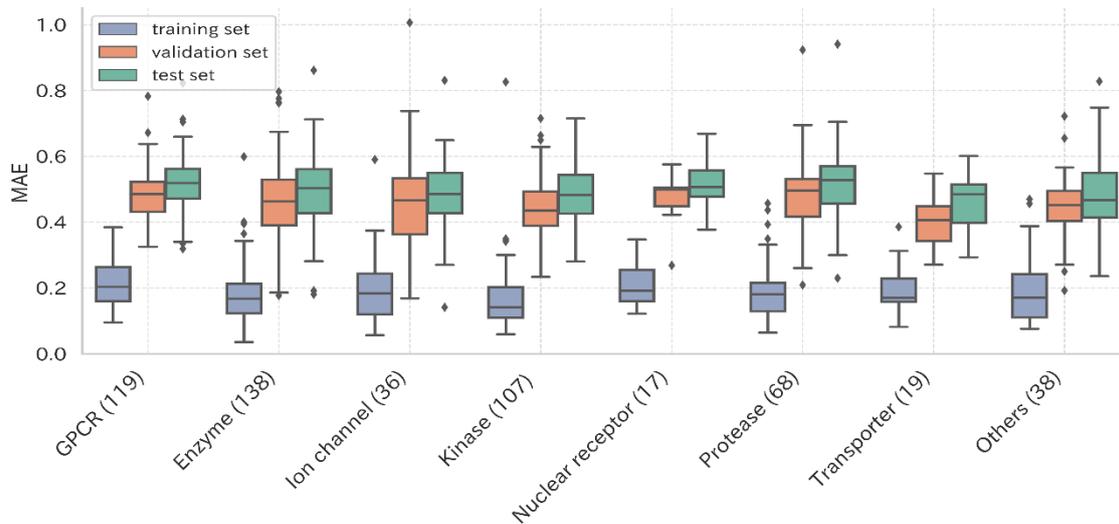


図6 542種類の標的タンパク質に対する回帰モデルのMAE

542種類の標的タンパク質に対する、タンパク質ファミリー別の予測モデルの性能。X軸の名称の後の数字は各ファミリーに含まれる標的タンパク質の数を示している。

1.4 まとめと展望

ChEMBL から取得した 542 の標的タンパク質に対して、GCN アーキテクチャを適用し、化合物の 2 次元構造情報から抽出した特徴量のみを用いて定量的な活性予測モデルを構築することに成功した。高精度での予測が可能となった理由としては、主にハイパーパラメータの探索範囲を、活性分類タスクで報告された範囲よりも拡張したことが考えられる。本研究で良好な予測性能を示した回帰モデルの多くはグラフ畳み込み層が 1 層であり、4 層のグラフ畳み込み層を持つモデルは、ハイパーパラメータ探索時に 3 層のモデルをその性能で上回るものはなかった。一般に言われる“畳み込み層が多いほど予測性能が良い”といった傾向とは異なるものであったが、これは主にプーリング層の演算による影響と推察される。アンサンブルモデルは個々のモデルに比べて予測性能を向上させたことから、個々のモデルは化合物構造の特徴を少しずつ異なった形で認識し、それらを統合することで予測精度の向上につながった可能性が示唆される。

KekuleScope データセットを用いて構築した GCN モデルの予測性能は KekuleScope (CNN) モデルと同程度で、また間接的に、ベースライン手法として比較対象にされることが多い RF モデルや FNN モデルとも同程度の予測性能を示した。KekuleScope データセットにおけるデータ準備スキームや質的測定データの取り扱いとは異なるため単純に比較することは難しいものの、興味深いことに、ChEMBL (リリース 25) を用いて構築した GCN モデルは、CNN、RF および FNN モデルよりも優れた性能を示していた。GCN モデルも CNN モデルも共に予測性能が良いことを鑑みると、化学構造にはその活性を説明するに足る十分な情報が備わっているとの仮定には一定の妥当性があると考えられる。

様々なデータソースから収集されたデータベースには、様々な実験条件下での計測値やノイズが含まれており、それらの条件を考慮することで活性予測モデルの性能が向上したとの報告もある⁵³。本研究では方法の項に記載したフィルターのみを使用したため、さらなる検討の余地があると考えられる。例えば、同じ化合物-標的タンパク質で複数の実測値がある場合、本研究では既報を参考に最大値を用いたが^{6,9,18}、中央値^{11,23,53}や平均値⁴³を用いる報告もあり、データの前処理方法の違いにより回帰モデルの予測性能がさらに向上することが期待される。

本研究では回帰モデルの構築に加えてスキュアフォールドの多様性も定量化した。その結果、スキュアフォールドの多様性は回帰モデルの性能に大きな影響を及ぼさないことを明らかにした。さらに、回帰モデルの妥当性を検証するために行ったバーチャルスクリーニングにより、既存薬であるシタロプラムと同等の SERT 阻害活性を有する新規化合物を同定するにいたった。よって、本研究で得た回帰モデルは薬剤のリポジショニング、オフターゲットの発見、医薬品開発の初期段階における探索戦略の優先順位付け、ポリファーマコロジカルな薬物の発見、生体機能の分子メカニズムの解明を支援するツール化合物の探索などに有用なヒントを提供するものと考えている。

第2章 受容体親和性予測値を用いた有害事象および原因標的タンパク質の予測

2.1 序論

有害事象（AE）とは、医薬品が投与された患者に生じたあらゆる好ましくない医療上の出来事とされている⁵⁴。AEが原因で入院あるいは入院期間の延長が引き起こされ、ときに障害や死亡にもつながることがあるなど、その社会的および経済的損失は大きい⁵⁵。さらにAEは、効果不足、商業的理由に続く、非臨床試験や臨床試験の撤退の主原因でもある⁵⁶。

予期せぬAEの発現リスクを最小化するために、前臨床段階では治療標的となる生体分子に加え、オフターゲットと呼ばれる望ましくない生体分子に対する薬理作用評価が行われている⁵⁷。しかし、全てのオフターゲットへの影響を網羅的に実測するのは時間的・費用的にも非現実的である。その結果、安全性試験を通過した医薬品であっても、市販後に想定していなかったAEが発現し、医療安全上の問題となることがある^{58,59}。そのためAEを予防する方法や、創薬の早い段階で効率的にAEを発見する手法に対する必要性は高い。

医薬品の使用によって生じたAEの情報は、医療専門家や一般使用者からの自発的な報告の形で集積され、例えば米国FDA（Food and Drug Administration）の公開するFAERS（FDA Adverse Event Reporting System）⁶⁰や、日本の医薬品医療機器総合機構（PMDA）の公開するJADER（Japanese Adverse Drug Event Report database）が存在する⁶¹。これらの自発報告データは医薬品安全管理活動に活用され、臨床試験では検出困難な稀にしか起こらないAEの特定や、多薬剤併用とAE発現上昇への影響などの解析に用いられている^{62,63}。また、薬剤併用によって単剤使用時よりもAE発現が低減する傾向に着目した結果、新規創薬標的タンパク質の提案や既存薬剤の新たな適応の提案に至るといった報告も存在する^{64,65}。このような「医療ビッグデータ」は、新たな薬理作用の洞察を与えるデータソースとしても活用されているが、対象になるのはデータベースに記録された既存の薬剤に限られるという限界が存在する。

AE自発報告データベースに加え、米国の医薬品添付文書に記載されたAE情報を構造化したデータベースであるSIDER（Side Effect Resource）もAE解析に有用なデータソースである⁶⁶。AE発症は化合物が種々の（オフ）ターゲットに結合することが原因であるとの考え方から、DrugBank⁶⁷やMatador⁶⁸といった公開データベースをもとに構築された薬剤-標的タンパク質相互作用のバイナリプロファイルを入力として、機械学習などの手法により、SIDERに記載のAEを予測する手法が報告されている^{69,70}。これらの報告で用いられたバイナリプロファイルは、標的タンパク質への作用の有無を単純化したものである。薬理作用の類似性情報を基にしてAEを予測できるため、予測適応範囲が既存の薬剤に限定されないメリットがある。しかし、標的タンパク質への定量

的な作用強度という重要な情報をそぎ落とし、設定した閾値以下（以上）の強度の薬剤を全て同列に扱うため、AE回避や薬剤リポジショニングなどの優先順位付けを行うには、バイナリではない定量的な情報による予測モデルのほうがより有用である。

Ietswaartらは、薬剤-標的タンパク質相関について前述のシンプルなバイナリ記載を一步進め、そのバイナリプロファイルを用いたAE発現確率予測モデルを発表した⁷¹。2,134種類の薬剤の184種の標的タンパク質に対する*in vitro*薬理作用を実測し、AC₅₀（IC₅₀、EC₅₀など複数種をまとめたものを指している）実測値に応じて、標的タンパク質ごとに0-3 μM、3-30 μMおよび30 μMの3クラスについて1もしくは0となるプロファイルを設定した。この半定量的な特徴量の設計はAEの発現と薬理作用の強度の関係を考慮することを可能にし、さらに、重篤な不整脈であるトルサード・ド・ポアント（TdP）の危険因子として知られるhuman ether-a-go-go（hERG）カリウムチャンネルなど、複数の関連タンパク質の特定にも成功している。

様々なデータソースから収集された薬理作用データベースには実験条件などの違いによる様々なノイズが含まれていることを考えると⁵³、実験条件のそろった実測値に基づくAE予測モデルの有用性は高い。しかし、Ietswaartらの用いた実測値は全薬剤・全標的タンパク質の組み合わせの30%程度であり、PubChemBioAssaysのような公開データベースにいたっては、薬理試験実施率は0.52%程度であるという事実を考えると（登録されている実測値の総数7,559,601を、全化合物489,575 x 試験の種類2,937で除したときの割合）⁷²、AE予測モデルの入力値として実測値以外のデータソース、例えば薬理試験の実施率100%に相当するような親和性予測モデルによる予測値を用いることの有用性は大きい。

以上のことから本章では、第1章で構築した親和性予測モデルの出力するpIC₅₀値を入力値として、AEの発現確率を出力するAE予測モデルの構築を検討した。具体的には、JADERに報告されている薬剤を対象として、各薬剤につき542種類のpIC₅₀値からなる親和性プロファイルを構築した。次にJADERのAE報告データを1、0に二値化することでAE発現確率へと変換し、親和性プロファイルからAE発現確率を予測するモデルを構築した。さらに、構築したAE予測モデルの出力結果に対して、どの標的タンパク質がどの程度影響を及ぼしたかを解析することにより、AEの発生原因となりえる標的タンパク質の抽出と解釈を試みた。なお、第1章で回帰モデルと記載したものを、AE予測モデルと区別しやすくするため本章では親和性予測モデルと記載する。

2.2 方法

データセット

PMDA のウェブサイトより JADER の 2020 年 1 月公開版（2004 年 4 月-2019 年 9 月までに PMDA が受理した症例が登録されている）を取得し解析に用いた⁶¹。JADER は患者症例一覧テーブル（demo）、医薬品情報テーブル（drug）、副作用情報テーブル（reac）および原疾患テーブル（hist）の 4 種のデータテーブルから構成されている。全てのテーブルに渡り、使用される医療用語は MedDRA（Medical Dictionary for Regulatory Activities）の基本語（PT）に従っている。MedDRA とは日米欧医薬品規制ハーモナイゼーション国際会議にて開発された医療用語集であり、表 6 にあるように、薬剤の適応症や副作用を表す用語を 5 階層の概念にて分類したものである。JADER の 2020 年 1 月公開版は MedDRA22.1 に準拠したものであったため、本研究においても 22.1 に準拠することとし、本章で用いる AE 表記は MedDRA22.1 の日本語表記（MedDRA/J）を用いた。

表 6 MedDRA の階層と分類

階層	英語表記	日本語表記
SOC	System Organ Class	器官別大分類
HLGT	High Level Group Terms	高位グループ語
HLT	High Level Terms	高位語
PT	Preferred Terms	基本語
LLT	Lowest Level Terms	下層語

JADER では、1 つの症例に複数の医薬品や有害事象のデータが存在する場合、セミコロン区切りで 1 つにまとめて登録されていることがある。そのようなデータについては「セミコロンの数 + 1」の数だけの別々のデータに分割したうえで解析に用いた。例えば症例 1 に「高血圧;転倒」と登録されていた場合、「高血圧」および「転倒」という 2 つのデータに分割した。

医薬品情報テーブルの「投与開始日」および副作用情報テーブルの「有害事象の発現日」のデータが共に年月（yyyymm）あるいは年月日（yyyymmdd）の形式に従って登録されており、かつ、「有害事象の発現日」以前に投与開始された医薬品のデータをのみを解析対象とした。日付データが年月の形式の場合には当該月の 15 日と読み替えた。日付データが 8 桁より大きい場合は 9 桁目以降を削除した。

医薬品情報テーブルの「医薬品名（一般名）」には、いわゆる一般名（成分名）だけでなく、省略した一般名（塩酸塩などの塩の記載を省略するなど）や販売名などが登録されていることがある。これらの記載のばらつきは、成分辞書データベース（メディカルデータベース社⁷³）を用いて統一した一般名にマッピングした。配合剤については、含まれる薬剤成分ごとにデータを分割し、配合剤成分データベースおよび成分／系統重複・医薬品データベース（メディカルデータベース社⁷³）を用いて同様に統一化した。例えば症例2に「アムロジピンベシル酸塩・アトルバスタチンカルシウム水和物」と登録されていた場合、「アムロジピンベシル酸塩」と「アトルバスタチンカルシウム水和物」の2つのデータに分割した。一律に統一化できなかったデータは目視確認の上、手作業で統一化した。

親和性プロファイル（特徴量）の構築

KEGG⁷⁴が提供する日本医薬品一般的名称（JAN）および PubChem CID を用いて、統一化した一般名に対する化学構造式を PubChem などのオープンデータベースより取得し、さらに IJC により、第1章の方法の節で記載した手順に従い、重原子数 1,000 以下の薬剤について canonical SMILES（以下 SMILES）に変換した。SMILES の存在しない薬剤を含む症例は解析の対象から外した。例えば症例3に「トシリズマブ（遺伝子組み換え）」と「アムロジピンベシル酸塩」が登録されている場合、重原子数が 10,208 のトシリズマブは SMILES 変換の対象外となるため（SMILES の存在しない薬剤）、症例3はアムロジピンベシル酸も含めて解析対象外とした。第1章で構築した 542 種類の標的タンパク質に対する親和性予測モデルを用い、各 SMILES につき 542 次元の pIC_{50} 予測値からなるベクトルを構築した。同一症例内で AE 発現以前に複数の薬剤が使用されていた場合は、それら薬剤の SMILES から算出されたベクトルの要素を加算し、542 次元のユニークな親和性プロファイルを構築した。加算方法は受容体占有率の考え方にに基づき設定した。濃度作用曲線を考えると例えば $pIC_{50} = 9$ の薬剤の効力は 6 の薬剤のものより 1000 倍高い。仮に両剤が同濃度で投与された場合、 $pIC_{50} = 6$ の薬剤の受容体占有率に与える影響が 10^6 とすると $pIC_{50} = 9$ のそれは 10^6 の 1000 倍（すなわち 10^9 ）である。この考え方にに基づき、 pIC_{50} 予測値を 10 のべき乗に変換してから加算し、加算後の値を 10 を底とする対数に変換することで、薬剤の組合せごとに独自の親和性プロファイルを構築した（図7）。

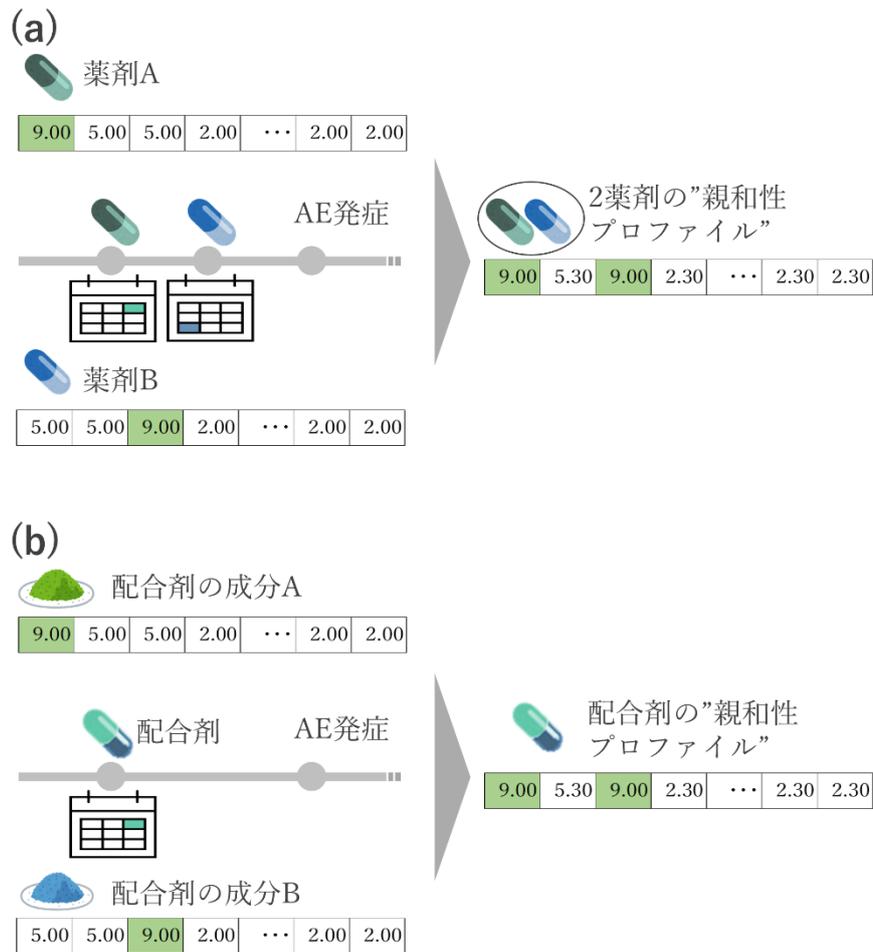


図7 親和性プロファイルの構築

(a) AE発現以前に複数の薬剤が使用されていた場合は、それぞれの親和性プロファイルの要素を10のべき乗に変換してから足し合わせ、足し合わせた値を10を底とする対数に変換することで、薬剤の組み合わせごとに1つのユニークな親和性プロファイルを構築した。

(b) AE発現以前に配合剤が使用されていた場合は、各成分薬剤の親和性プロファイルの要素を(a)と同様に足し合わせ、配合剤の親和性プロファイルを構築した。

教師ラベル（目的変数）の構築

MedDRA の標準検索式である SMQ (Standardised MedDRA Queries) 22.1 版を用い、抽出した AE 報告 (PT) を SMQ にマッピングした。SMQ とは、ある定義された医学的状态または関心のある領域に関連する MedDRA 用語 (通常は PT) をグループ化したものである。SMQ に属す PT には「狭域」と「広域」の 2 種類が存在するが、関心の対象となった状態を示す可能性が極めて高い症例の検出に用いるとされている「狭域」の PT を本研究の解析対象とした。階層構造をもつ SMQ の場合は、1 つ以上の狭域 PT が直接紐づいている SMQ を解析対象とした。例えば、うつ病および自殺/自傷 (SMQ) は 2 階層からなる SMQ だが、本研究での解析対象はレベル 2 の自殺/自傷、および、うつ病 (自殺/自傷を除く) の 2 種類とした (図 8)。JADER から取得した PT が全くマッピングされなかった SMQ を取り除いた結果、179 種の SMQ に PT がマッピングされていた。本章で扱う“AE”とは、以降特に断りのない限り、マッピング後の SMQ を指すこととする。

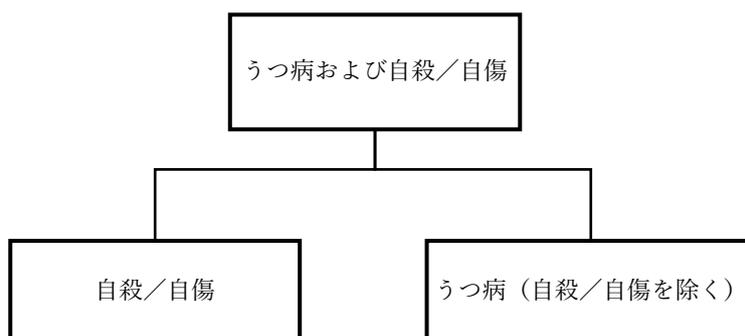


図 8 うつ病および自殺/自傷 (SMQ) の階層構造

レベル 1 のうつ病および自殺/自傷にはレベル 2 のサブ SMQ が 2 種類紐づく。レベル 2 のサブ SMQ である自殺/自傷には狭域 PT が紐づき、うつ病 (自殺/自傷を除く) には狭域と広域の PT が紐づいている。本研究ではサブ SMQ の狭域 PT に JADER より抽出した PT をマッピングし解析に用いた。

1 つの親和性プロファイルに複数の AE が報告される場合があるが、本研究ではそれらの AE は独立して発現する事象として捉え、既報を参考に AE 予測モデルの教師ラベル (目的変数; 1 と 0 に二値化したもの) を構築した⁷¹。まず親和性プロファイル別の AE 総報告回数 N_p を算出した。次に、ある親和性プロファイルに対しある AE がランダムに X 回報告されるのであれば、 X は二項分布 $B(N_p, p)$ 、 $p = 1/179$ に従うという帰無仮説を立てた (179 は AE の数)。この帰無仮説のもと、全ての AE に対する P 値を算出し、Benjamini-Hochberg の False Discovery Rate (FDR) 法により補正した。

FDR 補正後の P 値が 0.05 未満の場合に帰無仮説を棄却し、当該の親和性プロファイルにおける当該の AE の教師ラベルに 1 を設定し、それ以外は 0 とした。二項検定は python の scipy1.1.0 を、FDR 補正は statsmodels0.9.0 を使用した。

XGBoost モデル

XGBoost1.4.1⁷⁵ を用い、親和性プロファイルを入力値として、AE の発現確率を出力するモデルを AE 別に構築した。XGBoost は勾配ブースティングと呼ばれる、アンサンブル学習と決定木を組み合わせた手法で、弱学習器を直列的に連結させることで予測精度を上げるというものである。特徴量のスケールなしにモデルを構築してもモデルの予測性に影響しないアルゴリズムであるため、親和性プロファイルの値はそのまま使用した。

本研究では、AE 発症前に使用していた薬剤の組合せから得られた親和性プロファイル別に AE の教師ラベルを設定した。その中には、単に AE 実報告が存在しないため (AE 発現日や薬剤投与開始日のデータが欠損していたものを含む) 教師ラベル 0 と設定されるものが多く見受けられた。このような「みなしラベル」データは AE 予測モデル構築においてノイズとなりえることが考えられたため、予備的に教師ラベルが 1 と 0 の親和性プロファイルの間で共通薬剤が存在する場合に限り、共通薬剤を含むラベル 0 の親和性プロファイルを解析対象から外したモデル構築を試した。その結果、解析対象データ数が減少するものの、「みなしラベル」全てを用いた場合に比べ、予測性能に改善傾向が見られた。よって、実際のモデル構築時には同様の扱いを全ての AE について適応した。例えばアトルバスタチンカルシウム水和物の親和性プロファイルに対し「横紋筋融解症／ミオパチー (SMQ)」の教師ラベルが 1 と設定されたとする。もしアトルバスタチンカルシウム水和物とアムロジピンベシル酸塩 (配合剤もしくは 2 種の単剤) から構成した親和性プロファイルに対し当該 AE の教師ラベルが 0 と設定された場合、両者には共通してアトルバスタチンカルシウム水和物が存在しているおり教師ラベルが異なるため、当該 AE のモデル構築時には、後者の教師ラベル 0 の親和性プロファイルを入力データから除外した。

XGBoost のハイパーパラメータはベイズ最適化法と 5 分割クロスバリデーションを組み合わせることで最適化した。具体的には、共分散関数に Matérn カーネル、獲得関数に「期待改善度」を用い、ホールドアウトした validation set の precision-recall 曲線 (PR 曲線) 下の面積 (PR-AUC) を最大化するパラメータセットを求めた。過学習を防ぐため、各ハイパーパラメータ探索過程において、PR-AUC 値の改善が連続して 50 回観測されない場合はそこで学習を打ち切った。探索したハイパーパラメータの種類と探索範囲は表 7 に示した通りであり、表に記載のないものはデフォルト設定を用いた。ハイパーパラメータ探索には python の BayesianOptimization1.2.0 を、クロスバリデーションには scikit-learn0.24.2 の StratifiedKFold を利用した。

120回のハイパーパラメータ探索計算を実施した後、ホールドアウトした5つの validation set の平均 PR-AUC 値が最も高くなるモデルを最終モデルとした。最終モデルの予測性能は、受信者操作特性曲線（ROC 曲線）の下の面積（ROC-AUC、以降 AUC）の値にて評価した。179種類の AE のうち、教師ラベル1のデータ数が5未満である AE を除き、最終的には166種類の AE について予測モデルを構築した。

表7 XGBoost のハイパーパラメータと探索範囲

ハイパーパラメータ	探索範囲
Max_depth	[3-12.99]
Gamma	[0.01-5.0]
Eta (learning_rate)	0.05
N_estimators	50,000
Subsample	[0.5-1.0]
Colsample_bytree	[0.3-1.0]
Min_child_weight	[0-6]
Scale_pos_weight	[1.2-50]

Shapley Additive Explanations (SHAP) value

Shapley value という概念は、ゲームの最終結果に対して、プレイヤーの貢献の相対的な重要性に応じて、プレイヤー間で総利益を公平に分配することを目的として提案された。この概念の「最終結果」を予測モデルの予測値に、「プレイヤーの貢献」を入力特徴量の寄与に置き換え、予測モデルにおける特徴量の影響度を理解しようとするのが SHAP (SHapley Additive exPlanations) value である^{76,77}。

具体的には、ある特徴量*i*（標的タンパク質*i*に対する pIC₅₀ 予測値）の影響度は式 (11) の SHAP value (ϕ_i) で定義される。

$$\phi_i = \sum_{S \subseteq M \setminus \{i\}} \frac{|S|!(|M|-|S|-1)!}{|M|!} (f(S \cup \{i\}) - f(S)) \quad (11)$$

$f(S)$ は、特徴量セット S を用いて構築した AE 予測モデルにおける出力量に相当する。 M は特徴量の総数であり、 S は M から特徴量 i を除いた特徴量の組合せ、式 (11) の $f(S \cup \{i\}) - f(S)$ は特徴量 i が予測モデルに加わったことによる出力量の変化を示す。SHAP value は特徴量の影響度（大、小）に加えて方向（正の影響、負の影響）を表すことができる興味深い特徴を有する。正の影響を及ぼす特徴量は AE 発現に関与し、負の影響の特徴量は AE 発現の減少・抑制傾向に関与するとの解釈が可能であるため、本研究では、各 AE 予測モデルにおける SHAP value 絶対値の上位 20 の標的タンパク質

に着目し、その SHAP value の可視化によって AE 発現への関与が疑われる標的タンパク質を解釈することとした。SHAP value の算出および可視化には python の shap0.39.0 を使用した。

2.3 結果と考察

データセット

AE は、医薬品の使用開始後まもなく発現するものから、長期使用によって発現するものまで様々である。また、併用される医薬品によって AE の発現が変わることも報告されている^{64,65,78,79}。この事実により、本研究では AE 発現前に使用していた全ての薬剤がその AE 発現確率に関与しえると仮定し、AE 予測モデルを構築した。

AE 自発報告データベースとしては、規模の大きい FAERS が機械学習のデータソースとしてよく用いられている^{62,63}。しかし FAERS においては、1つの症例に何件 AE が報告されていようとも、ただ1つの AE 発現日が記載されているだけである。また、AE 発現日や医薬品投与開始日の報告の存在しないデータも多く、薬剤投与と AE 発現の時系列に注目する本研究の解析目的には合致しないと考えられた。そこで本研究では、日付データが FAERS よりも充実している JADER を用い、時系列が明確なデータを解析対象とした。

AE 自発報告データには、使用した医薬品の役割（被疑薬、併用薬もしくは相互作用薬）の情報が存在する。これは AE の報告者が判断した情報であり、判断基準の曖昧さや誤認などの課題を含んでいる。この課題を克服するには、1つは医療専門家による報告データのみを解析対象とする方法が考えられる⁷¹。また別のアプローチとしては、全ての医薬品を区別なく扱う方法も考えられる。後者のアプローチをとる報告としては、医薬品が原因で入院・死亡に至った事象を予測する試みにおいて被疑薬と相互作用薬を解析対象とすることで予測精度が上がったというものや⁶³、AE 発現日より遡って1か月以内に使用した全医薬品を考慮した AE 発現予測モデルというものが挙げられる⁸⁰。本研究は後者に似たアプローチをとるが、医薬品の役割や医薬品投与開始日のどちらにも制限を設けずに解析を行うこととした。

表 8 は、2.2 方法の節で述べた手順によって準備したデータのうち、医薬品の投与終了日の情報が存在するデータについて、医薬品の投与終了日から AE 発症日までの期間別、医薬品の役割別に、AE の報告件数を集計したものである。医薬品の投与終了日から1か月以内に発現した AE（868,909 件）が全体の 93%と大多数を占めている。これは Zhu らが、AE 発現日を起点に1か月以内に使用した医薬品を解析対象として選択した考えと近いものであり⁸⁰、“最近”使用した医薬品による AE が報告されやすいという現象を反映しているとも考えられた。被疑薬（411,453 件）に比べ併用薬（527,886 件）のほうが多いことから、AE 予測モデルの構築にあたり、併用薬の存在を考慮することの有用性も示唆された。興味深いことに、AE 報告者が“被疑薬”と認識した、投与終了日から5年以上経過した医薬品がわずかではあるが存在することも確認された（385 件）。

表 8 医薬品投与終了日から AE 発現日までの期間と AE 報告件数

投与終了日から				
AE 発現日までの期間	被疑薬 (件)	併用薬 (件)	相互作用薬 (件)	計 (件)
1 ヶ月以内	392,954	473,894	2,061	868,909
1 ヶ月超-6 ヶ月以内	12,338	37,418	39	49,795
6 ヶ月超-1 年以内	2,535	7,914	2	10,451
1 年超-3 年以内	2,626	6,397	25	9,048
3 年超-5 年以内	615	1,367	5	1,987
5 年超	385	896	0	1,281
計 (件)	411,453	527,886	2,132	941,471

親和性プロファイル (特徴量)

1 つの薬剤あたり 542 次元の pIC_{50} 予測値からなるベクトルを構築した。AE 発現以前に使用されていた薬剤が 1 種類のみ存在する場合にはその薬剤のベクトルを親和性プロファイルとみなした。複数存在する場合は、該当するベクトルの要素を 10 のべき乗に変換してから足し合わせ、足し合わせた値を 10 を底とする対数に変換することで、薬剤の組み合わせごとに 542 次元のユニークな親和性プロファイルを構築した (図 7)。その結果、ユニークな親和性プロファイルは 48,606 種類となった。構築した親和性プロファイルの AE 報告件数トップ 5、およびそのプロファイルを構成する薬剤名を表 9 に示した。最も多かった親和性プロファイルは造影剤のイオパミドール単独によるものだった。3 番目は FOLFOX 療法と呼ばれる 3 剤併用の化学療法薬剤のプロファイルだった。このような複合型のプロファイルは、1 種の薬剤-1 種の AE の関係を解析する既存の手法^{71,81}においては用いることのできない入力データの形であり、本研究の特徴の 1 つでもある。

表9 親和性プロファイル AE 報告件数トップ5

薬剤名	説明	件数
イオパミドール	非イオン性尿路・血管造影剤	4,278
イオメプロール	非イオン性造影剤	3,136
フルオロウラシル_レボホリナートカルシウム_オキサリプラチン ^a	FOLFOX 療法と呼ばれる 3 剤併用の化学療法に用いられる薬剤	2,845
イオヘキソール	非イオン性造影剤	2,537
ソラフェニブトシル酸	抗悪性腫瘍剤	2,086

(a) 複数の薬剤から成る親和性プロファイルにおいては、薬剤を”_”で結合して示している。

AE 予測モデル

AE の発現確率を出力するモデル 166 種類を構築し、AUC を指標として AE 予測モデルの予測性能を評価した。図 9a、b はホールドアウトした validation set の AUC を示したものである。AUC は[0-1]の値をとる指標であり、AE 予測モデルの予測性能がランダムなそれと変わらないのであれば 0.5、1 に近いほど予測性能がよいことを示す。本研究では 0.7 以上を良好な予測性能の目安として考えた。

本研究に用いたデータセットにおける教師ラベルデータは不均衡なものであり、教師ラベル 1、すなわち AE 発症ありの親和性プロファイル数が 5-10,477 件であるのに対し、教師ラベル 0 のそれは 1,465-47,767 件であった。このような不均衡さは予測精度に影響を与えることが考えられたため、AUC と教師ラベル 1 の数をプロットしたところ（図 9a）、教師ラベル 1 の数の増加とともに AUC も上昇することがわかった（表 10）。なお予備検討として、ラベル 1 と 0 の不均衡さに対処するための XGBoost のパラメータ（scale_pos_weight）の上限の引き上げ（上限値として 100、300、1000 を使用）、ラベル 1 データの人工的なオーバーサンプリング（SMOTE 法⁸²）およびラベル 0 データのアンダーサンプリング（Tomek’s Links 法⁸³）を試したが、AUC は同程度もしくは悪化傾向であったため、本研究で扱ったデータセットについては、2.2 方法の節に示した計算条件が予測性能の観点からも計算時間の観点からも妥当なものであると推定された。

図 9b は、教師ラベル 1 のデータが 100 件以上存在する AE98 種について、AUC を降順ソートしプロットしたものである。72 種（74%）の AE において AUC は 0.7 以上と良好な予測精度が得られていた。AUC の最高値は過敏症（SMQ）の 0.98 で、0.7 未満の AE の中には先天性、家族性および遺伝性障害（SMQ）（0.60）のように使用医薬品から予測することが非常に困難だと思われるものが含まれていた。

本研究では、AE 発現日以前に使用していた全ての薬剤がその AE 発現確率に影響しえると仮定し、親和性プロファイルを設計した。ゆえに1つの親和性プロファイルは1つの薬剤から構成されることもあり、10種以上の薬剤から構成されることもある。多剤併用とAE発現リスク上昇の関係が報告され臨床上的の問題になっているが⁷⁸、調査した限りにおいては6剤以上の併用を露わに扱うAE予測モデルの報告は数少ない。図9cは親和性プロファイルを構成する薬剤数別にAUCをプロットしたものである。1種の薬剤から構成される親和性プロファイル(1 drug; AUC 平均値 0.69)を除き、全ての構成においてAUC平均値は0.7を超えていた。構成薬剤数別の教師ラベル1のデータ数(平均値)をみると、9薬剤以上の構成(37件)に次いで1薬剤の構成(45件)が少ない。これは、AEが1薬剤だけの使用によって発現するというよりは、2薬剤以上の使用によって起こりやすくなる、あるいは2薬剤以上を使用する必要のある病態において起こりやすくなることを示唆するものと考えられる。10薬剤以上から構成される親和性プロファイルでの予測性能が良い理由としては、化学療法などのあらかじめ決められた薬剤の組み合わせがこの構成の中に多く存在し、親和性プロファイルのバリエーションとAEの組み合わせが限定された結果、予測しやすいデータとなっている可能性が考えられた。

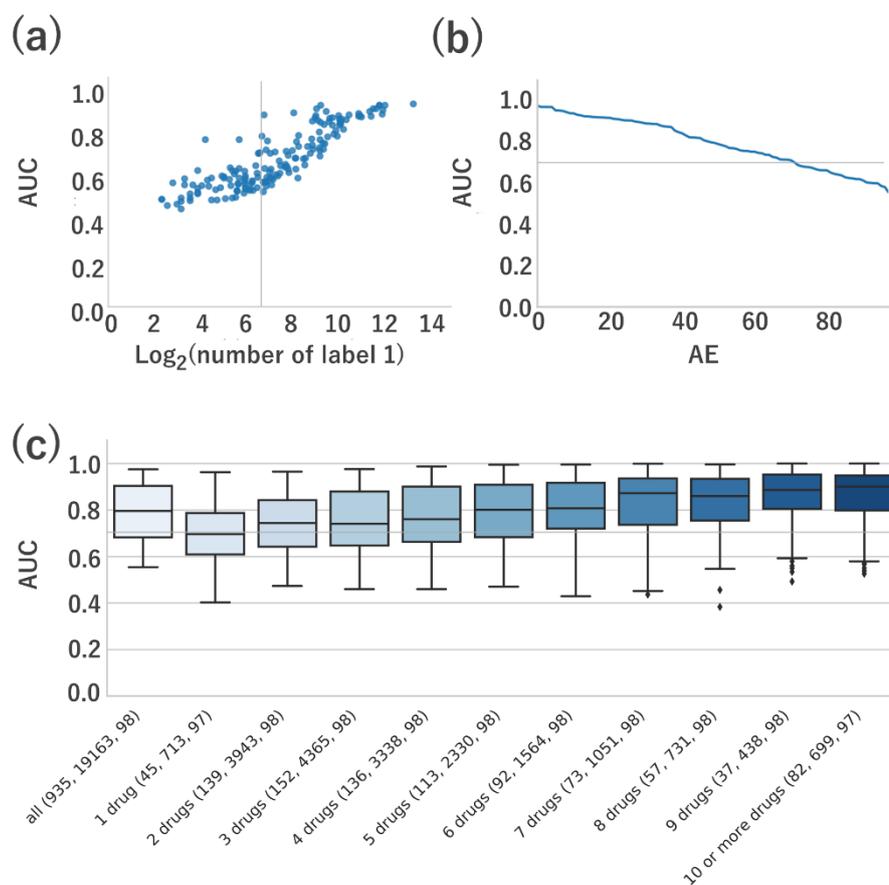


図9 AE予測モデルの予測性能

(a) AUCと教師ラベル1のデータ数(2を底とする対数)の散布図。教師ラベル1のデータが増えるほど予測性能が向上する傾向がある。垂直線は6.64($\log_2(100)$)を示す。

(b) 教師ラベル1のデータ数が100件以上のAE予測モデルのAUC。横軸は予測モデルの数を示す。水平線はAUC0.7を示す。

(c) 親和性プロファイルを構成する薬剤数別のAUCのボックスプロット。X軸の各名称の後の数字は左から順番に教師ラベル1の平均データ数、教師ラベル0の平均データ数および当該の構成数からなる親和性プロファイルが存在する予測モデル(AE)の数を示す。

表 10 AE 予測モデルの詳細

AE 名称	ラベル 1 ^a (件数)	ラベル 0 ^b (件数)	AUC
うつ病 (自殺/自傷を除く) (SMQ)	157	27,632	0.74
その他の虚血性心疾患 (SMQ)	294	23,605	0.71
アカシジア (SMQ)	43	44,001	0.66
アナフィラキシー反応 (SMQ)	3,755	5,192	0.97
アナフィラキシー/アナフィラキシー様ショック状態 (SMQ)	3,750	5,248	0.97
ギラン・バレー症候群 (SMQ)	15	47,111	0.55
ショック関連の循環あるいは心臓の状態 (トルサード ド ポアントを除く) (SMQ)	1,144	12,290	0.89
ジスキネジア (SMQ)	139	28,890	0.66
ジストニア (SMQ)	105	31,467	0.82
トルサード ド ポアント、ショック関連状態 (SMQ)	525	18,292	0.91
トルサード ド ポアント/Q T延長 (SMQ)	490	18,810	0.89
パーキンソン様事象 (SMQ)	114	31,102	0.67
リポジストロフィー (SMQ)	19	46,081	0.80
上室性頻脈性不整脈 (SMQ)	295	23,583	0.66
不整脈に関連する臨床検査、徴候および症状 (SMQ)	9	47,685	0.51
中枢神経系出血および脳血管性損傷に伴う状態 (SMQ)	33	45,813	0.62
中毒性-敗血症性ショック状態 (SMQ)	711	15,066	0.78
乳酸アシドーシス (SMQ)	120	30,250	0.76
事故および損傷 (SMQ)	607	16,709	0.84
伝導障害 (SMQ)	490	18,798	0.78
低カリウム血症 (SMQ)	286	23,578	0.71
低ナトリウム血症/S I A D H (SMQ)	630	16,558	0.89
低血糖性および神経性ショック症状 (SMQ)	607	16,723	0.77
低血糖 (SMQ)	734	15,056	0.94
偽膜性大腸炎 (SMQ)	169	27,096	0.63
先天性、家族性および遺伝性障害 (SMQ)	122	30,239	0.60
先天性および新生児不整脈 (SMQ)	6	47,751	0.49
免疫処置後の全身痙攣発作 (SMQ)	1,126	12,334	0.89
全身性エリテマトーデス (SMQ)	26	45,999	0.56
全身状態およびその他特定部位の前癌状態 (SMQ)	5	47,767	0.52
出血性あるいは虚血性と特定されない中枢神経系血管障害 (SMQ)	33	45,812	0.63
出血性中枢神経系血管障害 (SMQ)	1,078	12,494	0.92
出血関連用語 (臨床検査用語を除く) (SMQ)	4,178	4,443	0.94

(続き)

前庭障害 (SMQ)	49	40,202	0.56
効能／効果の欠如 (SMQ)	27	45,937	0.51
動脈の塞栓および血栓 (SMQ)	681	16,088	0.76
口腔咽頭のアレルギー状態 (SMQ)	132	29,075	0.65
口腔咽頭の感染 (SMQ)	93	32,578	0.62
口腔咽頭の新生物 (SMQ)	9	47,685	0.53
口腔咽頭の状態 (新生物、感染およびアレルギーを除く) (SMQ)	2,001	8,384	0.92
味覚および嗅覚障害 (SMQ)	110	31,396	0.60
呼吸不全 (SMQ)	618	16,600	0.80
喘息／気管支痙攣 (SMQ)	214	25,128	0.69
好酸球増加および全身症状を伴う薬物反応症候群 (SMQ)	542	18,061	0.92
好酸球性肺炎 (SMQ)	235	24,946	0.73
妊娠、分娩合併症と危険因子 (流産および死産を除く) (SMQ)	95	32,445	0.74
妊娠中絶および流産のリスク (SMQ)	87	35,431	0.68
尿細管間質性疾患 (SMQ)	233	24,965	0.66
強膜障害 (SMQ)	10	47,678	0.58
後腹膜線維症 (SMQ)	47	43,682	0.53
心不全 (SMQ)	1,095	12,401	0.87
心室性頻脈性不整脈 (SMQ)	480	19,121	0.81
心筋梗塞 (SMQ)	368	21,856	0.71
心筋症 (SMQ)	80	35,459	0.56
急性中枢性呼吸抑制 (SMQ)	608	16,675	0.72
急性腎不全 (SMQ)	2,065	8,081	0.92
急性膵炎 (SMQ)	313	22,882	0.62
悪性リンパ腫 (SMQ)	124	30,260	0.80
悪性乳房腫瘍 (SMQ)	62	39,489	0.65
悪性前立腺腫瘍 (SMQ)	30	45,889	0.57
悪性卵巣腫瘍 (SMQ)	12	47,670	0.51
悪性子宮／卵管腫瘍 (SMQ)	40	44,704	0.67
悪性度不明の血液学的腫瘍 (SMQ)	53	40,181	0.80
悪性度不明の非血液学的腫瘍 (SMQ)	64	37,151	0.63
悪性疾患関連の治療と診断の手法 (SMQ)	9	47,686	0.47
悪性疾患関連状態 (SMQ)	240	24,734	0.64
悪性症候群 (SMQ)	624	16,614	0.97
悪性皮膚腫瘍 (SMQ)	31	45,883	0.63
悪性肝臓腫瘍 (SMQ)	53	40,164	0.67
悪性胆道腫瘍 (SMQ)	15	47,109	0.54

(続き)

感染性肺炎 (SMQ)	1,349	11,132	0.90
感染性胆道系障害 (SMQ)	146	28,842	0.61
慢性腎臓病 (SMQ)	320	22,852	0.77
投薬過誤 (SMQ)	7	47,706	0.60
敗血症 (SMQ)	494	18,683	0.76
敵意／攻撃性 (SMQ)	64	37,137	0.62
新生児障害 (SMQ)	135	29,052	0.74
末梢性ニューロパチー (SMQ)	186	25,859	0.68
横紋筋融解症／ミオパチー (SMQ)	1,180	12,061	0.90
歯肉障害 (SMQ)	43	44,001	0.52
水晶体障害 (SMQ)	51	40,203	0.59
洞結節機能障害 (SMQ)	177	26,036	0.70
消化管の出血 (SMQ)	1,469	10,616	0.92
消化管の潰瘍 (SMQ)	1,093	12,530	0.90
消化管の穿孔、潰瘍、出血あるいは閉塞についての非特異的所見／ 処置 (SMQ)	49	40,207	0.56
消化管の穿孔 (SMQ)	317	22,903	0.82
消化管の閉塞 (SMQ)	397	21,298	0.71
消化管の非特異的機能障害 (SMQ)	35	45,808	0.52
消化管の非特異的炎症 (SMQ)	111	31,135	0.60
消化管の非特異的症状および処置 (SMQ)	2,148	7,534	0.91
涙器障害 (SMQ)	41	44,203	0.58
溶血性障害 (SMQ)	181	25,897	0.64
無顆粒球症 (SMQ)	1,986	8,354	0.93
生殖器の前癌状態 (SMQ)	10	47,678	0.62
生殖能障害 (SMQ)	12	47,670	0.52
甲状腺機能亢進症 (SMQ)	52	40,184	0.59
甲状腺機能低下症 (SMQ)	53	40,174	0.62
痙攣 (SMQ)	1,162	12,043	0.88
眼感染 (SMQ)	95	32,447	0.74
眼球運動障害 (SMQ)	64	37,141	0.57
眼窩周囲および眼瞼障害 (SMQ)	120	30,246	0.58
精神病および精神病性障害 (SMQ)	417	20,992	0.77
結膜障害 (SMQ)	79	35,896	0.60
網膜障害 (SMQ)	224	25,014	0.68
緑内障 (SMQ)	91	34,891	0.64
聴覚障害 (SMQ)	141	28,874	0.62

(続き)

肝不全、肝線維症、肝硬変およびその他の肝細胞障害 (SMQ)	2,908	6,183	0.92
肝感染 (SMQ)	136	28,983	0.62
肝臓に起因する胆汁うっ滞および黄疸 (SMQ)	990	13,215	0.85
肝臓に関連する凝固および出血障害 (SMQ)	98	32,178	0.74
肝臓関連臨床検査、徴候および症状 (SMQ)	4,075	4,457	0.92
肺高血圧症 (SMQ)	75	35,955	0.59
胃腸の前癌状態 (SMQ)	50	40,205	0.56
胆嚢関連障害 (SMQ)	141	28,887	0.59
胆石に関連する障害 (SMQ)	72	36,304	0.56
胆道障害 (SMQ)	618	16,650	0.75
胎児障害 (SMQ)	47	43,683	0.65
脂質異常症 (SMQ)	159	27,603	0.64
脱水 (SMQ)	221	25,057	0.75
脱髄 (SMQ)	123	30,239	0.62
腎血管障害 (SMQ)	12	47,670	0.54
腫瘍マーカー (SMQ)	17	46,999	0.58
腫瘍崩壊症候群 (SMQ)	74	35,968	0.57
腱障害および靭帯障害 (SMQ)	21	46,063	0.58
自殺/自傷 (SMQ)	548	18,039	0.95
良性肝新生物 (嚢胞およびポリープを含む) (SMQ)	8	47,703	0.49
薬剤離脱 (SMQ)	46	43,687	0.64
薬物乱用および依存 (SMQ)	39	44,798	0.61
虚血性中枢神経系血管障害 (SMQ)	892	14,012	0.87
虚血性大腸炎 (SMQ)	299	23,529	0.79
蛋白尿 (SMQ)	63	37,156	0.59
血栓性静脈炎 (SMQ)	23	46,054	0.62
血液の前癌状態 (SMQ)	79	35,895	0.61
血液学的悪性腫瘍 (SMQ)	264	24,206	0.75
血液量減少性ショック状態 (SMQ)	670	16,169	0.79
血管タイプ不明あるいは混合型の塞栓および血栓 (SMQ)	1,248	11,698	0.85
血管外漏出 (注射、注入および埋込み部位) (SMQ)	15	47,109	0.66
血管浮腫 (SMQ)	934	13,538	0.83
血管炎 (SMQ)	119	30,242	0.60
血行動態的浮腫、蓄水および体液過負荷 (SMQ)	1,029	12,988	0.82
視神経障害 (SMQ)	23	46,055	0.52
角膜障害 (SMQ)	112	31,123	0.92
認知症 (SMQ)	61	39,483	0.55

(続き)

造血障害による白血球減少症 (SMQ)	4,381	4,346	0.97
造血障害による血小板減少症 (SMQ)	2,760	6,624	0.94
造血障害による赤血球減少症 (SMQ)	210	25,204	0.67
造血障害による2種以上の血球減少症 (SMQ)	766	15,075	0.82
過敏症 (SMQ)	10,477	1,465	0.98
重症皮膚副作用 (SMQ)	3,115	6,080	0.95
間質性肺疾患 (SMQ)	3,339	5,781	0.95
関節炎 (SMQ)	110	31,393	0.55
静脈の塞栓および血栓 (SMQ)	582	17,336	0.90
非感染性下痢 (SMQ)	894	14,166	0.91
非感染性肝炎 (SMQ)	458	19,347	0.74
非感染性脳炎 (SMQ)	13	47,199	0.59
非感染性脳症／譫妄 (SMQ)	702	15,085	0.88
非感染性髄膜炎 (SMQ)	64	37,143	0.57
非特異的不整脈用語 (SMQ)	91	34,898	0.57
非特異的徐脈性不整脈用語 (SMQ)	5	47,767	0.52
非特異的頻脈性不整脈用語 (SMQ)	20	46,080	0.61
非血液学的悪性腫瘍 (SMQ)	800	14,819	0.89
骨壊死 (SMQ)	278	23,893	0.93
骨粗鬆症／骨減少症 (SMQ)	15	47,110	0.57
骨髄異形成症候群 (SMQ)	63	37,149	0.60
高血圧 (SMQ)	383	21,619	0.68
高血糖／糖尿病の発症 (SMQ)	815	14,233	0.90

(a) 二項検定にてAE発現が有意に認められた親和性プロファイルの数

(b) AE発現が報告されているが二項検定にて有意ではなかった、もしくはAE発現報告のない親和性プロファイルの数

SHAP value による AE 発現関与タンパク質の推定

AE の発現あるいは抑制に関与する標的タンパク質、およびその影響度を定量的に検討するため、構築した各 AE 予測モデルについて SHAP value を算出した。SHAP value は特徴量の予測結果への影響度を表現することができるため、SHAP value 絶対値の大きい上位 20 の標的タンパク質を抽出し解析対象とした。図 10a-c は各 AE 予測に対する標的タンパク質の SHAP value をヒートマップで表したものである。なおヒートマップによる可視化においては、教師ラベル 1 のデータ数が 100 未満の AE も含め、2 種類以上の AE にて抽出された標的タンパク質を対象とした。

■循環器系 AE ヒートマップ

図 10a には循環器系の AE にて抽出された標的タンパク質を示した。致死性の多形性心室頻拍である TdP や心電図 QT 延長への関与が知られている hERG カリウムチャンネルを本 AE 予測モデルにおいても抽出していることがわかった^{84,85}。心機能障害への関与が知られているアセチルコリン受容体⁸⁶、一酸化窒素合成酵素 (NOS)⁸⁷なども抽出されていた。

1990 年代初頭に抗ヒスタミン薬のテルフェナジンによる TdP^{88,89}が報告されて以来、心電図 QT 間隔の延長を伴う TdP による医薬品の市場撤退の原因の 1 つとして hERG 阻害作用が問題視されてきた。現在、新薬の開発において hERG 阻害作用のスクリーニングはルーチンで行うべきものとされている⁵⁷。

■精神神経系 AE ヒートマップ

図 10b には精神神経系の AE にて抽出された標的タンパク質を示した。ドパミン D2、D3 および D4 受容体など、ジスキネジア・ジストニア・パーキンソン様事象への関与が知られている標的タンパク質が抽出されていることがわかった。これらの AE ではセロトニン受容体も抽出されており、L-ドパ誘発のジスキネジアへのセロトニンの関与などの報告などからみても妥当性があると考えられた⁹⁰。また、呼吸抑制や術後せん妄を引き起こすことが問題となる、モルヒネなどのオピオイドの主作用点である μ オピオイド受容体も抽出されていた⁹¹。興味深いことに、薬物乱用および薬物依存 (SMQ) および薬剤離脱 (SMQ) において、コカインやアンフェタミンの標的タンパク質であるドパミントランスポータ (DAT) が抽出されていた⁹²。我が国ではコカインやアンフェタミンは処方医薬品としては存在せず、アンフェタミン類縁体が検査薬として使用されているのみである。そのため、JADER にはコカインによる AE 報告は存在せず、アンフェタミンによる薬物依存の報告も存在しない。一方で、注意欠陥・多動性障害 (ADHD) やナルコレプシーの治療薬として用いられるものの、その依存性が問題となっているメチルフェニデートについて、JADER 中で薬物依存の報告がなされてい

る。メチルフェニデートは DAT に作用することから、本 AE 予測モデルで DAT が抽出されたのはメチルフェニデートによる薬物依存の報告に起因するものと推察された。

■出血、血栓系 AE ヒートマップ

図 10c には出血・血栓系の AE にて抽出された標的タンパク質を示した。血液凝固因子阻害剤、シクロオキシゲナーゼ (COX) -1 阻害剤などの抗血小板薬や、組織型プラスミノゲンアクチベータ、一酸化窒素などの抗血栓作用物質といった、出血関連事象への関与が知られている薬剤の標的タンパク質が抽出されていた^{93,94}。AE 予測モデルにおいては COX-1 のみが抽出され、COX-2 は抽出されなかったが、COX-2 選択的な阻害剤は非選択的な阻害剤に比べ出血事象を起こしにくいと言われている報告と矛盾しない結果である⁹⁴。

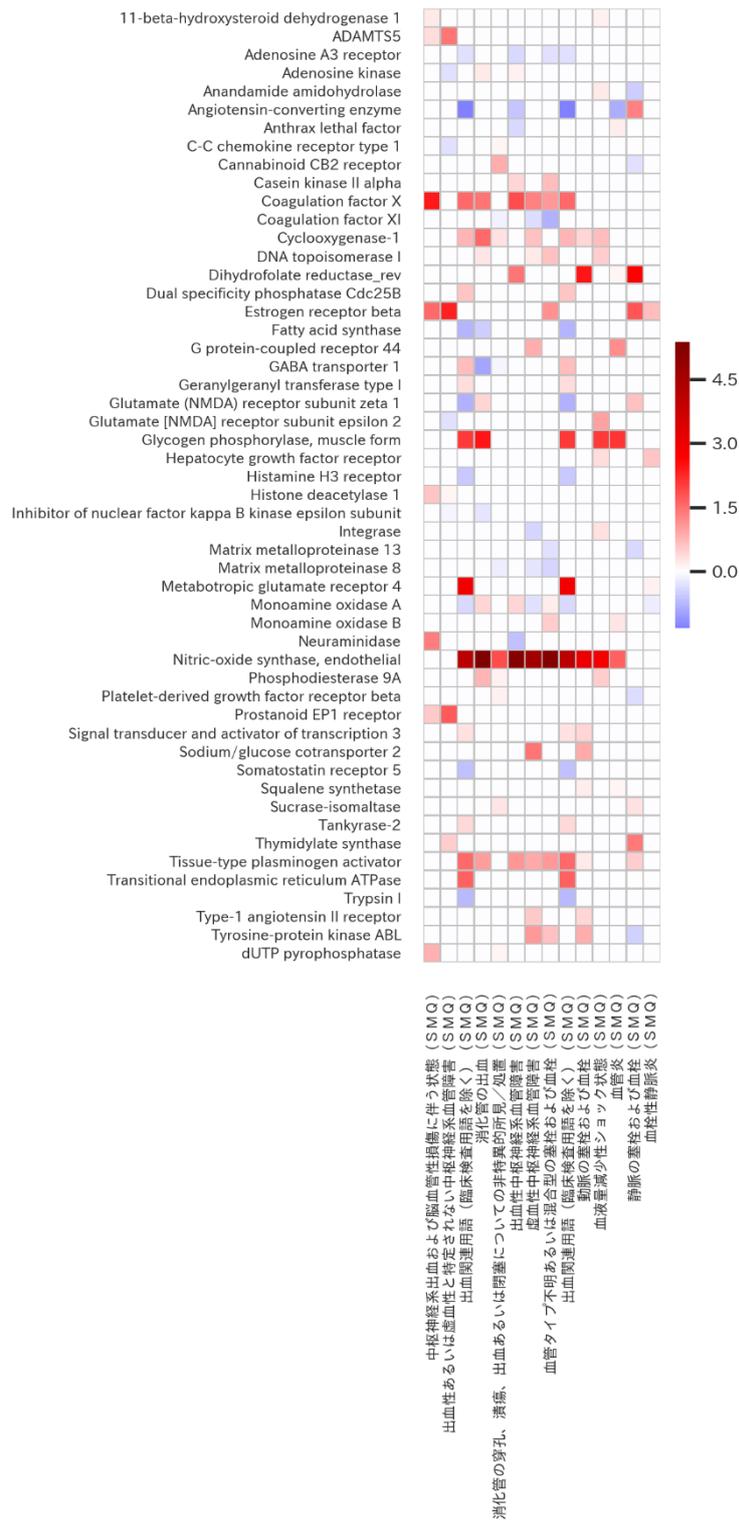


図 10 (c) 出血、血栓系の AE と抽出された標的タンパク質

構築した AE 予測モデルにより既知の標的タンパク質が認識されていることが確認されたため、次のステップとして、各 AE 予測モデルにおいて抽出された標的タンパク質、SHAP value とその正負の方向（すなわち AE 発現への正方向、負方向への影響度）、および当該標的タンパク質との親和性との関係を検証した。具体的には、SHAP value によって可視化された標的タンパク質と AE 発現の関係性が、実験動物あるいはヒトにおける既存の報告において直接的あるいは間接的に示されているかを調査した。既報が存在する場合には関係性に妥当性ありと判断することとした。

抽出された標的タンパク質のうち SHAP value の大きいものを中心に調査した結果、教師ラベル 1 のデータ数が 100 件以上の AE98 種類のうち、少なくとも 1 種類以上の標的タンパク質において AE 発現との関連性を示唆する報告の存在した AE は 69 種類（70%）であった。この結果も、本研究で構築した AE 予測モデルの予測妥当性を示すものと考えている。以下に、多くの AE 予測モデルで取り扱われる TdP について詳述し、次に精神神経系 AE の代表的なものについて詳述する。

■トルサード ポイント (TdP) /QT 延長 (SMQ)

TdP は突然死の原因にもなる重篤な不整脈であり、心電図 QT 延長は TdP を誘発する因子の 1 つである。図 11a 左は TdP/QT 延長 (SMQ) 予測モデルの出力結果への寄与度の高い 20 種類の標的タンパク質、およびその SHAP value の関係を、各標的タンパク質に対する親和性予測値に応じて色を変え示したものである。1 つの点は 1 つの親和性プロファイルを意味している。

薬剤誘発性の TdP の原因の 1 つに hERG カリウムチャネルの障害が関与することが知られている^{84,85}。hERG 遺伝子に変異の生じた先天性 QT 延長症候群の 2 型ではストレスまたは安静時でも失神・突然死が発生する可能性がある⁹⁵。本 AE 予測モデルにおいても TdP/QT 延長の発現に正に寄与する因子として hERG カリウムチャネルが抽出されてきた。図 11a 右は hERG に対する親和性予測値と SHAP value の関係である。この図からは、hERG に対する pIC₅₀ が 7 超の領域で TdP/QT 延長により寄与する可能性が示唆された。

この AE では内皮型一酸化窒素合成酵素 (eNOS) も AE 寄与因子として抽出されてきた。eNOS は心筋細胞膜上に発現しており、そのノックアウトマウスでは高血圧および心収縮異常を呈する^{87,96}。調べた限りにおいて TdP/QT 延長と eNOS との関係を論じた予測モデルは本報告が初めてだが、ノックアウトマウスの表現系を鑑みると、eNOS が TdP/QT 延長の発症に直接あるいは間接的に関与している可能性は高いと思われた。図 11a 右の図からは eNOS に対する pIC₅₀ が 6.5 近辺で SHAP value が最大になることがわかる。eNOS の親和性予測モデルの予測性能 MAE は 0.46 ± 0.034 (std) であり、その予測誤差を加味する必要があるものの、6.5 という pIC₅₀ は古典的な eNOS 阻害剤の N^G-ニトロアルギニン (IC₅₀ = 1.2 μM) と同程度の活性強度であり⁹⁷、

この活性強度範囲の阻害であっても TdP/QT 延長への影響を注意する必要性を示唆するものと考えられた。

■自殺／自傷（SMQ）

死の思考や自殺念慮は、うつ病患者に多く見られる症状であり、インターフェロン製剤、副腎皮質ステロイド剤、経口避妊薬、抗ヒスタミン薬などによる薬剤惹起性うつ病も知られている⁹⁸。

図 11b 左に示された標的タンパク質の中には、抗うつ薬の作用点である SERT、うつ病患者の不安症状を抑える目的で一緒に処方されることの多い抗不安薬の作用点である γ -アミノ酪酸（GABA）の受容体の 1 つ、GABA-A 受容体が存在する。表面上は治療標的タンパク質が抽出されたように見えるが、抗うつ薬と抗不安薬の併用使用により自殺、自傷のリスクが増加したとの調査結果が存在することより⁹⁹、これらの標的タンパク質が抽出されたことは疫学調査と矛盾しないと考えられた。

興味深いことに、本 AE では SHAP value が負の（すなわち AE 発現の抑制に寄与する）ヒストン脱アセチル化酵素 6（HDAC6）や血管内皮増殖因子受容体 2（VEGFR2）が抽出されてきた。図 11b 右に示すように、 pIC_{50} がそれぞれ 5 超、6 超の活性強度において AE 発現が抑制されていた。既報において、HDAC6 阻害剤はマウスで抗うつ効果を示すことや¹⁰⁰、VEGFR2 ノックアウトマウスでは恐怖条件付け試験ですくみ行動が減少することが示されている¹⁰¹。これらの報告を併せて考えると、作用メカニズムは異なるものの、HDAC6 および VEGFR2 がヒトの自殺／自傷（SMQ）関連の治療標的になりえる可能性が示唆され、本 AE 予測モデルが AE 発現原因の推定のみならず、新規治療標的のヒントを与えるツールとして機能する可能性を示す例だと考えられる。

■うつ病（自殺／自傷を除く）（SMQ）

この AE は上述の自殺／自傷（SMQ）とも関連性がある。図 11c 左には自殺／自傷（SMQ）と同様 SERT や GABA-A 受容体が抽出されており、セロトニン 2A（5-HT_{2A}）受容体も両 AE に共通して抽出されていた。5-HT_{2A} は催幻覚薬の作用点であるが、アゴニストには抗うつ効果が認められていることから¹⁰²、AE 予測モデルの抽出結果と当該報告とは矛盾しないと考えられる。また中枢神経系に高発現のニコチン性アセチルコリン受容体 $\alpha 4\beta 2$ の阻害薬は、抗うつ薬ケタミンのセロトニン神経活性化作用を消失させるが¹⁰³、本標的タンパク質も AE 予測モデルで抽出されており、予測モデルの妥当性が示唆される。

■精神病および精神病性障害（SMQ）

精神病とは妄想・幻覚・著しい思考散乱性言語・興奮行動が認められる精神異常と定義され、統合失調症など様々な精神疾患やコカインなどの違法薬物およびドパミン作動薬などの治療薬により惹起される¹⁰⁴。

図 11d 左に抽出された標的タンパク質の中には、抗精神病薬の作用点であるドパミン D2、D4 受容体やセロトニン受容体が存在していたが、この理由としては、抗精神病治療薬を減量・中断した際の原因の悪化などが AE 報告に含まれていることが想像された。本研究では解析対象として医療専門家が“被疑薬”として報告していない医薬品の情報も“被疑薬”と区別せずに用いている。そのため、例えば抗精神病治療薬（“併用薬”として報告）を使用中の患者において、原疾患の症状を悪化させるような薬剤（“被疑薬”）を追加投与した結果生じた AE 報告を、AE 予測モデルが認識した可能性も推測される。

本 AE においても、AE 発現を抑制する傾向の標的タンパク質が抽出されてきており、その 1 つがプロスタグランジン E 合成酵素である。プロスタグランジン E の受容体である EP1 受容体のノックアウトマウスにおいては、コカイン誘発運動量亢進が減弱し、ドパミン D2 受容体アンタゴニストによるカタレプシーも減弱するという報告が存在する¹⁰⁵。もう 1 つはインテグリン $\alpha\beta3$ である。インテグリン $\beta3$ は SERT と結合することが知られている。また、ドパミン神経伝達異常がその要因として知られる ADHD¹⁰⁶ にインテグリン $\beta3$ の L33P の遺伝子変異が関連している¹⁰⁷。さらにインテグリン $\alpha\beta3$ の機能阻害は、抗うつ薬のシタロプラムの尾懸垂試験における無動時間減少効果（抗うつ効果）を阻害することから¹⁰⁸、インテグリン $\alpha\beta3$ に対する作用がセロトニンおよびドパミン神経伝達に影響を及ぼし、最終的に精神症状の発症を予防しえる可能性が考えられる。

■認知症（SMQ）

薬剤性の認知症の報告数は多くなく、教師ラベル 1 の数が 100 件未満であるが、図 11e に示したように、AE 発現に正に寄与する標的タンパク質としてグレリン受容体および transient receptor potential cation channel subfamily V member 4 (TRPV4) が得られていた。グレリン O-アシルトランスフェラーゼ欠損マウスが海馬依存性空間記憶障害を示すという報告や¹⁰⁹、TRPV4 ノックアウトマウスにおいて認知機能障害が起こるとの報告から¹¹⁰、ヒトでの因果関係については推測の域をでないものの、一定程度の妥当性のある標的タンパク質を抽出しているように思われた。TRPV4 についてはいまだに強力な阻害剤の報告は数少ないが¹¹¹、今後このような化合物が開発されることで AE 予測モデルが抽出した標的タンパク質と AE の関係の解明も進むものと期待される。

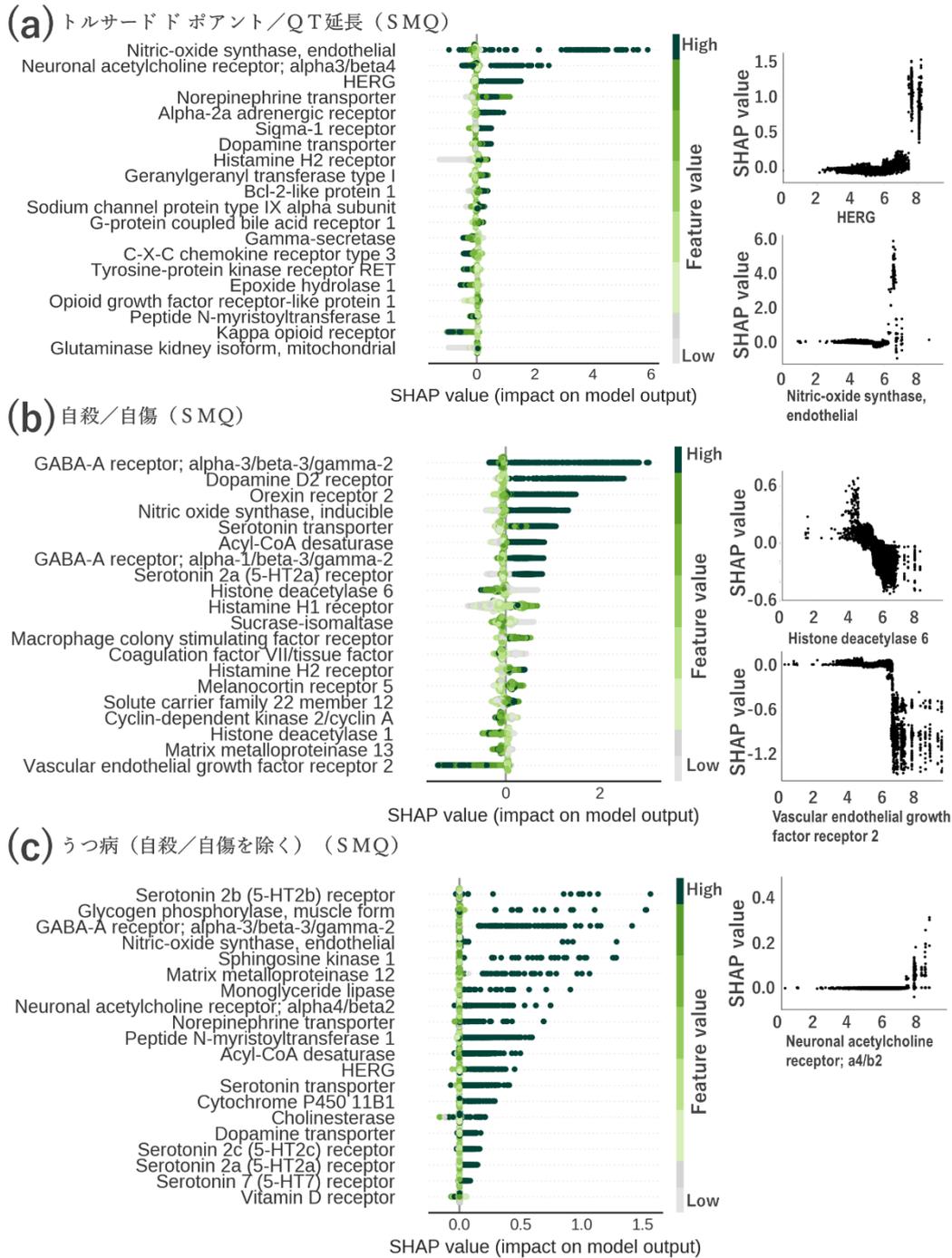


図 11 (a) - (c) AE と抽出された標的タンパク質

予測モデル出力結果に対する寄与度の高い 20 種類の標的タンパク質と SHAP value。1 つの点は 1 つの親和性プロファイルを示す。

(左) 標的タンパク質に対する親和性予測値に応じて濃い緑から灰色に変化

(右) SHAP value (縦軸) と pIC_{50} (横軸)

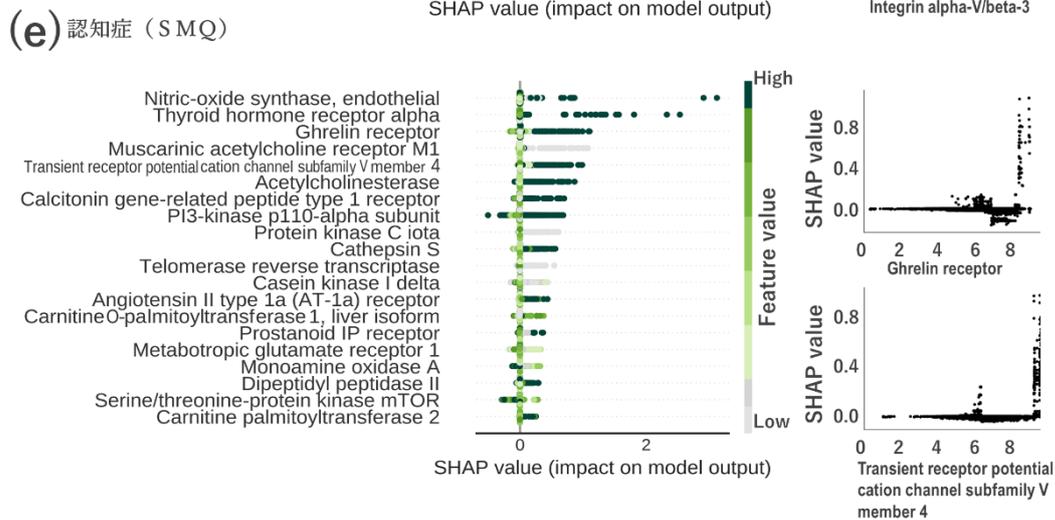
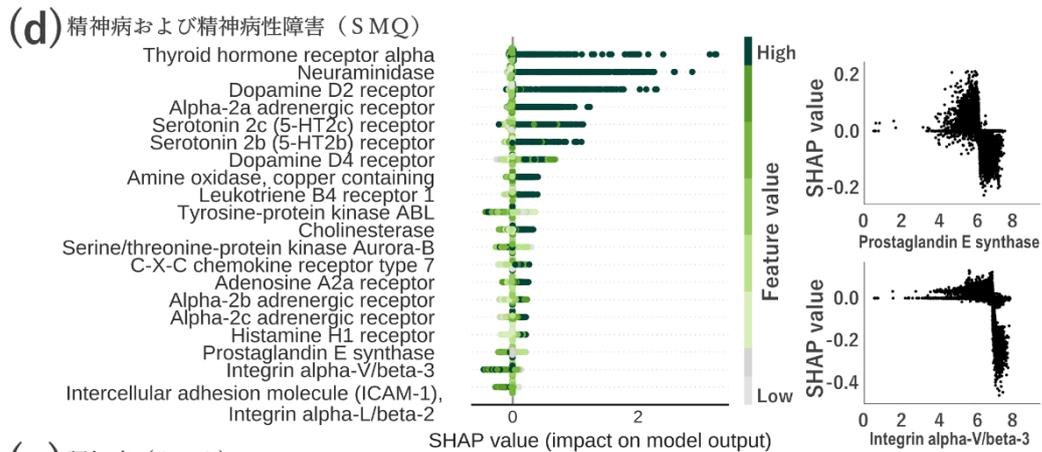


図 11 (d) (e) AE と抽出された標的タンパク質

予測モデル出力結果に対する寄与度の高い 20 種類の標的タンパク質と SHAP value。1つの点は1つの親和性プロファイルを示す。

(左) 標的タンパク質に対する親和性予測値に応じて濃い緑から灰色に変化

(右) SHAP value (縦軸) と pIC_{50} (横軸)

2.4 まとめと展望

JADER に報告されている医薬品を対象に、542 種類の pIC_{50} 予測値からなる親和性プロファイルという薬理活性を入力特徴量として、166 種類の AE 予測モデルを構築することに成功した。100 件以上の教師ラベル 1 のデータが存在していた AE98 種のうち、72 種 (74%) については AUC が 0.7 以上の予測精度を達成した。AE 予測モデルの出力結果に対する特徴量の寄与度を SHAP value によって解析したところ、TdP/QT 延長不整脈に対する hERG の関与など、既知の標的タンパク質を認識していることが見出された。興味深いことに、SHAP value による解析を通じて、AE 発現を抑制する可能性が示唆される標的タンパク質も複数見いだすことができた。自殺/自傷 (SMQ) における HDAC6 や VEGFR2 など、動物実験による検証が困難な事象も含まれていたが、そのマウスにおける表現型からはヒトでの有効性が示唆されていた。これら結果は、AE 予測モデルが AE 発現の原因推定ツールとしての役割のみならず、新規治療標的のヒントを与えるツールとしても機能する可能性を示すものである。

本研究で新たに採用した SHAP value と pIC_{50} 予測値との関係から、薬理活性強度と AE 発現 (あるいは AE 発現抑制) の関係を定量的に推測することが可能となった。従来の定性的な AE 予測とは一線を画す点であり、初期の創薬探索フェーズにおける AE 発現のリスクに着目した探索優先順位づけなどに有用であると考えられる。本研究で構築した AE 予測モデルは AE 発現確率を出力するものであるが、将来的にはこの出力機能を発展させた、例えば AE 発現頻度を定量的に予測するモデルも有用であろう。そしてその際には、JADER のような AE 自発報告データに加え、例えば診療報酬請求データベースなどの薬剤使用母集団が明らかなデータベースが、AE 発現頻度情報のデータソースとして活用されることが重要となると考える。

AE 自発報告データは医療専門家だけでなく一般使用者からの報告も含んでいる。そのため、解析データの正確性を担保するために、医療専門家による報告で、かつ、医療専門家が“被疑薬”と判断したデータのみを教師ラベル 1 として採用する報告も存在する⁷¹。本研究では、AE 発現日以前に使用していた全ての薬剤が AE 発現確率に寄与すると仮定したが、この既報に従い、“被疑薬”のみからなる親和性プロファイルを用いて AE 予測モデルを構築したところ、意外なことに“被疑薬”、“併用薬”、“相互作用薬”の全てを用いたモデルの予測性能と同等あるいは若干低下するという結果であった。さらに重要なことに、SHAP value にて AE 発現への寄与度の大きい標的タンパク質を調べたところ、2.3 結果と考察の節で述べた標的タンパク質の、特に AE 発現抑制に寄与するものが“被疑薬”モデルでは検出されない傾向がみられた。このことは“併用薬”、“相互作用薬”として報告された医薬品の情報を AE 予測モデルに活用することの有用性を示すものと考えられる。

教師ラベルを設定するにあたり、本研究では二項検定と FDR 補正による判定を行った。薬剤疫学分野では AE シグナル検出の目的で報告オッズ比 (ROR) や比例報告比 (PRR) を用いることが多いため、予備的に ROR と PRR の組合せによる教師ラベル

を作成し、TdP/QT 延長 (SMQ) を例として AE 予測モデルを構築し比較したところ、AUC、precision、recall の 3 種の指標において下回る結果となった。ROR、PRR などの不均衡分析においては、分割表であらわされる 4 種のグループいずれにおいても 1 件以上の親和性プロファイルが存在しなければならない制限がある。また、興味の対象となる AE を発現したという報告が、当該の親和性プロファイルにおいて 3 件以上存在することも必要である。これらの制限下では、特定の親和性プロファイル以外には報告されていない希な AE 報告などは解析対象から除外されてしまうため、結果として解析対象データの不足により AE 予測性能が低下したことが想定される。

AE 予測モデルには限界も存在するが、以下の 5 点について今後さらなる改善が可能であると考えている。

1 自発報告のデータ数

報告数と予測性が相関している傾向がみられたことから、JADER の規模の小ささに起因する予測精度の低下が生じている可能性が示唆される。したがって、将来的にはより規模の大きいデータソースを活用することで AE 予測モデルのさらなる性能向上が期待できる。

2 AE 自発報告データに存在する報告バイアス

例えばベンゾジアゼピン系薬剤は薬物依存を起こすことが知られているが、JADER にはあまり報告されていない。結果、薬物乱用および依存 (SMQ) の予測モデルにおいては、ベンゾジアゼピン系薬剤であっても教師ラベルの 0 のものが大多数であった (7 種のベンゾジアゼピン系医薬品のいずれかを含む親和性プロファイルが 3,386 種類存在したが、教師ラベルが 1 となったのは 14 種類のみだった)。自発報告に依らない、例えば診療報酬請求データベースなどを活用することで、報告バイアスを緩和できる可能性が考えられる。

3 AE の独立性

本研究では各 AE は独立した事象と仮定したが、自殺/自傷 (SMQ) とうつ病 (自殺/自傷を除く) (SMQ) など、相互に関連した AE も存在していると思われる。解析データ数と予測性能に相関傾向が見られたことから、将来的には関連 AE のグループ化により解析対象データを増加させることで、予測性能を向上させられる可能性がある。また、"被疑薬"、"併用薬"および"相互作用薬"を同列に扱うことで初めて検出された標的タンパク質が存在したことから (つまり、"被疑薬"のみを対象とした AE 予測モデルでは検出されなかった標的タンパク質を検出しえたことから)、解析対象の薬剤バリエーションの増加により、AE 発現と標的タンパク質の関係においてさらなる洞察がえられる可能性も考えられる。

4 SMQ のカバー範囲

現在公開されている SMQ は全ての AE をカバーしているわけではない。SMQ に代わる AE の適切なグルーピング方法は未だ解決していない課題である。しかし、実臨床における新たな知見に応じて毎年新たな SMQ が開発されていることから、そのような SMQ について継続的に AE 予測モデルを拡充することが重要だと考えられる。

5 重原子数が 1,000 超の医薬品

本研究で対象外とした重原子数が 1,000 を超える医薬品、例えば抗体医薬品などは、特定の標的タンパク質への親和性および特異性が極めて高いと考えられる。したがって、抗体医薬品などについては「治療標的タンパク質に対する親和性実測値」と「治療標的以外の生体分子に対する親和性予測値は 0」とするような親和性プロファイルを設定することも考えられる。これにより、AE 予測モデルに用いるデータの質・量が拡充され、さらに、重原子数が 1,000 超の医薬品が関与する AE の洞察も得られる可能性がある。

以上、本章では、第 1 章で構築した親和性予測モデルの出力する 542 種の標的タンパク質に対する親和性プロファイルを仲介役とし、166 種類の AE 予測モデルの構築に成功した。また、予測確率に寄与する因子の解析を通じて、AE への関与が疑われる標的タンパク質を検出しえることを示した。

全ての”オフターゲット”と呼ばれる生体分子への影響を *in vitro*、*in vivo* 実験によって評価するのは時間的・費用的に現実的ではない。よって本研究で構築した AE 予測モデルは、医薬品探索・開発戦略の優先順位付けなどにおいて、実測値を補足する第 2 の情報源として有用なものになると考えられる。また、AE 発現への関与が疑われる標的タンパク質が浮かび上がったことから、本 AE 予測モデルは新たな医薬品開発標的の提案・探索ツールの 1 つとしても活用できるものと考えられる。

総括

本研究では、近年充実してきた化学構造式や活性実測値のオープンデータベース、および有害事象（AE）自発報告ビッグデータを活用し、化学構造式から種々生体分子に対する親和性を予測し、さらにその親和性予測値を仲介役としてヒトでの AE を予測する機械学習モデルの構築を達成した。

第 1 章

ChEMBL から抽出した 542 種類の多様な標的タンパク質に対して、GCN 手法により、化合物の 2 次元構造情報から抽出した特徴量のみを用いて、定量的な親和性予測モデルを構築することに成功した。既存の画像認識手法に基づく親和性予測モデルと比較したところ、予測性能も構築モデルの種類も共に上回っていた。スキヤフォールドの多様性と予測性の関係を検証したところ、スキヤフォールドの多様性が予測性能に与える影響は限定的であり、親和性予測モデルの未知スキヤフォールド化合物に対する予測適応範囲の広さが示された。予測値の妥当性を検証するために行ったバーチャルスクリーニングにより、既存の抗うつ薬シタロプラムと同等の SERT 阻害活性を有する新規化合物を同定するにいたった。

第 2 章

JADER に報告されている医薬品を対象に、AE 発現前に使用していた全医薬品を入力情報として取り扱った定量的な親和性プロファイルを用いて、166 種類の AE 予測モデルを構築した。100 件以上の教師ラベル 1 のデータが存在していた AE98 種のうち、72 種（74%）について AUC が 0.7 以上と、良好な予測精度を有する AE 予測モデルの構築に成功した。SHAP value を用いた特徴量の寄与度解析を通じて、AE 発現に寄与する標的タンパク質として、ヒトの AE への関連が知られている既知標的タンパク質を抽出し得たのみならず、AE 発現および AE 発現抑制への関与が疑われる新規標的タンパク質を抽出しえる可能性も示した。

以上、著者は化学構造式から生体分子への親和性を高精度に予測する機械学習モデルを確立し、その親和性予測値によりヒトでの薬効・AE が予測可能であることを示した。本研究で得た 2 種類の予測モデルは、既存医薬品のリポジショニングや作用未知の化合物のヒトでの薬理作用を予測するプラットフォームとしての活用が期待できるものである。

謝辞

本研究に際しまして、終始懇切丁寧な御指導と御鞭撻を賜りました京都大学大学院薬学研究科 金子周司教授に謹んで感謝の意を表します。また、直接御指導を賜り、多くの有益な御助言を頂きました京都大学大学院薬学研究科 白川久志准教授並びに永安一樹助教に心より感謝の意を表します。

本研究の遂行に当たり、第1章の *in vitro*、*in vivo* 試験においてきれいなデータを出してくださいました渋井紀宏学士、安藤千紘学士、高山海都学士に心より感謝いたします。また解析データの準備にあたり種々ご協力いただきました古田晴香学士、山本浩貴学士、栢沼玄さんに心より感謝するとともに、京都大学医学部附属病院薬剤部 中川貴之准教授、京都大学医学研究科 浅岡希美助教ならびに京都大学大学院生体機能解析学分野の諸氏に感謝いたします。

本研究の機会を与えてくださったメディカルデータベース株式会社の亀田裕文氏、三原洋一氏、杉平直子氏に心より感謝するとともに、元メディカルデータベース株式会社の池本良平氏に感謝いたします。本研究に際し、在職中のメディカルデータベース株式会社の皆様に多大なご配慮をいただきました。心より感謝いたします。

最後に、研究生活を終始支えてくれた家族に心より感謝いたします。

発表論文目録

第1章

Prediction of pharmacological activities from chemical structures with graph convolutional neural networks.

(グラフ畳み込みニューラルネットワークを用いた化学構造式からの親和性予測)

Miyuki Sakai, Kazuki Nagayasu, Norihiro Shibui, Chihiro Andoh, Kaito Takayama, Hisashi Shirakawa, Shuji Kaneko

Scientific Reports (2021) 11:525 に掲載

第2章

Prediction of adverse drug events and associated molecular targets with affinity profiles.

(受容体親和性予測値を用いた有害事象および原因標的タンパク質の予測)

未発表

参考文献

1. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* **25**, 1097–1105 (2012).
2. Duvenaud, D. *et al.* Convolutional networks on graphs for learning molecular fingerprints. <https://arxiv.org/abs/1509.09292>.
3. Wu, Z. *et al.* MoleculeNet: A benchmark for molecular machine learning. *Chem. Sci.* **9**, 513–530 (2018).
4. DeepChem. <https://github.com/deepchem/deepchem>.
5. Altae-Tran, H., Ramsundar, B., Pappu, A. S. & Pande, V. Low Data Drug Discovery with One-Shot Learning. *ACS Cent. Sci.* **3**, 283–293 (2017).
6. Cai, C. *et al.* Deep Learning-Based Prediction of Drug-Induced Cardiotoxicity. *J. Chem. Inf. Model.* **59**, 1073–1084 (2019).
7. Cheng, W. & Ng, C. A. Using Machine Learning to Classify Bioactivity for 3486 Per- and Polyfluoroalkyl Substances (PFASs) from the OECD List. *Environ. Sci. & Technol.* **53**, 13970–13980 (2019).
8. Rodríguez-Pérez, R. & Bajorath, J. Prediction of Compound Profiling Matrices, Part II: Relative Performance of Multitask Deep Learning and Random Forest Classification on the Basis of Varying Amounts of Training Data. *ACS Omega* **3**, 12033–12040 (2018).
9. Miyazaki, Y., Ono, N., Huang, M., Altaf-Ul-Amin, M. & Kanaya, S. Comprehensive Exploration of Target-specific Ligands Using a Graph Convolution Neural Network. *Mol. Inform.* **39**, 1900095 (2020).
10. Mayr, A. *et al.* Large-scale comparison of machine learning methods for drug target prediction on ChEMBL. *Chem. Sci.* **9**, 5441–5451 (2018).
11. Bosc, N. *et al.* Large scale comparison of QSAR and conformal prediction methods and their applications in drug discovery. *J. Cheminform.* **11**, 4 (2019).
12. Unterthiner, T. *et al.* Deep learning as an opportunity in virtual screening. in *Proceedings of the deep learning workshop at NIPS* vol. 27 1–9 (2014).
13. Gomes, J., Ramsundar, B., Feinberg, E. N. & Pande, V. S. Atomic Convolutional Networks for Predicting Protein-Ligand Binding Affinity. <http://arxiv.org/abs/1703.10603> (2017).
14. Karimi, M., Wu, D., Wang, Z. & Shen, Y. DeepAffinity: interpretable deep learning of compound–protein affinity through unified recurrent and convolutional neural networks. *Bioinformatics* **35**, 3329–3338 (2019).
15. Öztürk, H., Özgür, A. & Ozkirimli, E. DeepDTA: Deep drug-target binding affinity prediction. *Bioinformatics* **34**, i821–i829 (2018).
16. Wang, X. *et al.* Dipeptide Frequency of Word Frequency and Graph Convolutional Networks for DTA Prediction. *Front. Bioeng. Biotechnol.* **8**, 267 (2020).
17. Liu, P., Li, H., Li, S. & Leung, K.-S. Improving prediction of phenotypic drug response on cancer cell lines using deep convolutional network. *BMC Bioinformatics* **20**, 1–14 (2019).
18. Whitehead T. M., Irwin, B. W. J., Hunt, P., Segall, M. D. & Conduit, G. J. Imputation of Assay Bioactivity Data Using Deep Learning. *J. Chem. Inf. Model.* **59**, 1197–1204 (2019).
19. Feinberg, E. N. *et al.* PotentialNet for Molecular Property Prediction. *ACS Cent. Sci.* **4**, 1520–1530 (2018).
20. Karlov, D. S., Sosnin, S., Fedorov, M. V. & Popov, P. GraphDelta: MPNN Scoring Function for the Affinity Prediction of Protein-Ligand Complexes. *ACS Omega* **5**, 5150–5159 (2020).

21. Wu, J. *et al.* Precise modelling and interpretation of bioactivities of ligands targeting G protein-coupled receptors. *Bioinformatics* **35**, i324–i332 (2019).
22. Wang, X. *et al.* Molecule Property Prediction Based on Spatial Graph Embedding. *J. Chem. Inf. Model.* **59**, 3817–3828 (2019).
23. Lenselink, E. B. *et al.* Beyond the hype: deep neural networks outperform established methods using a ChEMBL bioactivity benchmark set. *J. Cheminform.* **9**, 45 (2017).
24. ChemAxon. <https://chemaxon.com/>.
25. RDKit. <https://www.rdkit.org/>.
26. Jiménez, J. & Ginebra, J. pyGPGO: Bayesian Optimization for Python. *J. Open Source Softw.* **2**, 431 (2017).
27. Xu, Y., Pei, J. & Lai, L. Deep Learning Based Regression and Multiclass Models for Acute Oral Toxicity Prediction with Automatic Chemical Feature Extraction. *J. Chem. Inf. Model.* **57**, 2672–2685 (2017).
28. Yang, K. *et al.* Analyzing Learned Molecular Representations for Property Prediction. *J. Chem. Inf. Model.* **59**, 3370–3388 (2019).
29. Kwon, S., Bae, H., Jo, J. & Yoon, S. Comprehensive ensemble in QSAR prediction for drug discovery. *BMC Bioinformatics* **20**, 521 (2019).
30. Bemis, G. W. & Murcko, M. A. The properties of known drugs. 1. Molecular frameworks. *J. Med. Chem.* **39**, 2887–2893 (1996).
31. Godden, J. W. & Bajorath, J. Differential Shannon Entropy as a Sensitive Measure of Differences in Database Variability of Molecular Descriptors. *J. Chem. Inf. Comput. Sci.* **41**, 1060–1066 (2001).
32. Schneider, P. & Schneider, G. Privileged Structures Revisited. *Angew. Chemie Int. Ed.* **56**, 7971–7974 (2017).
33. Asano, M. *et al.* SKF-10047, a prototype Sigma-1 receptor agonist, augmented the membrane trafficking and uptake activity of the serotonin transporter and its C-terminus-deleted mutant via a Sigma-1 receptor-independent mechanism. *J. Pharmacol. Sci.* **139**, 29–36 (2019).
34. Ramamoorthy, S. *et al.* Antidepressant- and cocaine-sensitive human serotonin transporter: Molecular cloning, expression, and chromosomal localization. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 2542–2546 (1993).
35. Nishitani, N. *et al.* Manipulation of dorsal raphe serotonergic neurons modulates active coping to inescapable stress and anxiety-related behaviors in mice and rats. *Neuropsychopharmacology* **44**, 721–732 (2019).
36. Mervin, L. H. *et al.* Target prediction utilising negative bioactivity data covering large chemical space. *J. Cheminform* **7**, 51 (2015).
37. Romeo, G. *et al.* New pyrimido[5,4-b]indoles as ligands for α 1-adrenoceptor subtypes. *J. Med. Chem.* **46**, 2877–2894 (2003).
38. Koutsoukas, A., Monaghan, K. J., Li, X. & Huan, J. Deep-learning: Investigating deep neural networks hyper-parameters and comparison of performance to shallow methods for modeling bioactivity data. *J. Cheminform.* **9**, 42 (2017).
39. Willmott, C. J. & Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **30**, 79–82 (2005).
40. Chai, T. & Draxler, R. R. Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* **7**, 1247–1250 (2014).
41. Li, Q., Han, Z. & Wu, X.-M. Deeper Insights into Graph Convolutional Networks for Semi-Supervised Learning. in *32nd AAAI Conference on Artificial Intelligence, AAAI 2018* 3538–3545 (2018).
42. Goh, G. B., Siegel, C., Vishnu, A., Hodas, N. O. & Baker, N. Chemception: A Deep Neural Network with Minimal Chemistry Knowledge Matches the

- Performance of Expert-developed QSAR/QSPR Models. <https://arxiv.org/abs/1706.06689>.
43. Cortés-Ciriano, I. & Bender, A. KekuleScope: prediction of cancer cell line sensitivity and compound potency using convolutional neural networks trained on compound images. *J. Cheminform.* **11**, 41 (2019).
 44. Uesawa, Y. Quantitative structure–activity relationship analysis using deep learning based on a novel molecular image input technique. *Bioorganic Med. Chem. Lett.* **28**, 3400–3403 (2018).
 45. Hirohara, M., Saito, Y., Koda, Y., Sato, K. & Sakakibara, Y. Convolutional neural network based on SMILES representation of compounds for detecting chemical motif. *BMC Bioinformatics* **19**, 526 (2018).
 46. Nidhi, Glick, M., Davies, J. W. & Jenkins, J. L. Prediction of Biological Targets for Compounds Using Multiple-Category Bayesian Models Trained on Chemogenomics Databases. *J. Chem. Inf. Model.* **46**, 1124–1133 (2006).
 47. Shang, J. *et al.* Comparative analyses of structural features and scaffold diversity for purchasable compound libraries. *J. Cheminform.* **9**, 25 (2017).
 48. Li, Y., Zhang, L. & Liu, Z. Multi-objective de novo drug design with conditional graph generative model. *J. Cheminform.* **10**, 33 (2018).
 49. Paricharak, S. *et al.* Data-driven approaches used for compound library design, hit triage and bioactivity modeling in high-throughput screening. *Brief. Bioinform.* **19**, 277–285 (2018).
 50. Zhang, Y. & Lee, A. A. Bayesian semi-supervised learning for uncertainty-calibrated prediction of molecular properties and active learning. *Chem. Sci.* **10**, 8154–8163 (2019).
 51. Robinson, M. C., Glen, R. C. & Lee, A. A. Validating the validation: reanalyzing a large-scale comparison of deep learning and machine learning models for bioactivity prediction. *J. Comput. Aided. Mol. Des.* **34**, 717–730 (2020).
 52. Tatsumi, M., Groshan, K., Blakely, R. D. & Richelson, E. Pharmacological profile of antidepressants and related compounds at human monoamine transporters. *Eur. J. Pharmacol.* **340**, 249–258 (1997).
 53. Tarasova, O. A. *et al.* QSAR Modeling Using Large-Scale Databases: Case Study for HIV-1 Reverse Transcriptase Inhibitors. *J. Chem. Inf. Model.* **55**, 1388–1399 (2015).
 54. 薬食安発0328007. <https://www.pmda.go.jp/files/000143457.pdf>.
 55. Sultana, J., Cutroneo, P. & Trifirò, G. Clinical and economic burden of adverse drug reactions. *J. Pharmacol. Pharmacother.* **4**, S73 (2013).
 56. Hwang, T. J. *et al.* Failure of investigational drugs in late-stage clinical development and publication of trial results. *JAMA Intern. Med.* **176**, 1826–1833 (2016).
 57. Bowes, J. *et al.* Reducing safety-related drug attrition: The use of in vitro pharmacological profiling. *Nat. Rev. Drug Discov.* **11**, 909–922 (2012).
 58. Siramshetty, V. B. *et al.* WITHDRAWN—a resource for withdrawn and discontinued drugs. *Nucleic Acids Res.* **44**, D1080–D1086 (2016).
 59. Amouzadeh, H. R. *et al.* Clinical Implications and Translation of an Off-Target Pharmacology Profiling Hit: Adenosine Uptake Inhibition In Vitro. *Transl. Oncol.* **12**, 1296–1304 (2019).
 60. FAERS. <https://www.fda.gov/drugs/surveillance/questions-and-answers-fdas-adverse-event-reporting-system-faers>.
 61. 副作用が疑われる症例報告に関する情報 | 独立行政法人 医薬品医療機器総合機構. <https://www.pmda.go.jp/safety/info-services/drugs/adr-info/suspected-adr/0005.html>.
 62. Zhou, Y. *et al.* Hypoglycemia associated with direct-acting anti-hepatitis C virus drugs: An epidemiologic surveillance study of the FDA adverse event reporting system (FAERS). *Clin. Endocrinol. (Oxf)*. (2021) doi:10.1111/CEN.14660.

63. Anastopoulos, I. N., Herczeg, C. K., Davis, K. N. & Dixit, A. C. Multi-Drug Featurization and Deep Learning Improve Patient-Specific Predictions of Adverse Events. *Int. J. Environ. Res. Public Health* **18**, 2600 (2021).
64. Nagashima, T., Shirakawa, H., Nakagawa, T. & Kaneko, S. Prevention of antipsychotic-induced hyperglycaemia by vitamin D: a data mining prediction followed by experimental exploration of the molecular mechanism. *Sci. Rep.* **6**, 26375 (2016).
65. Nagaoka, K. *et al.* Striatal TRPV1 activation by acetaminophen ameliorates dopamine D2 receptor antagonist-induced orofacial dyskinesia. *JCI Insight* **6**, e145632 (2021).
66. Kuhn, M., Letunic, I., Jensen, L. J. & Bork, P. The SIDER database of drugs and side effects. *Nucleic Acids Res.* **44**, D1075–D1079 (2016).
67. Wishart, D. S. *et al.* DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Res.* **46**, D1074–D1082 (2018).
68. Günther, S. *et al.* SuperTarget and Matador: resources for exploring drug-target relationships. *Nucleic Acids Res.* **36**, D919–D922 (2008).
69. Wang, C. S. *et al.* Detecting Potential Adverse Drug Reactions Using a Deep Neural Network Model. *J. Med. Internet Res.* **21**, e11016 (2019).
70. Mizutani, S., Pauwels, E., Stoven, V., Goto, S. & Yamanishi, Y. Relating drug-protein interaction network with drug side effects. *Bioinformatics* **28**, i522–i528 (2012).
71. Ietswaart, R. *et al.* Machine learning guided association of adverse drug reactions with in vitro target-based pharmacology. *EBioMedicine* **57**, 102837 (2020).
72. Gaulton, A. *et al.* The ChEMBL database in 2017. *Nucleic Acids Res.* **45**, D945–D954 (2017).
73. メディカルデータベース株式会社. <https://www.medicaldb.co.jp/>.
74. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res.* **45**, D353–D361 (2017).
75. Chen, T. & Guestrin, C. XGBoost: A scalable tree boosting system. in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (2016).
76. Lundberg, S. M., Erion, G. G. & Lee, S.-I. Consistent Individualized Feature Attribution for Tree Ensembles. <http://arxiv.org/abs/1802.03888> (2018).
77. Lundberg, S. M. *et al.* From Local Explanations to Global Understanding with Explainable AI for Trees. *Nat. Mach. Intell.* **2**, 56–67 (2020).
78. Kojima, T. *et al.* High risk of adverse drug reactions in elderly patients taking six or more drugs: Analysis of inpatient database. *Geriatr. Gerontol. Int.* **12**, 761–762 (2012).
79. Masnoon, N., Shakib, S., Kalisch-Ellett, L. & Caughey, G. E. What is polypharmacy? A systematic review of definitions. *BMC Geriatr.* **17**, 230 (2017).
80. Zhu, A. *et al.* A Super-combo-drug test (SupCD-T) to Detect Adverse Drug Events and Drug Interactions from Electronic Health Records in the Era of Polypharmacy. *Stat. Med.* **39**, 1458–1472 (2020).
81. Galeano, D., Li, S., Gerstein, M. & Paccanaro, A. Predicting the frequencies of drug side effects. *Nat. Commun.* **11**, 4575 (2020).
82. Chawla, N. V., Bowyer, K. W., Hall, L. O. & Kegelmeyer, W. P. SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002).
83. Tomek, I. Two Modifications of CNN. *IEEE Trans. Syst. Man Commun.* **6**, 769–772 (1976).
84. Lupoglazoff, J. M. *et al.* Long QT syndrome in neonates: conduction disorders associated with HERG mutations and sinus bradycardia with KCNQ1 mutations. *J. Am. Coll. Cardiol.* **43**, 826–830 (2004).

85. Sanguinetti, C. M. & Tristani-Firouzi, M. hERG potassium channels and cardiac arrhythm. *Nature* **440**, 463–469 (2006).
86. Wang, N., Orr-Urtreger, A., Chapman, J., Rabinowitz, R. & Korczyn, A. D. Deficiency of nicotinic acetylcholine receptor beta 4 subunit causes autonomic cardiac and intestinal dysfunction. *Mol. Pharmacol.* **63**, 574–580 (2003).
87. Liu, V. W. T. & Huang, P. L. Cardiovascular roles of nitric oxide: A review of insights from nitric oxide synthase gene disrupted mice. *Cardiovasc. Res.* **77**, 19–29 (2008).
88. Monahan, B. P., Ferguson, C. L., Killeavy, E. S., Lloyd, B. K., Troy, J., & Cantilena, L. R. Torsades de Pointes Occurring in Association With Terfenadine Use. *JAMA* **264**, 2788–2790 (1990).
89. Roy, M. L., Dumaine, R. & Brown, A. M. HERG, a Primary Human Ventricular Target of the Nonsedating Antihistamine Terfenadine. *Circulation* **94**, 817–823 (1996).
90. Svenningsson, P. *et al.* Eltoprazine counteracts l-DOPA-induced dyskinesias in Parkinson's disease: a dose-finding study. *Brain* **138**, 963–973 (2015).
91. Boom, M. *et al.* Non-analgesic effects of opioids: opioid-induced respiratory depression. *Curr. Pharm. Des.* **18**, 5994–6004 (2012).
92. Ashok, A. H., Mizuno, Y., Volkow, N. D. & Howes, O. D. Association of Stimulant Use With Dopaminergic Alterations in Users of Cocaine, Amphetamine, or Methamphetamine: A Systematic Review and Meta-analysis. *JAMA Psychiatry* **74**, 511–519 (2017).
93. Sweidan, A. J. *et al.* Coagulopathy reversal in intracerebral haemorrhage. *Stroke Vasc. Neurol.* **5**, 29–33 (2020).
94. Nørgård, B. *et al.* COX-2-selective inhibitors and the risk of upper gastrointestinal bleeding in high-risk patients with previous gastrointestinal diseases: a population-based case-control study. *Aliment. Pharmacol. Ther.* **19**, 817–825 (2004).
95. Roden, D. M. Long-QT syndrome. *N. Engl. J. Med.* **358**, 169–176 (2008).
96. Huang, P. L. *et al.* Hypertension in mice lacking the gene for endothelial nitric oxide synthase. *Nature* **377**, 239–242 (1995).
97. Cowart, M. *et al.* Nitroaromatic Amino Acids as Inhibitors of Neuronal Nitric Oxide Synthase. *J. Med. Chem.* **41**, 2636–2642 (1998).
98. PMDA. 薬剤惹起性うつ病. <https://www.pmda.go.jp/files/000144134.pdf>.
99. Jeong, H. E., Jeon, H. L., Oh, I. S., Kim, W. J. & Shin, J. Y. Risk of mortality associated with concomitant antidepressant and benzodiazepine therapy among patients with depression: a population-based cohort study. *BMC Med.* **18**, 387 (2020).
100. Jochems, J. *et al.* Antidepressant-like properties of novel HDAC6-selective inhibitors with improved brain bioavailability. *Neuropsychopharmacology* **39**, 389–400 (2014).
101. De Rossi, P. *et al.* A critical role for VEGF and VEGFR2 in NMDA receptor synaptic function and fear-related behavior. *Mol. Psychiatry* **21**, 1768–1780 (2016).
102. Dong, C. *et al.* Psychedelic-inspired drug discovery using an engineered biosensor. *Cell* **184**, 2779–2792 (2021).
103. Nishitani, N. *et al.* Raphe AMPA receptors and nicotinic acetylcholine receptors mediate ketamine-induced serotonin release in the rat prefrontal cortex. *Int. J. Neuropsychopharmacol.* **17**, 1321–1326 (2014).
104. Introductory Guide for Standardised MedDRA Queries (SMQs) Version 22.1. (2019).
105. Kitaoka, S. *et al.* Prostaglandin E2 acts on EP1 receptor and amplifies both dopamine D1 and D2 receptor signaling in the striatum. *J. Neurosci.* **27**, 12900–12907 (2007).

106. Swanson, J. M. *et al.* Etiologic subtypes of attention-deficit/hyperactivity disorder: Brain imaging, molecular genetic and environmental factors and the dopamine hypothesis. *Neuropsychol. Rev.* **17**, 39–59 (2007).
107. Dohn, M. R. *et al.* The Gain-of-Function Integrin $\beta 3$ Pro33 Variant Alters the Serotonin System in the Mouse Brain. *J. Neurosci.* **37**, 11271–11284 (2017).
108. Pan, H., Dohn, M. R., Kingston, R. & Carneiro, A. M. D. Integrin $\alpha v \beta 3$ function influences citalopram immobility behavior in the tail suspension test. *Front. Neurosci.* **13**, 70 (2019).
109. Hornsby, A. K. E. *et al.* Unacylated-Ghrelin Impairs Hippocampal Neurogenesis and Memory in Mice and Is Altered in Parkinson’s Dementia in Humans. *Cell reports. Med.* **1**, 100120 (2020).
110. Diaz-Otero, J. M. *et al.* Transient receptor potential vanilloid 4 channels are important regulators of parenchymal arteriole dilation and cognitive function. *Microcirculation* **26**, e12535 (2019).
111. Cheung, M. *et al.* Discovery of GSK2193874: An Orally Active, Potent, and Selective Blocker of Transient Receptor Potential Vanilloid 4. *ACS Med. Chem. Lett.* **8**, 549–554 (2017).