薬 剤 学, **79** (2), 89–93 (2019)

≪若手研究者紹介≫

機械学習を利用した薬物の経皮吸収性予測

馬 場 廣 海* Hiromi Baba
マルホ株式会社 探索研究部 探索動態グループ セクションチーフ

1. は じ め に

私は現在，マルホ株式会社の京都 R & D センター内にある探索研究部にて計算機を用いた薬剤学研究ならびに創薬研究に従事している．大学の学部，大学院修士課程で有機化学を専攻していた私が，格式ある日本薬剤学会学会誌の本コラムに執筆させていただくまでの経緯とともに，掲題の研究をご紹介したい．

2. 学生時代の研究と気づき

私は理数科に所属していた高校生の頃から，ある現象を記述する際には，その現象に特徴的な表層的表現ではなく，その現象を支配する最も根本的な原理から組み立てて説明したいとの思いがあった．その観点で，いずれ薬学や工学など産業に通じる応用科学的学問に進む可能性があるとしても，最初に学ぶのは基礎科学（理学）にしたいと考え，京都大学理学部を選択したことを今でも覚えている．入学当時の 2000 年代前半までに，卒業生の自然科学系ノーベル賞数が多かったこと，実家の大阪府東大阪市から通学可能な距離（後に電車での通学を省略した く下宿したが…）であったことも少なからず影響していたと思う．

さて，志望大学に入学した後，各種理学について大学の講義・演習はもちろんのこと，種々の教科書を用いた自学にて真面目に勉学に勤しんだ．また，興味のあったコンピュータプログラミング（C/C++ 言語）の学習や，PC の分解あるいは PC の自作（CPU，マザーボードなど各パーツを選定・接続して PC を組み上げること）などを行っていた．コンピュータプログラムは，ある課題に対して，いくつかの同様の処理をまとめて対応可能とした汎用化モジュールを作成することになるが，そのこと自体が現象を原理から説明したいという私の性に合っていたのかもしれない．物理化学の講義課題であった気体の状態方程式の解釈において，気体挙動のシミュレーションプログラムを作成し，その解析結果が評価されたことが嬉しかった．

卒業研究では，最も学問的に好きであった有機化学を選択し，集合有機分子機能研究室（主宰：大須賀篤弘教授）に所属した．ロジウムという遷移金属を触媒的に用いたポルフィリンの修飾反応の開発[1]というテーマで，忍久保洋助教授（現：名古屋大学大学院工学研究科教授）にご指導いただいた．ポルフィリンの反応位置ならびに反応基質によって劇的に反応形式が遷移する興味深い結果が得られ，芳香環ポルフィリン，カルボニル反応剤と遷移金属の 3 者が関係する複雑な反応も，電子の挙動から端的に説明できる理学の素晴らしさに気づいたとき，当時の私にとって，当研究が理学部で学んできたことの集大成となったことを感じた．

大学院進学時には，興味を感じていた天然物の全

*2007 年京都大学理学部卒業，2009 年京都大学大学院薬学研究科修士課程修了．同年，マルホ株式会社入社．以来，薬物動態研究と計算創薬研究を担当．2017 年京都大学大学院薬学研究科博士後期課程修了，博士（薬科学）．2015 年日本薬剤学会第 30 年会最優秀発表者賞，2017 年第 33 回日本 DDS 学会学術集会優秀発表賞，2018 年第 34 回日本 DDS 学会学術集会優秀発表賞，2018 年日本薬剤学会第 43 回製剤・創剤セミナー Post-doctoral Presentation Award 受賞．連絡先：〒600–8891 京都府京都市下京区中堂寺粟田町 93 番地
E-mail: baba_dfq@mii.maruho.co.jp

合成研究を実施している「研究室」を選択したいと考えていた．京都大学薬学研究科に薬品合成化学研究室があることを知り，主宰する富岡清教授（現：京都大学名誉教授）に進学希望をお伝えした際,「全合成は楽しいよ」と優しそうに仰ってくださったことから進学を決意した．なお，大層厳しい先生であること，研究室は一晩中，人と明かりが消えることのない不夜城と化していることを知ったのは進学後であった．朝に合成反応をかけ，日中に反応進行の追跡とセミナー等を行い，夜に反応終結を確認し，深夜にカラム精製となり，ときに数リットルに及ぶ溶媒の留去と NMR 分析を終えるのは，健康的な学生が登校してくる朝という生活を続けていた（2, 3 日帰宅しない日もあった）．当時, 同様の生活にあった同期と先輩方とは，深夜に有機化学や物理化学，将来の夢を熱く語っており，今にして思えば青春であったように思う．苦楽を共にした同期や先輩とは今でも連絡を取り合う仲であり，科学や将来の相談をすることもある．アイディアは，人と話し，ときに批判にさらされて洗練されていくものであると強く思うようになったのも，この時分である．修士課程の 2 年間は，現在の研究にも活用できる有機化学という学問の習熟の他に，かけがえのない仲間を得ることのできた大切な時間であったように思う．

3. 計算薬剤学・計算創薬との出会い

博士課程と就職で迷っていた私に，多分野が混じり合い新しい研究領域を開拓していけること，温かな人間関係を築いている社風があることを説いていただいたことから，京都に研究所のあるマルホ株式会社に就職することに決めた．当時，合成化学系の部署に所属するものとばかり思っていたが，配属されたのは薬物動態研究グループであった．そこでは，これまで実践したことのなかった ADME・PK に関する各種 *in vitro*/*in vivo* 試験に携わることができ，大変興味深かったことを覚えている．入社後半年ほどが経過し，新しい研究分野の体系を理解し，吸収することに精一杯であった時期を過ぎると，とりわけ ADME 試験において化学的バックグラウンドが有効に機能すること，また，経皮吸収性の改善に向けた製剤的施策など，それぞれの研究分野において未解決の課題があることに気づき，研究者としての楽しみが見つかったことに喜びを感じていた．各種実験方法やその改善点，結果解釈，研究目標などに意見する生意気な私に，良きアドバイザーとして理解・支援していただいたのが上司であった高原淳一博士（現：探索動態グループマネージャー；当時からいくつもの社内異動をほぼ同じくしており，現在も私の上司）である．

その上司から，化学的素養とコンピュータプログラミングスキルを有しているとの理由でいただいたテーマが,「機械学習を利用した薬物の経皮吸収性の予測法の開発」である．これが，計算薬剤学に携わるきっかけとなったテーマであり，現在もライフワークの一環として，その発展的研究に取り組んでいる．本テーマは後述するとして，その後，私は薬物動態研究グループから先端的製剤技術開発を担当する部署に移り，弊社の製剤研究者である堀沢栄次郎博士（日本薬剤学会 製剤の達人）とともに各種アカデミアや海外企業との新規製剤技術の共同開発に取り組むことになった．堀沢博士の豊かな経験と鋭利な推察にいつも驚かせられながら，軟膏，クリーム，ローション，テープ剤などコンベンショナルな製剤の製法・特徴を学んだのもこの頃である．現在においても，注目する先端製剤技術についての議論を随時続けており，研究者としての賦質と修練を試されているようで，如何なるときも対応できるように日々文献等を用いて勉学に励むようにしている．その後，化学，薬物動態，各種デジタル技術のスキルを活用できると判断され，探索研究を担当する部署に移り，各種製剤の特性を計算的に解明する計算製剤研究のほか，量子化学・AI を利用した受容体–候補薬物の結合親和性解析, ファーマコフォア解析, 化合物クラスタリング，ADME・毒性についての定量的構造物性相関（QSPR）モデル構築，薬物動態モデリングならびに分子動力学シミュレーションなどの計算創薬も担当している．企業での研究を早期に発表するためには，所属企業から公表承認の取得と学会の先生方にお認めいただくことが必要になるが，担当した研究の中で，幸運にも公表の機会に恵まれた「機械学習を利用した薬物の経皮吸収性の予測法の開発」についてご紹介したく思う．

4. 機械学習を利用した薬物の経皮吸収性の予測法の開発

薬物の経皮吸収性は，各種外用医薬品の開発など

において，その有効性・リスク評価の観点から，極めて重要視されるものである．従来，薬物の経皮吸収性は，摘出皮膚を装着した拡散セルを用いた *in vitro* 透過性試験や，げっ歯類・ミニブタ等を用いた *in vivo* 経皮吸収性試験により評価してきた．一方，コンピュータ計算によって薬物の経皮吸収性を推算する *in silico* 予測法[2]では，上記実験手法と比較して，スループット（時間・コスト）の観点で有利なだけでなく，生体試料・実験動物の使用回避という倫理上の優位性を有すること，さらに，デザイン段階の薬物も評価可能であることなどの利点がある．これらの特徴は，多種の候補薬物の効率的評価が求められる医薬品開発の初期ステージにおいて，極めて有利に機能する．

そこで，まず 1960 年代以降の *in vitro* 皮膚透過性試験の文献を集積した上で，当該試験にて算出される透過係数の値を収集した．各文献間の誤差を最小限とするため，摘出ヒト皮膚と拡散セルを用いていること，各種透過促進技術を利用していないことなど選定基準を設定し，それを厳密に適用することでデータの選別を行った．結果として，投与液を水溶液とした 211 種の薬物のヒト皮膚透過係数を含む，当時で最大規模の新規データベースを構築することができた．データベース内の全化合物の 3 次元モデリングを行い，その構造に基づいて 2,732 種類もの分子記述子を算出した．この 211×2,732 の記述子マトリクスと皮膚透過性の実測値について，非線形の機械学習法であるガウスカーネルを用いたサポートベクター回帰（SVR）やランダムフォレスト（RF）法を適用することで予測モデルの構築を試みることとした．まず，モデル検証を実施するため，あらかじめデータベースを学習用セット（169 薬物）と検証用セット（42 薬物）に 8：2 の割合でランダム分割した．適切なモデル検証のためには，学習用セットと検証用セットの化学的特徴に偏りがないことが必要となるため，主成分分析を用いて両セットのケミカルスペースを解析し，両者に差がないことを確認した（図 1）．

不要な分子記述子の予測モデルへの組み込みは，計算負荷・予測のロバスト性の観点から不利である．しかしながら，経皮吸収性に関連する各種メカニズムの詳細が完全に解明されていない状況において，経皮吸収性を説明するための分子記述子をあらかじめ指定することは容易ではない．そこで，本研究では，学習用セットに対し 10-fold 交差検定法における予測性能を指標とした分子記述子の逐次選定法（greedy algorithm）を採用することにした．学習用セットに対し，上記の記述子選定法と機械学習法（SVR，RF）を組み合わせて適用した結果，過剰適合が適切に回避されながら，未知サンプルとなる検証用セット内の薬物の経皮吸収性を正確に推算（予測－実測の間の $R^2 \geq 0.88$，RMSE ≤ 0.39）する優れた予測モデルを構築[3]することができた（図 2）．

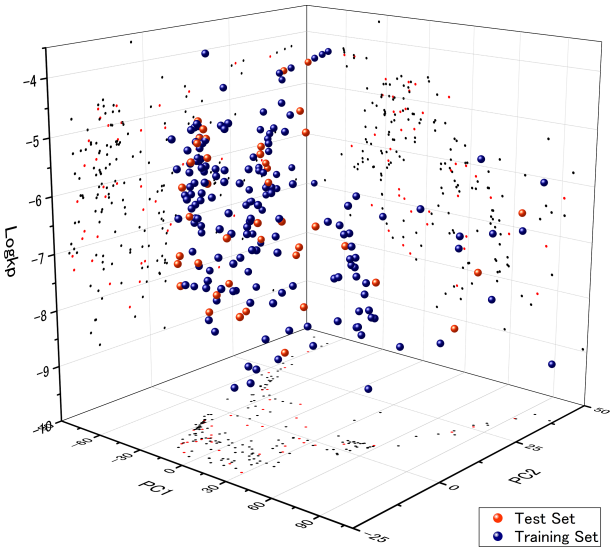

図 1　学習用セットと検証用セットのケミカルスペース比較（PC1：第 1 主成分，PC2：第 2 主成分，Logkp：ヒト皮膚透過係数）．許諾を得て文献 3）より引用．

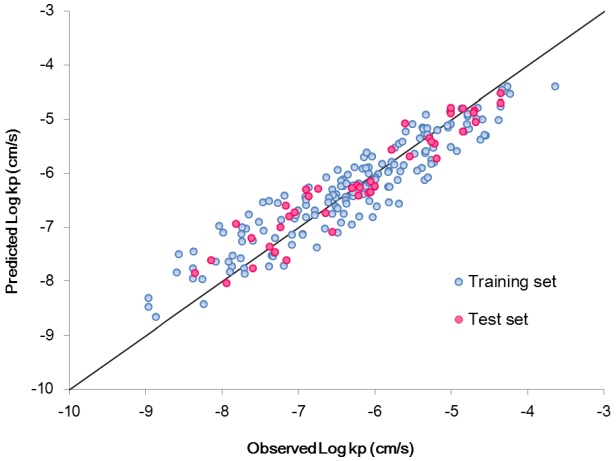

図 2　SVR モデルにおける薬物の経皮吸収性（Logkp）の予測-実測の関係性．許諾を得て文献 3）より引用．

本結果は，幾重もの素過程を経る複雑な経皮吸収性を薬物の化学構造から予測できたことを意味しており，溶媒組成・pH などのより複雑な投与液環境の影響を包括的に推算する予測モデル構築への道を拓くものである．なお，本研究は，京都大学化学研究所バイオインフォマティクスセンター生命知識工学研究領域（主宰：馬見塚拓教授）に受託研究員として所属し，馬見塚教授のご指導の下，完成させたものである．

5. 投与液環境を考慮した予測モデルへの拡張

一般に，pH や溶媒組成といった投与液環境は，薬物の皮膚透過性に甚大な影響を与えることが知られている．例えば，ketrolac や ibuprofen などは，投与液の溶媒組成を変更することで，ヒト皮膚透過係数が 100 倍以上変化することが知られている．また，サリチル酸については，投与液 pH が 5.13 と 2.10 の場合で，ヒト皮膚透過係数が 60 倍程度異なる．しかしながら，既存の経皮吸収性予測モデルのほぼ全ては溶媒環境を考慮しておらず，外用製剤開発における予測モデルの実用的価値が高いとは言い難い状況であった．

そこで，溶媒組成ならびに投与液 pH の影響を評価するためのヒト皮膚透過性の新規データセットを先述の方法と同様に編纂した．サンプル数はそれぞれ 412 ならびに 322 であり，それらのデータベース内の薬物のケミカルスペースは，外用薬として用いられる化合物群と比較しても遜色のないものであった．ここで，溶媒組成の影響を考慮した予測モデルについては，各投与液の特徴を主薬の記述子と溶媒の記述子で表すことを考えた．とりわけ，溶媒の記述子に関しては，投与液の溶媒構成比に従い各溶媒の記述子を加重平均化することで，溶媒組成の影響を組み込むこととした．また，投与液 pH の影響に関しては，主薬の投与液 pH における水–オクタノール分配係数（LogD）を利用することとした．先述と同様の方法にて記述子選択と予測モデル構築を行い，結果として，検証用セットにおける予測－実測の間の R^2 値が 0.89 を超える予測モデル[4, 5]を確立できた．構築した予測モデルについて，近年導入されてきた 10 種以上の統計的指標を利用して検証した結果，いずれの指標のクライテリアも満足することが明らかとなり，予測性能に対する一定の信頼性についても確認することができた．

また，一次微分感度解析による記述子影響度評価法を開発[4, 5]したことにより，一般に困難である非線形機械学習モデル内の記述子の影響度を評価することを可能とした（図 3）．当該方法論を用いて解析した結果，溶媒組成・投与液 pH に関するいずれの予測モデルに対しても，最大の影響度を呈した記述子は主薬の水–オクタノール分配係数であることが明らかとなった．水溶液投与モデルにおいて，当該記述子の影響度の大きさはかねてから指摘されてきたが，格段に複雑な溶媒効果を扱う予測モデルにおいても，なお当該記述子の重要性が示唆されたことは極めて興味深い知見であり，より実製剤に近い投与液環境の影響を包括的に予測する方法論の開発に対する重要な示唆を与えるものと考えている．

なお，当該研究は，京都大学大学院薬学研究科の博士後期課程で薬品動態制御学分野に在籍した際に行った研究である．主宰する橋田充教授（現：京都大学高等研究院物質–細胞統合システム拠点特定教授・京都大学名誉教授）とともにモデリング & シミュレーション技術を用いた薬物動態解析の第一人者

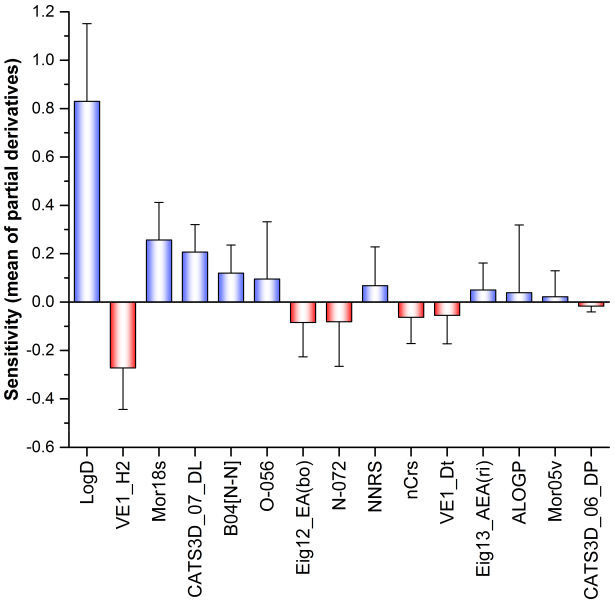

図 3　一次微分感度解析法による投与液 pH 考慮の SVR モデル内の記述子の影響度．許諾を得て文献 5）より引用．記述子の算出においては，LogD に対して ADMET Predictor version 7.2（Simulations Plus Inc., http://www.simulations-plus.com/software/admetpredictor/），それ以外の記述子に対しては Dragon version 7.0.2（Kode srl, https://chm.kode-solutions.net/）を用いた．

である山下富義教授に，処方の特徴化や記述子の影響度解析法に関して種々のご助言とご指導をいただいて完成させたものである．

6. おわりに

上記研究は，経皮吸収性に焦点を当てたものであるが，手法自体は他の ADME・毒性などの予測に対しても汎く適用できるものと考える．また，完全に computational な処理のみによって予測が可能であり，仮想薬物や仮想処方でさえも推算可能な当方法論に対して，研究対象としての興味深さとともに，効率性が強く求められる医薬品の研究開発業務上の有用性・将来性を感じており，現在も研究を進めている．日本薬剤学会の発展に少しでも資する形で成果を残すことができるなら幸甚である．

ご高名で極めてご多用な先生方がお時間を割いてご指導くださったことに深く感謝し，少しでも報いることができるように，また研究遂行にあたり種々の支援をいただいた所属会社に貢献できるように，日々学び，研鑽し続けたいと強く考えている．また，研究に集中できる環境を整えてくれた妻に，日々癒しをくれる息子に，幼い頃から勉強をすることを支援してくれた両親と兄に，ここで感謝の気持ちを記しておきたいと思う．

引用文献

1) H. Baba, J. Chen, H. Shinokubo, A. Osuka, Efficient rhodium-catalyzed installation of unsaturated ester functions onto porphyrins: site-specific Heck-type addition versus conjugate addition, *Chem. Eur. J.*, **14**, 4256–4262 (2008).
2) H. Baba, J. Takahara, F. Yamashita, *In Silico* Approach. In: K. Sugibayashi, ed., Skin Permeation and Disposition of Therapeutic and Cosmeceutical Compounds, Springer, pp. 363–376 (2017).
3) H. Baba, J. Takahara, H. Mamitsuka, In silico predictions of human skin permeability using nonlinear quantitative structure-property relationship models, *Pharm. Res.*, **32**, 2360–2371 (2015).
4) H. Baba, J. Takahara, F. Yamashita, M. Hashida, Modeling and prediction of solvent effect on human skin permeability using support vector regression and random forest, *Pharm. Res.*, **32**, 3604–3617 (2015).
5) H. Baba, Y. Ueno, M. Hashida, F. Yamashita, Quantitative prediction of ionization effect on human skin permeability, *Int. J. Pharm.*, **522**, 222–233 (2017).