

# 日本大学医学部 臨床試験研究センター研究紀要

---

Bulletin of the Clinical Trials Research Center,  
Nihon University School of Medicine

日本大学医学部臨床試験研究センター  
Clinical Trials Research Center, Nihon University School of Medicine

## 目次

ビッグデータを用いた研究基盤-バイオバンク・ジャパンの活用 .....	高橋泰夫 他	1
機械学習を活用した薬物性肝障害のリスクに影響を及ぼす薬物間相互作用の検出 .....	秋元勇人 他	6
低ナトリウム血症を起こしにくい抗うつ薬の同定 .....	長島卓也 他	15
日本大学医学部臨床情報統合データベースに基づき疾患リスク解析を行う 機械学習パイプラインの構築 .....	早川隆 他	19
DLGAP1 ノックアウトマウスの表現型解析 .....	皆川貴美乃 他	25

## INDEX

Research platform using big data - utilization of BioBank Japan .....	Yasuo TAKAHASHI et.al	1
Detection of drug-drug interaction on drug-induced liver injury using supervised machine-learning algorithms .....	Hayato AKIMOTO et.al	6
Identifying antidepressants less likely to cause hyponatremia .....	Takuya NAGASHIMA et.al	15
A machine learning pipeline for risk analysis based on the integrative database of clinical information in NUSM .....	Takashi HAYAKAWA et.al	19
Behavioral characterization of DLGAP1 knockout mice with SHIRPA primary screen .....	Kimino MINAGAWA et.al	25

## ビッグデータを用いた研究基盤-バイオバンク・ジャパンの活用

高橋泰夫<sup>1)</sup>, 皆川貴美乃<sup>1)</sup>, 秋元勇人<sup>1)2)</sup>, 長島卓也<sup>1)2)</sup>,  
早川隆<sup>1)2)</sup>, 浅井聡<sup>1)2)</sup>

## Research platform using big data - utilization of BioBank Japan

Yasuo TAKAHASHI<sup>1)</sup>, Kimino MINAGAWA<sup>1)</sup>, Hayato AKIMOTO<sup>1)2)</sup>,  
Takuya NAGASHIMA<sup>1)2)</sup>, Takashi HAYAKAWA<sup>1)2)</sup>, Satoshi Asai<sup>1)2)</sup>

キーワード : clinical database, real world data, deep phenotyping, deep learning, machine learning, bioinformatics, data science, omics data,

### 要旨

ビッグデータであるパーソナルデータ等の利活用の観点から、データ利活用ルールの策定・整備が急速に進められている。このような背景を踏まえ、ビッグデータを用いた研究の実現に向けて必要な基盤や環境として、以下の3点が重要である。1) deep phenotyping に最適化した医療ビッグデータの整備。2) ゲノムデータ及びオミックスデータの利活用。3) データサイエンスによる新しい作業仮説の創出。この戦略を支えるのが、臨床データベースの拡充とバイオバンク・ジャパンネットワークの連携強化、並びに、パイオインフォマティクスデータの活用と機械学習・深層学習等のデータサイエンスの基盤の整備である。この研究アプローチに基づいて新たに組織された研究体制は、臨床疫学研究からヒトの生物学研究まで広範囲の研究ニーズに対応し、疾患メカニズムの発見や個別化医療の樹立、創薬に貢献するものと期待される。

### 1. はじめに

国は個人情報及びプライバシーの保護に配慮したデータ利活用の推進に向けて、個人情報保護法の改正等（令和2年改正及び令和3年改正、令和4年4月1日施行）を含めて、政令・規則・ガイドライン等の必要なルールの整備を進めている。このような背景を基にビッグデータであるパーソナルデータの活用へ向けた活動が促進することが見込まれる。この潮流は研究開発領域にも波及し、オープンサイエンスの促進に向けた取り組みとして公的資金により得られた研究データの管理・利活用の推進が挙げられている。先駆事例として国立研究開発法人日本医療研究開発機（AMED）が策定したデータシェアリングポリシーが示すようにゲノム情報等は原則としてデータシェアリング

が求められることから、近い将来に国はオミックスデータを統合するような公的データベースを整備して公開することが見込まれる。このような状況を踏まえて、医療及びゲノム情報等のビッグデータを利活用して国際的価値の向上に資する研究を本学において推進する際、どのような研究戦略を打ち立てればよいだろうか。本稿では、医療データの利活用に向けて我々が取り組んでいる研究スキームの一例を紹介する。その趣旨は、過去18年間に渡る膨大な医療ビッグデータと「オーダーメイド医療実現化プロジェクト」のゲノム・オミックスデータを最大限に活用しつつ、最近のトレンドであるデータサイエンスを用いて新たな知見の創造を図るものであり、この研究目的を達成するための新しい研究基盤及び戦略的な運営を実行

<sup>1)</sup> 日本大学医学部臨床試験研究センター

<sup>2)</sup> 日本大学医学部生体機能医学系薬理学分野  
高橋泰夫 : takahashi.yasuo@nihon-u.ac.jp

する研究体制について概観した。

## 2. 日本大学臨床データベースとバイオバンク・ジャパン (BBJ) との連携

日本大学医学部は医療情報の研究応用と産業化への有効利用を推進するとともに、臨床疫学・医学情報学分野の研究促進のため、日本大学医学部臨床データベース (Nihon University School of Medicine's Clinical Data Warehouse: NUSM's CDW) (以下「本学臨床 DB」という。) を構築し 2004 年から運用を開始した。本システムは病院情報システムの医事情報、検査情報、薬剤オーダーリング情報等を定期的に収集・統合し、個人情報保護のためにデータを匿名加工情報に変換して格納している<sup>1)</sup>。本学臨床 DB はデータクリーニングを含む管理・運用等の長年の知識の蓄積により、研究に利活用するリアルワールドデータを最適な仕様で提供する環境を整備しており、疾患および治療の実態調査や薬剤の有効性・安全性調査、疫学研究等の臨床研究に、臨床医学系分野の研究者や製薬企業によって実際に利活用されています。また、2003 年に文部科学省のリーディングプロジェクトとして開始された BBJ「オーダーメイド医療実現化プロジェクト」(2003 年～2018 年) への協力参加にあたり、医療情報の標準化とデータの電子化収集、及び管理・運用、情報セキュリティー等の業務に関しても、本学臨床 DB は当該プロジェクトの重要な研究基盤として有効に活用されています。BBJ は、「オーダーメイド医療実現化プロジェクト」が終了した後も、診療情報を追加収集・解析することを目的とする BBJ ネットワークを立ち上げた<sup>2)</sup>。本研究ネットワークを通じて、本学と BBJ との間で双方向の情報提供の環境が整備された。このような環境の変化を踏まえ、BBJ のジェノタイプデータと本学臨床 DB の連携を強力に推進することにより、新しい研究領域への利活用の可能性を検討する。例えば、日本大学が BBJ に提供した約 1 万 9 千人の DNA は SNP (Single Nucleotide Polymorphism) データとしてマッチングが可能となり、本学の臨床 DB を用いて臨床情報を詳細に解析し、その情報を用いて BBJ の SNP データを解析することで、より複雑な疾患リスクや予後、薬剤感受性に関わるバリエーションやマ

ーカーを探索する研究が可能となる。すなわち、これまでの臨床 DB の利活用は臨床疫学研究が主流であったが、今後は疾患メカニズムの発見等、ヒトの生物学研究への利活用の選択肢が増えることになる。これにより時代のニーズの変化に柔軟に合わせて研究テーマの選択が可能となり、本学の臨床 DB は研究戦略構築の際の大きな強み (Strength) となることが期待される。

## 3. ビッグデータの活用を推進するための国内環境の進展

ヘルスケア分野における医療ビッグデータの利活用が着目されるようになって数年が経過したが、データの二次利用に関する社会的な理解が醸成されておらず、かつ、プライバシーに関する強い懸念等から日本ではデータの整備、データの利用環境の整備及び実際のデータの利活用は十分に進んでいない。このような状況を踏まえ、政府はデジタル基盤を構築するため包括的なデータ戦略を策定し、パーソナルデータ利活用の面から個人情報保護法の改正等 (令和 2 年改正及び令和 3 年改正) の施行に伴い、令和 4 年 4 月 1 日からオプトアウト提供できる個人データの範囲や届出事項等の変更等、政令・規則・ガイドライン等の必要なルールを整備した。今回の個人情報保護法の改正を受けて、人を対象とする生命科学・医学系研究に関する倫理指針も改正され (令和 4 年 4 月 1 日施行)、インフォームド・コンセント手続 (情報の取得・利用・提供) が例外要件ごとに規定された。また、仮名加工情報や個人関連情報の取扱い、第三者への研究情報提供等の手続も追加され、越境データの流通増大に伴う新たなリスクへの対応、並びに、AI・ビッグデータ時代への対応等、円滑なデータの流通・共有を促進するための制度の整備が進められている。また、公的資金により得られた研究データの管理・利活用についてはオープン・アンド・クローズ戦略を踏まえて、研究開発を行う機関は 2025 年までにデータポリシーを策定することが進められている。最近の潮流はオープンデータ (公共データを機械判読可能な形式で二次利用可能なルールの下公開すること) を推進する観点から、適切なエンバーゴ (時限付き非公開) 期間を設定して原則公開し利活用することが求められる。AMED は 2016 年度にデータシェアリング

ポリシーを定め、オーダーメイド医療実現化プロジェクト等の一部の事業ではデータシェアリングを原則として義務付けている。このデータシェアリングポリシーに基づいて、BBJ は約 18 万人の SNP データ、約 3000 人の全ゲノムシーケンス解析の結果を NBDC (National Bioscience Database Center) を通じて制限公開している。

#### 4. 医療ビッグデータを用いた研究戦略の方向性

本学臨床 DB の効果的かつ効率的な活用促進へ向けて、新たな研究シーズを創造、並びに、新たな価値の創造を目標として、以下の方向で研究の加速化を図る。

##### 1) ゲノム・オミックスデータの統合

日本大学が BBJ に提供した約 1 万 9 千人の DNA は SNP データとして制限公開されており利用可能である。また、BBJ はナイチンゲールヘルスジャパン株式会社と共同で最大 65 万の血清検体を対象に大規模メタボローム解析を実施することをプレスリリース (令和 3 年 10 月) した。BBJ は AMED のデータシェアリングポリシーの対象事業であるため、この共同事業の結果は原則データシェアリングされることが決まっており、ゲノムデータと同様に日本大学が提供した血清サンプルのメタボローム解析の結果は数年以内に利用可能となることが十分に見込まれる。また、web 上で公開されている Biotechnology 情報 (DrugBank, GWAS Catalogue, HaploReg, NCBI, EBI, DDBJ, KEGG 等) をバイオインフォマティクスデータとして集約し、本学の臨床 DB へ統合する。これにより genetic variants と clinical phenotype の相関を解析する(図 1)。ここで表現型と遺伝子型のつ

ながりを解釈するものとしてバイオインフォマティクスが重要である。

##### 2) 深い表現型 (deep phenotyping)

本学の臨床 DB を用いて詳細に表現型の解析を行い、特定の病態を呈する集団や特定の薬物に反応が期待される集団(subpopulation)に分類し、より複雑な疾患リスクや予後、薬剤感受性に関わるバリエーション等の解明に取り組む。これにより個別化医療の実装を目指す。

##### 3) データサイエンスによる新しい作業仮説の創出

deep phenotyping とゲノム・オミックスデータ、バイオインフォマティクスデータを集約し、機械学習及び深層学習を用いて解析することによって新たな作業仮説を生み出し、新規の疾患概念やメカニズム等を明らかにすることを目指す。さらに、科学性が担保され得る質の高い開発研究を目指し、新たな創造された知見を検証する研究を提案し、基礎研究 (リバーストランスレーショナル研究) 又は臨床研究 (トランスレーショナル研究) への橋渡しを実施する (図 2)。

このように、既存のリソース、すなわち、18 年間の膨大な臨床情報を収集した本学臨床 DB 及び「オーダーメイド医療実現化プロジェクト」のゲノムデータや関連するオミックスデータを最大限に活用しつつ、データサイエンス等を中心とするテクノロジーのトレンドを踏まえ、研究開発を促進するための新たな研究プラットフォームの構築を進めている。本研究方針に基づき新たな研究体制が整備され短期間であるが、いくつかの研究において成果を公表した<sup>3)5)</sup>。

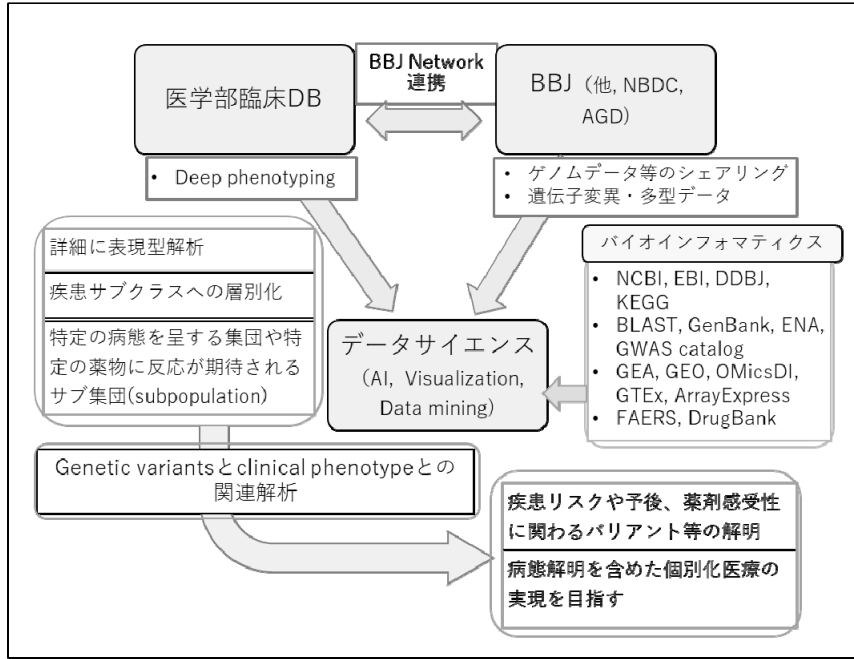


図 1. ビッグデータを活用した研究アプローチ

データサイエンスを用いて解析したヒトの表現型とゲノム及びオミックスデータとの関連等から疾患リスクや予後、薬剤感受性に関わるバリエーション等を解明する研究を推進し、疾患メカニズムの発見や個別化医療の実装を目指す。

AI, Artificial intelligence; AGD, AMED Genome group sharing Database; BBJ, Biobank Japan; DB, Database; DDBJ, DNA Data Bank of Japan; EBI, European Bioinformatics Institute; KEGG, Kyoto Encyclopedia of Genes and Genomes; NBDC, National Bioscience Database Center; NCBI, National Center of Biotechnology Information.

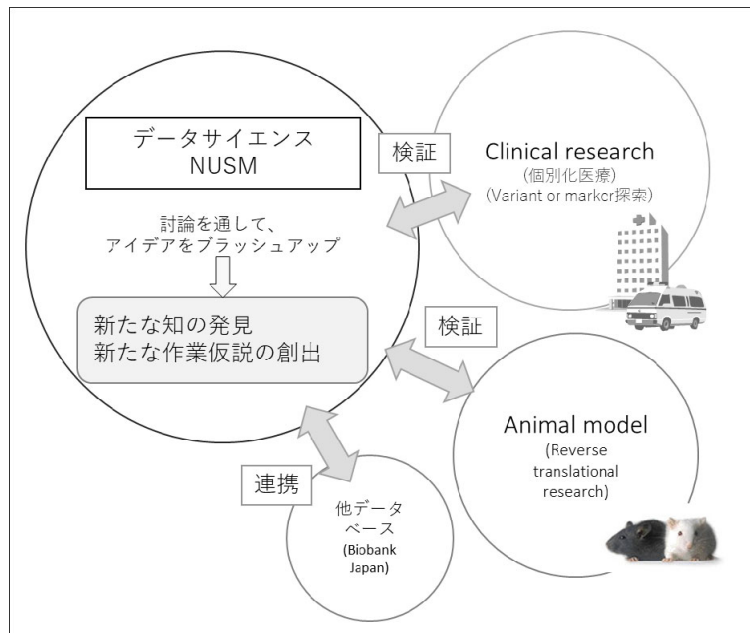


図 2 データサイエンスを用いた研究スキーム

データサイエンスを用いて解析した結果をブラッシュアップし、新たな作業仮説・知見を創出する。さらに検証研究を支援し、基礎研究 (リバーストランスレーショナル研究) 又は臨床研究 (トランスレーショナル研究) への橋渡しを実施する。NUSM, Nihon University School of Medicine.

## 5. 結語

本稿では、ビッグデータを用いた研究戦略の方向性について述べた。その骨子はヒトの表現型解析と様々なオミックスデータ等を含むバイオインフォマティクスデータを統合し、これをデータサイエンスで解析することであり、臨床疫学研究からヒトの生物学研究まで広範囲の研究ニーズに対応し、疾患メカニズムの発見や個別化医療の実装、創薬に貢献するものと期待できる。

## 6. 文献

- 1) Takahashi Y, Nishida Y, Asai S. Utilization of health care databases for pharmacoepidemiology. *Eur J Clin Pharmacol*. 2012;68(2):123-9.
  - 2) 高橋泰夫, 長島卓也, 浅井聡. 日本大学医学部 220 万人 15 年間の医療ビッグデータを基にした臨床薬理学的研. *Jpn J Clin Pharmacol Ther*. 2021;52(2):51-54
  - 3) Akimoto H, Nagashima T, Minagawa K, Hayakawa T, Takahashi Y, Asai S. Signal Detection of Potential Hepatotoxic Drugs: Case-Control Study Using Both a Spontaneous Reporting System and Electronic Medical Records. *Biol Pharm Bull*. 2021;44(10):1514-23.
  - 4) Takahashi Y, Yamazaki K, Kamatani Y, Kubo M, Matsuda K, Asai S. A genome-wide association study identifies a novel candidate locus at the DLGAP1 gene with susceptibility to resistant hypertension in the Japanese population. *Sci Rep*. 2021;11(1):19497.
  - 5) Nagashima T, Hayakawa T, Akimoto H, Minagawa K, Takahashi Y, Asai S. Identifying Antidepressants Less Likely to Cause Hyponatremia: Triangulation of Retrospective Cohort, Disproportionality, and Pharmacodynamic Studies. *Clin Pharmacol Ther*. 2022;111(6):1258-67.
- URLs
- 6) BioBank Japan  
[https:// bioba nkjp. org/ en/ index. html;](https://biobank.jp/en/index.html)
  - 7) DDBJ  
[https://www.ddbj.nig.ac.jp/index.html;](https://www.ddbj.nig.ac.jp/index.html)
  - 8) DrugBank  
[https://go.drugbank.com;](https://go.drugbank.com)
  - 9) EBI  
[https://www.ebi.ac.uk;](https://www.ebi.ac.uk)
  - 10) GWAS Catalog  
[https:// www. ebi. ac. uk/ gwas/;](https://www.ebi.ac.uk/gwas/)
  - 11) HaploReg  
[http:// www. broad insti tute. org/ mamma ls/ haplo reg/ haplo reg. php;](http://www.broadinstitute.org/mammas/haploreg/haploreg.php)
  - 12) KEGG  
[https://www.genome.jp/kegg/kegg\\_ja.h tml;](https://www.genome.jp/kegg/kegg_ja.html)
  - 13) NBDC  
[https://humandbs.biosciencedbc.jp;](https://humandbs.biosciencedbc.jp)
  - 14) NCBI  
[https://www.ncbi.nlm.nih.gov;](https://www.ncbi.nlm.nih.gov)



# 機械学習を活用した薬物性肝障害のリスクに影響を及ぼす 薬物間相互作用の検出

秋元勇人<sup>1)2)</sup>, 長島卓也<sup>1)2)</sup>, 皆川貴美乃<sup>2)</sup>, 早川 隆<sup>1)2)</sup>,  
高橋泰夫<sup>2)</sup>, 浅井 聡<sup>1)2)</sup>

## Detection of drug-drug interaction on drug-induced liver injury using supervised machine-learning algorithms

Hayato AKIMOTO<sup>1)2)</sup>, Takuya NAGASHIMA<sup>1)2)</sup>, Kimino MINAGAWA<sup>2)</sup>,  
Takashi HAYAKAWA<sup>1)2)</sup>, Yasuo TAKAHASHI<sup>2)</sup>, Satoshi ASAI<sup>1)2)</sup>

### 要旨

薬物性肝障害(DILI)は薬物性有害事象の1つであり、特定の2種薬物の組み合わせがDILIの報告頻度を増加させることが複数の臨床研究において報告されている。これらの研究では伝統的な統計解析手法である多重ロジスティック回帰分析(MLR)が利用されている。しかしながらこの解析手法はデータに過剰適合することが懸念される。そこで本研究では、汎化性能の向上が期待できる機械学習を利用し、肝逸脱酵素の1種である血清アラニンアミノ基転移酵素(ALT)の異常な増加のリスクに複合的な影響を及ぼす2種薬物の組み合わせを検出することを目的とした。まず、日本大学医学部臨床情報データベース(NUSM's CDW)から58,413人分の患者情報を抽出した。この情報を訓練データとテストデータにランダムに分割し、ALTの異常な増加のリスクに影響する薬物の組み合わせを次の3つの機械学習アルゴリズムを用いて検出した: MLRモデル、ロジスティックラッソ(LASSO)回帰モデル、勾配ブースティング決定木(XGBoost)モデル。まず、MLRモデルは訓練データに対し過剰適合していた。ロジスティックラッソ回帰モデルでは、ジクロフェナクとファモチジンの組み合わせはALTの異常な増加のリスクの上昇と関連していた。加えて、XGBoostモデルにおいてもジクロフェナクとファモチジンの組み合わせはALTの異常な増加を予測するための重要な因子であった。

### 1. はじめに

薬物性肝障害(DILI)は代表的な薬物有害事象の1つであり、急性肝不全の原因の1つでもある。個々の薬物あるいは薬物の治療クラス(抗菌薬等)のDILIリスクとの関連は世界中の臨床研究で調査されており、著者らにおいても日本大学臨床情報データベース(NUSM's CDW)を用いて調査してきた。<sup>1-3</sup> これらの臨床研究ではHy's lawに適合した患者をDILI患者(Hy's law cases)とみなしている。Hy's law ケースは一般に血清アラニンアミノ基転移酵素(ALT)が基準値上限の3倍を超え、かつ、血清総ビリルビン(TBL)が基準値上限の2倍を超える患者である。しかしながら、Hy's law ケースは非常に稀な症例であり、DILIを誘発する薬物はTBLの上昇を伴わずにALTの基準値上限の3倍

を超える上昇を引き起こすことが多い。<sup>4</sup> したがって、DILIの潜在的なリスク要因を評価する際には血清ALTの異常な増加に注目することが重要である。

2004年、イギリス人を対象とした臨床研究において2種類以上の肝毒性薬物の使用では、1種類のみと比較し、急性肝不全のリスクが6倍高いことが示唆され、これを契機に2種類の薬物の併用がDILIの報告頻度を変化させるかを調査した臨床研究が複数報告されている。<sup>5-7</sup> これらの研究の解析手法は伝統的な解析手法である多重ロジスティック回帰分析(MLR)が採用されている。MLRモデルは性別や年齢等の共変量を調整しつつ薬物とDILIリスクとの関連を評価することが可能であるが、

<sup>1)</sup> 日本大学医学部生体機能医学系薬理学分野

<sup>2)</sup> 日本大学医学部臨床試験研究センター

秋元勇人: akimoto.hayato@nihon-u.ac.jp

解析に供したデータに過剰適合 (over-fitting) することが知られている。<sup>8</sup> 機械学習は大規模データにおいて複数の変数間の複雑な関係を取り扱うことが可能であり、従来の統計解析手法の代替として応用できる。実際、機械学習アルゴリズムは DILI の予測に使用されている。<sup>9,10</sup> そこで本研究では、機械学習アルゴリズムを用いて特定の 2 種薬物間に DILI のリスクを大幅に増加させる相互作用があるかを検出することを目的とする。

## 2. 対象及び方法

### 2. 1. データソース

本研究は 2004 年 4 月から 2021 年 7 月までに集積された日本大学臨床情報データベース (NUSM' s CDW) を活用した症例対照研究であり、データマイニング指向の研究でもある。NUSM' s CDW は一元化されたデータリポジトリであり、次の 4 つのデータファイルから構成される：患者基本情報、診断情報、検査値情報、投薬情報。

### 2. 2. 解析対象患者および 2 値アウトカム

まず、NUSM' s CDW より 90 日以内に少なくとも 3 回以上肝機能検査を実施し、初回および 2 回目の ALT 値が基準値範囲内 (4.0-44.0 U/L) であった患者 122,285 名を抽出した。これら患者のうち、最後の測定日 (指標日) の ALT 値が基準値上限の 3 倍 (>132.0 U/L) を超えた患者を症例群 (outcome=1; case group) に割り当てた。他方、指標日における ALT 値もまた基準値範囲内であった患者は対照群 (outcome=0; control group) とみなした。次に、症例群および対照群の患者のうち、下記の除外基準のいずれかを満たす患者を除外した。最終的に 58,413 名分の患者情報が解析対象となった。

- 1) 18 歳未満の患者
- 2) 初回および指標日の肝機能検査値に欠損がある患者
- 3) 薬物に起因しない肝障害の既往がある患者
- 4) 何らかの肝保護薬を使用している患者

### 2. 3. 特徴量

本研究では全 1,375 特徴量が機械学習に供された：6 つの患者基本情報、5 つの診断歴、180 種類

の薬物の使用状況 (2 値データ)、これらの薬物で構成される 1,050 の積項 (二次元交互作用項)、134 の薬物治療クラス。6 つの基本情報は年齢、性別、併用薬剤数、NUSM' s CDW を構成する 3 病院 (日本大学医学部付属板橋病院、旧日本大学医学部付属練馬光が丘病院、旧駿河台日本大学病院) で構成された。診断歴は急性肝不全のリスク要因として知られている高血圧、糖尿病、脂質異常症、心不全、敗血症の 5 つで構成された。180 種類の薬物の使用状況とは、指標日前 90 日以内に新たに処方が開始され、かつ、その処方が指標日以降まで継続していた薬物を DILI の被疑薬とみなし、その被疑薬の使用の有無を指す。すなわち、症例群および対照群において 180 種類の薬物が対象期間内に高頻度 (100 名以上の患者)、かつ、新たに処方が開始されたことを示す。1,050 の積項とは、これら 180 種類の使用薬物のうち、高頻度で併用されていた薬物の組み合わせに注目し、その 2 種類の薬物の使用状況 (2 値データ; 0 or 1) の積を算出したものである。134 の薬物治療クラスに関しては指標日前 90 日以内に新たに処方が開始されていたが、その使用が低頻度であった薬物 909 種類を世界保健機構が公開している解剖治療化学分類法 (ATC system) の 4<sup>th</sup> level を用いて化学・治療法・薬学サブグループへ分類し、その後 134 の薬物治療クラスへ割り当てた。この 134 の薬物治療クラスは、そのクラスに該当する薬物のいずれかを使用していれば 1、そうでなければ 0 とし、特徴量として取り扱った。最終的に 58,413 行×1,375 列の二次元データセットが生成され、機械学習に供した。

### 2. 4. 機械学習モデルの構築とモデル評価

まず、機械学習に供する二次元データセットを無作為に訓練データ (70%; N=40,889) とテストデータ (30%; N=17,524) に分割した。血清 ALT の異常な増加のリスクへの個々の薬物およびそれら積項の影響を評価するために、訓練データを用いて 3 つの機械学習モデルを構築した：MLR モデル、ロジスティックラッソ (least absolute shrinkage and selection operator; LASSO) 回帰モデル、勾配ブースティング (extreme gradient boosting; XGBoost) 決定木モデル。いずれの機械学習モデルにおいても応答変数は ALT の異常な増加の有無であり、説明変数は先に示した 1,375 特徴量とした。ロジスティック LASSO 回帰モデルおよび XGBoost

モデルを構築する際、未知データへの予測性能を向上させるために 10 分割交差検証を実施した。ロジスティック LASSO 回帰モデルにおけるハイパーパラメータはラムダ( $\lambda$ )であり、10 分割交差検証によって誤分類率が最小となる  $\lambda_{\min}$  を用いてロジスティック LASSO 回帰モデルを構築した。XGBoost モデルのハイパーパラメータは、学習率を 0.1 に固定後 10 分割交差検証により最適な木の数 (nrounds) を決定し、学習率及び nrounds を固定した状態でその他のチューニングパラメータをグリッドサーチすることにより決定した。

各機械学習モデルの予測パフォーマンスを評価するために、受信者操作特性曲線の曲線下面積 (AUROC) を算出後、DeLong テストでアルゴリズム間の比較をした。更に、本研究では症例群と対照群の患者数に大きな偏りがあるため、precision-recall 曲線下面積 (AUPR) も算出した。

## 2. 5. ALT の異常な増加リスクの上昇に関連する 2 種薬物間の相互作用の検出

特定の 2 種薬物の併用が ALT の異常な増加リスクの上昇に関連するかを評価するための準備として、ロジスティック LASSO 回帰モデルの偏回帰係数から各特徴量の調整オッズ比 (aOR) を算出した。次にこの aOR の 95% 信頼区間を見積もるためにブートストラップパーセンタイル法を用いた。重複を許すリサンプリングを実行することで計 2,000 個のブートストラップ標本を訓練データから生成した。これらブートストラップ標本それぞれに対し、上記と同様の手順でロジスティック LASSO 回帰分析を実行し、1 特徴量に対し 2,000 個のパラメータ推定値を得た後、昇順に並べたパラメータ推定値の 50 番目から 1950 番目までの間を 95% ブートストラップパーセンタイル信頼区間 (95% BootCI) とした。

## 2. 6. 統計解析

症例群および対照群間の患者背景の比較には、データの分布状況に応じて、連続尺度に対してはウェルチの t 検定もしくはウィルコクソンの順位和検定を実行し、分類尺度に関しては  $\chi^2$  検定を実

行した。統計解析の有意水準を 5% とし、すべての統計解析には R ソフトウェア (Ver. 4.0.2) を用いた。

## 3. 結果

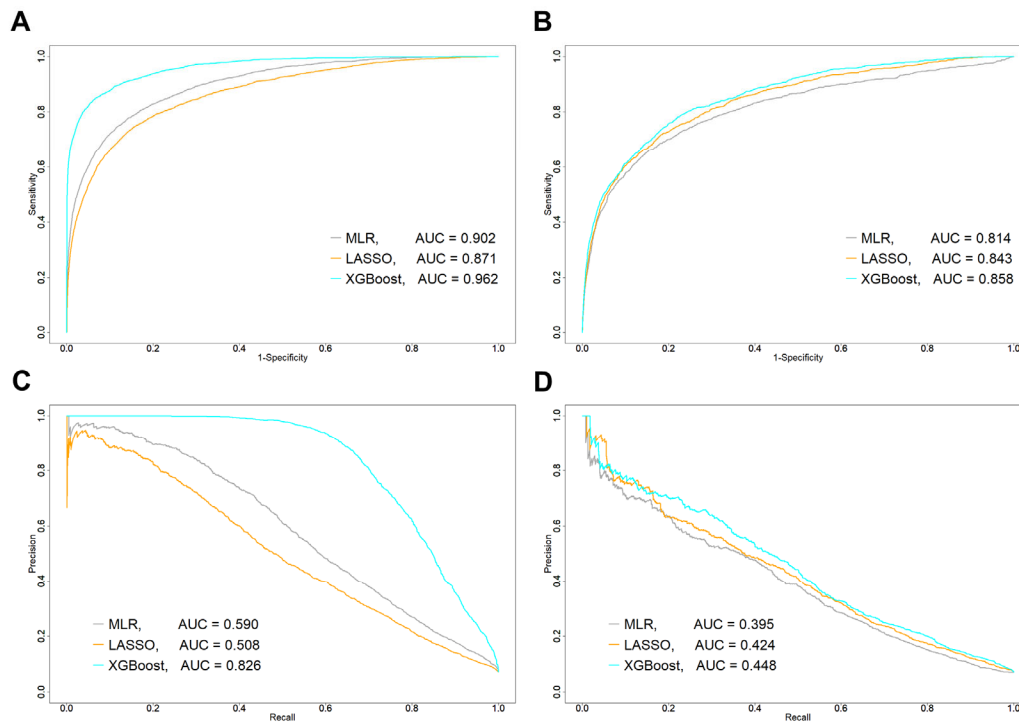
全 58,413 名分の患者情報が NUSM' s CDW から抽出され、これらの患者を症例群 (N=4,152) と対照群 (N=54,261) に割り当てた。Table 1 には 2 群間の患者背景の比較を示す。症例群および対照群の年齢はそれぞれ、平均 63.8 歳 (標準偏差 16.6 歳) および 57.0 歳 (19.6 歳) であり、症例群の方が高齢であった。また、男性の割合や併用薬剤数に関しても症例群の方が高かった。これら以外の患者背景についてはすべての項目に関して症例群と対照群で有意な違いが認められた ( $P < 0.001$ )。

次に、各機械学習アルゴリズムの予測パフォーマンスを示す。訓練データにおいて、XGBoost モデル (0.962; 95%CI, 0.958-0.965) > MLR モデル (0.902; 0.896-0.908) > ロジスティック LASSO 回帰モデル (0.871; 0.864-0.878) の順に AUROC は高かった (Figure 1A; いずれも  $P < 0.001$ )。しかしながら、テストデータ (未知のデータ) に対する AUROC は降順で、XGBoost モデル (0.858; 0.846-0.868) > ロジスティック LASSO 回帰モデル (0.843; 0.830-0.854) > MLR モデル (0.814; 0.798-0.828) であった (Figure 1B; いずれも  $P < 0.001$ )。AUPR に関しても同様に、XGBoost は訓練・テストデータに関わらず最も高値であり、MLR モデルは訓練データに関してはロジスティック LASSO 回帰モデルよりも高値であるが、テストデータへの適合は低かった。すなわち、MLR モデルは訓練データに対しては LASSO 回帰モデルよりも AUROC および AUPR が高値であるが (Figure 1C)、未知のデータに対しては一転して低値であった (Figure 1D)。このことから MLR モデルは訓練データに過剰適合している。

**Table 1. Comparison of patient characteristics between case and control groups.**

Characteristics	Case group (N=4,152)	Control group (N=54,261)	P value
Age (years), mean (SD)	63.8 (16.6)	57.0 (19.6)	<0.001
Male, n (%)	2,639 (63.6)	21,941 (40.4)	<0.001
Number of concomitant drugs, median (IQR)	7.0 (4.0-10.0)	4.0 (2.0-6.0)	<0.001
Hospital, n (%)			<0.001
Itabashi	3,425 (82.5)	42,311 (78.0)	
Hikarigaoka	370 (8.9)	6,107 (11.3)	
Surugadai	357 (8.6)	5,843 (10.8)	
Medical history, n (%)			
Hypertension	1,203 (29.0)	9,399 (17.3)	<0.001
Diabetes	1,412 (34.0)	15,035 (27.7)	<0.001
Dyslipidemia	737 (17.8)	6,930 (12.8)	<0.001
Heart failure	871 (21.0)	5,623 (10.4)	<0.001
Sepsis	319 (7.7)	616 (1.1)	<0.001
Liver function tests, median (IQR)			
ALT (U/L)	182.0 (150.0-276.2)	14.0 (11.0-20.0)	<0.001
AST (U/L)	171.0 (105.0-342.0)	19.0 (15.0-23.0)	<0.001
TBL (mg/dL)	0.8 (0.5-1.5)	0.5 (0.4-0.7)	<0.001
ALP (U/L)	361.0 (239.0-618.0)	209.0 (165.0-264.0)	<0.001

Unpaired two-tailed Welch's *t*-test was performed for differences in age. Wilcoxon rank-sum test was performed for differences in number of concomitant drugs and liver function test values. Chi-squared test was performed for differences in categorical variables such as sex, hospital, and medical history. Abbreviations: ALP, alkaline phosphatase; ALT, alanine aminotransferase; AST, aspartate aminotransferase; IQR, interquartile range; SD, standard deviation; TBL, total bilirubin; U/L, units per liter.

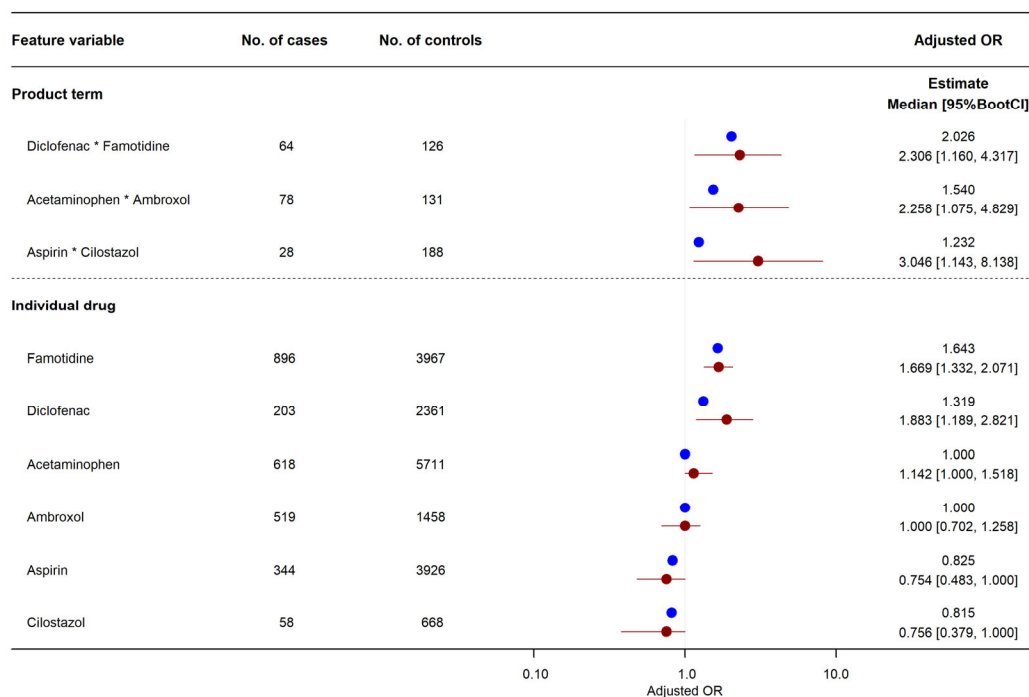


**Figure 1. Comparison of predictive performance of each machine-learning algorithm between training and testing sets.**

(A) Area under receiver operating characteristic curve (AUROC) in training set and (B) testing set. (C) Area under precision-recall curve (AUPR) in training set and (D) testing set. Abbreviations: AUC, area under the curve; LASSO, least absolute shrinkage and selection operator; MLR, multiple logistic regression; ROC, receiver operating characteristic curve; XGBoost, extreme gradient boosting.

ロジスティック LASSO 回帰モデルにおいて、特徴量として扱われた 1,050 の積項のうち、6 種の薬物による 3 つの積項は ALT の異常な増加リスクの上昇と有意に関連していた。3 つの積項およびこれらの積項を構成する 6 種薬物の ALT の異常な増加リスクへの影響を Figure 2 に示す。3 つの積項のうち、最もリスクの上昇と関連していたのはジクロフェナクとファモチジンの組み合わせであった (aOR, 2.026; 95%BootCI, 1.160-4.317)。次いで、アセトアミノフェンとアンブロキシール

(1.540; 1.075-4.829)、アスピリンとシロスタゾール(1.232; 1.143-8.138)であった。ジクロフェナクとファモチジンに関しては、積項のみならず、個々の薬物においても ALT の異常な増加リスクの上昇と有意に関連していた：それぞれ、aOR, 1.319 (95%BootCI, 1.189-2.821) および 1.643 (1.332-2.071)。このことから、ジクロフェナクとファモチジンを併用した際の ALT の異常な増加に関する aOR は 4.389 (点推定値; 個々の薬物の効果および積項の積)であった。

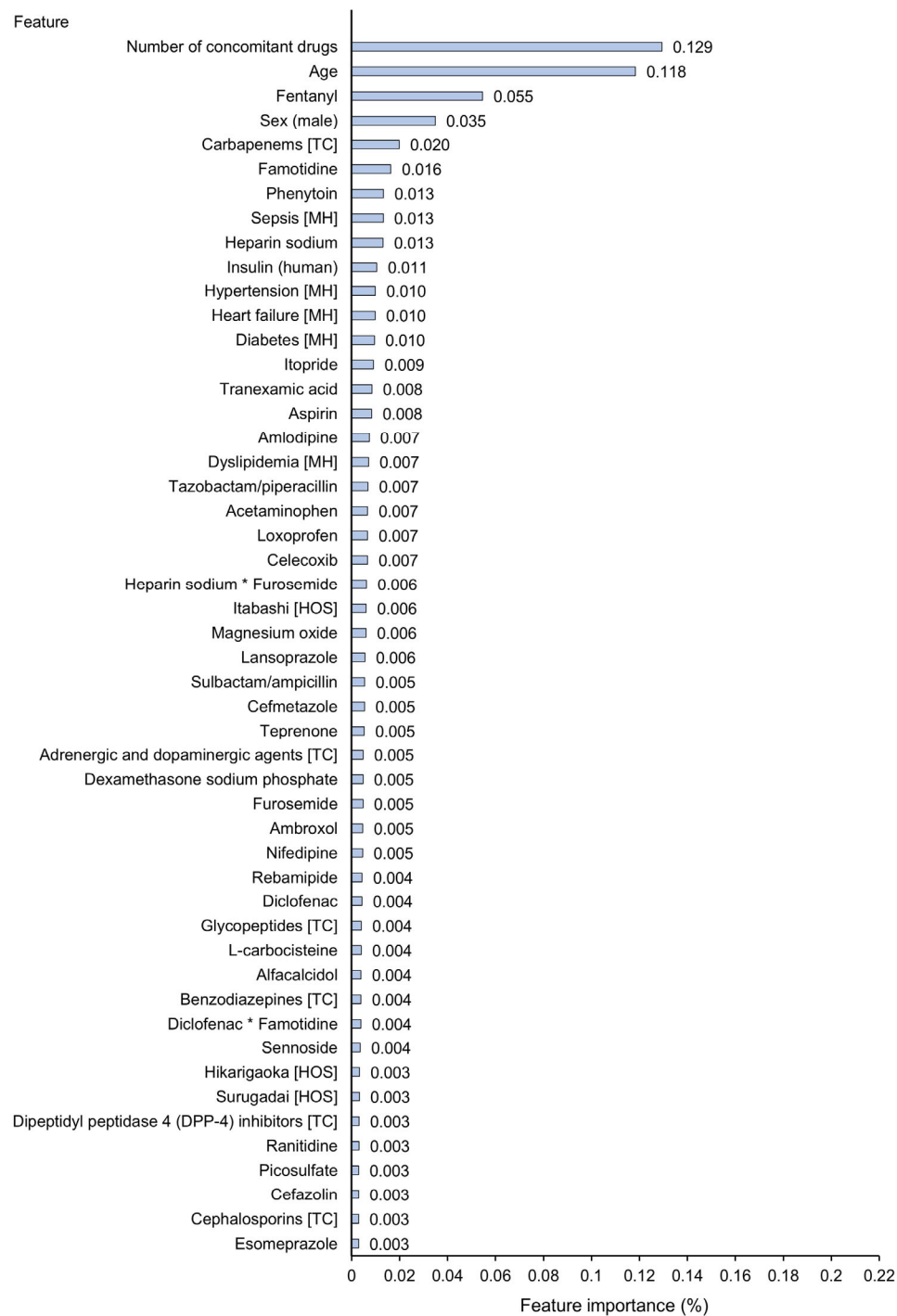


**Figure 2. Effects of individual drugs and their product terms on risk of abnormal elevation of ALT in logistic least absolute shrinkage and selection operator (LASSO) regression model.**

Three product terms had a lower limit of adjusted odds ratio 95%BootCI >1. Blue and red circles represent estimated adjusted odds ratio in original training set and median of adjusted odds ratio in 2,000 bootstrap replicates, respectively. Red horizon indicates adjusted odds ratio 95%BootCI. Abbreviations: 95%BootCI, 95% bootstrap percentile confidence interval; OR, odds ratio.

XGBoost モデルにおいて、ある患者が症例群か対照群のどちらであるかを予測するために重要な特徴量、すなわち特徴量重要度(feature importance)を全 1,375 特徴量の中から上位 50 個に注目し Figure 3 に示す。この上位 50 の特徴量重要度において積項は 2 つ存在していた：ヘパリンナトリウムとフロセミド(feature importance,

0.006)、ジクロフェナクとファモチジン(0.004)。更には、これら 2 つの積項を構成している 4 種薬物の個々の効果もまた上位 50 にすべて含まれていた：ヘパリンナトリウム(0.013)、フロセミド(0.005)、ジクロフェナク(0.004)、ファモチジン(0.016)。



**Figure 3. Top 50 important features for prediction of abnormal elevation of alanine aminotransferase in extreme gradient boosting (XGBoost) tree model.**

The first column includes the top 50 important features of all the features that were actually used in the XGBoost tree model. Feature importance (Gain) score indicates how important each feature was in the construction of boosted decision trees within the XGBoost model. A higher value of this metric when compared to another feature suggests that it is relatively more important for construction of a predictive model. The sum of all importance scores included in a trained XGBoost tree model is 1. The symbol \* indicates a product term of two drugs. Abbreviations: HOS, hospital; MH, medical history; TC, therapeutic class.

#### 4. 考察

本研究では、患者の医療情報が豊富に蓄積されている NUSM's CDW を用いることにより、従来の多変量解析手法である MLR はデータに対して過剰適合を引き起こし得ること、近年注目されている機械学習アルゴリズムはデータへの過剰適合を回避しつつイベントの予測に寄与している要因の探索にも有用であることを示した。MLR と同様、ロジスティック LASSO 回帰モデルは線形回帰ベースのアルゴリズムである。一方で、XGBoost モデルは線形回帰ではなく、決定木ベースのアルゴリズムを用いている。本研究ではこれらの異なる2つのアルゴリズム間で共通して、ジクロフェナクおよびファモチジンの個々の影響だけでなく、これら薬物間の交互作用項もまた ALT の異常な増加リスクの上昇に関連することが示唆された。

米国食品医薬品局が公表している the Drug Induced Liver Injury Rank (DILIRank) dataset によれば、ジクロフェナクは最も強く DILI が懸念される薬物群 (Most-DILI-concern drug category) に属しており、このカテゴリに属する薬物は DILI のリスクが高く、DILI の原因薬物としても同定されている。<sup>11</sup> 本邦においてもジクロフェナクは重篤な肝障害患者に対しては禁忌であり、このことはジクロフェナクの個々の効果として本研究においても検出された。ファモチジンは DILI の懸念が小さい薬物群 (Less-DILI-concern drug category) に属しているが、ファモチジンは DILI の原因薬物であるとみなされており、添付文書の警告欄にも DILI に関する記載がある。実際、本研究ではファモチジンは ALT の異常な増加リスクの上昇に有意に関連することが示されている。本研究で新たに得られた知見として、ジクロフェナクとファモチジンを近い時期に併用すると ALT の異常な増加リスクが個々の薬物の影響の和よりも大きくなる、すなわちこの2種薬物間には DILI のリスクに対して薬物間相互作用が存在する可能

性が示唆された。

#### 5. 結語

本研究では、患者の医療情報であるリアルワールドデータと人工知能技術の1つである機械学習を組み合わせることにより、膨大な情報の中から注目する副作用に影響を及ぼし得る薬物間相互作用の検出を試みた。今後さらなる検討が必要ではあるものの、本研究で新たに得られた知見として、ジクロフェナクとファモチジンの併用はこれらの薬物の個々の効果の和よりも大きく ALT の異常な増加のリスクを上昇させる可能性が示唆された。

#### 6. 文献

- 1) Björnsson E, Olsson R. Outcome and prognostic markers in severe drug-induced liver disease. *Hepatology* 2005; 42: 481–489.
- 2) Chalasani N, Bonkovsky HL, Fontana R, et al. Features and outcomes of 899 patients with drug-induced liver injury: The DILIN prospective study. *Gastroenterology* 2015; 148: 1340–1352.
- 3) Akimoto H, Nagashima T, Minagawa K, et al. Signal detection of potential hepatotoxic drugs: Case-control study using both a spontaneous reporting system and electronic medical records. *Biol Pharm Bull* 2021; 44: 1514–1523.
- 4) Church RJ, Watkins PB. In silico modeling to optimize interpretation of liver safety biomarkers in clinical trials. *Exp Biol Med* 2018; 243: 300–307.
- 5) de Abajo FJ, Montero D, Madurga M, et al. Acute and clinically relevant drug-induced liver injury: A population based case-control study. *Br J Clin Pharmacol* 2004; 58: 71–80.
- 6) Suzuki A, Yuen NA, Ilic K., et al. Comedications alter drug-induced liver injury reporting frequency:



- Data mining in the WHO VigiBase™. *Regul Toxicol Pharmacol* 2015; 72: 481–490.
- 7) Yazici C, Mutlu E, Bonkovsky, HL, et al. Risk factors for severe or fatal drug-induced liver injury from amoxicillin-clavulanic acid. *Hepatol Res* 2015; 45: 676–682.
- 8) Kim SM, Kim Y, Jeong K., et al. Logistic LASSO regression for the diagnosis of breast cancer using clinical demographic data and the BI-RADS lexicon for ultrasonography. *Ultrasonography* 2018; 37: 36–42.
- 9) Low Y, Uehara T, Minowa Y., et al. Predicting drug-induced hepatotoxicity using QSAR and toxicogenomics approaches. *Chem Res Toxicol* 2011; 24: 1251–1262.
- 10) Zhu X. W, Sedykh A, Liu SS. Hybrid in silico models for drug-induced liver injury using chemical descriptors and in vitro cell-imaging information. *J Appl Toxicol* 2014; 34: 281–288.
- 11) Chen M, Suzuki A, Thakkar S, et al. DILrank: the largest reference drug list ranked by the risk for developing drug-induced liver injury in humans. *Drug Discov Today* 2016; 21: 648–653.

## 低ナトリウム血症を起こしにくい抗うつ薬の同定

長島卓也<sup>1)2)</sup>, 早川隆<sup>1)2)</sup>, 秋元勇人<sup>1)2)</sup>, 皆川貴美乃<sup>2)</sup>,  
高橋泰夫<sup>2)</sup>, 浅井聡<sup>1)2)</sup>

### Identifying antidepressants less likely to cause hyponatremia

Takuya NAGASHIMA<sup>1)2)</sup>, Takashi HAYAKAWA<sup>1)2)</sup>, Hayato AKIMOTO<sup>1)2)</sup>,  
Kimino MINAGAWA<sup>2)</sup>, Yasuo TAKAHASHI<sup>2)</sup>, Satoshi ASAI<sup>1)2)</sup>

#### 要旨

抗うつ薬は副作用として低ナトリウム血症を引き起こすことが知られているが、各抗うつ薬のリスクについては統一見解が得られていない。そこで筆者は、①後方視的コホート研究、②不均衡解析、③薬剤疫学・薬力学解析の3種を組み合わせたトライアングレーション研究により、低ナトリウム血症を起こしにくい抗うつ薬の同定を試みた。その結果、ミルタザピン、フルボキサミン、ミルナシブランは、①血清ナトリウム値の低下が認められず、②低ナトリウム血症報告の有意な増加が認められず、③セロトニントランスポーターに対する結合親和性が低いという①②を説明可能な薬力学的特性を有していることから、他の抗うつ薬に比べ低ナトリウム血症を起こしにくいと考えられた。

#### 1. はじめに

抗うつ薬はうつ病をはじめとする様々な疾患の治療に広く用いられている。中でも新世代抗うつ薬は、旧世代抗うつ薬で問題視されていた心毒性等の副作用を起こしにくいことから第一選択薬として推奨されている。しかしながら近年、一部の副作用については新世代抗うつ薬の方が旧世代抗うつ薬よりも起こりやすいことが明らかとなってきた。低ナトリウム血症はそのような副作用の典型例であり、進行すると痙攣や意識障害といった重篤な症状を引き起こす。加えて、低ナトリウム血症の初期症状(食欲低下、倦怠感等)はうつ症状と類似しているため、軽症のうちその発生を正しく見極めるのは困難である。故に、抗うつ薬誘発低ナトリウム血症のリスクを最小限に抑えるには、低ナトリウム血症を起こしにくい抗うつ薬を予め同定しておくことが重要である。

抗うつ薬と低ナトリウム血症の関係については複数の海外エビデンスが存在するが、各抗うつ薬のリスクについては統一見解が得られていない。

そこで筆者は、①後方視的コホート研究、②不均衡解析、③薬剤疫学・薬力学解析の3種を組み合わせたトライアングレーション研究により、低ナトリウム血症を起こしにくい抗うつ薬の同定を試みた。

#### 2. 対象及び方法

##### 2. 1 後方視的コホート研究

データソースには日本大学臨床データベース(Nihon University School of Medicine's Clinical Data Warehouse [NUSM's CDW], 2004-2020)を用いた。抗うつ薬誘発低ナトリウム血症は主に高齢者において起こりやすいことが知られているため、解析対象は60歳以上の高齢者とした。曝露因子は新世代抗うつ薬の初回処方とし、アウトカムは投与開始から30日以内における血清ナトリウム値の変化を用いて評価した。群間の患者背景の違い(年齢、性別、合併症、併用薬、投与前の血清ナトリウム値)は多変量回帰分析によ

<sup>1)</sup> 日本大学医学部生体機能医学系薬理学分野

<sup>2)</sup> 日本大学医学部臨床試験研究センター

長島卓也 : nagashima.takuya@nihon-u.ac.jp

り調整した。 $P < 0.05$ を統計学的に有意と判定した。本研究は日本大学医学部倫理委員会の承認を得た上で実施した(承認番号: P21-01-0)。

## 2. 2 不均衡解析

データソースには医薬品医療機器総合機構が公開している副作用報告データベース(Japanese Adverse Drug Event Report [JADER] database, 2004-2020)を用いた。解析対象は後方視的コホート研究と同じく60歳以上の高齢者とした。曝露因子は新世代抗うつ薬の使用とし、アウトカムは低ナトリウム血症/抗利尿ホルモン不適合分泌症候群(SIADH)の報告オッズ比を用いて評価した。群間の患者背景の違い(年齢, 性別, 合併症, 併用薬)は傾向スコアマッチングにより調整した。 $P < 0.05$ を統計学的に有意と判定した。

## 2. 3 薬剤疫学・薬力学解析

薬剤疫学・薬力学解析(pharmacoepidemiological-pharmacodynamic [PE-PD] analysis)は、薬剤疫学データと薬力学データを組み合わせることで有害事象の分子メカニズムを探索する新たな手法である。本研究では、各抗うつ薬の血清ナトリウム値の変化又は低ナトリウム血症の報告オッズ比と、各標的分子に対する結合親和性との間の相関を評価した。標的分子としてはセロトニン・ノルアドレナリン・ドパミントランスポーター及びセロトニン・アドレナリン・ドパミン・ヒスタミン・ムスカリン受容体を解析対象とした。 $P < 0.05$ を統計学的に有意と判定した。

## 3. 結果

### 3. 1 後方視的コホート研究

新世代抗うつ薬のクラス間比較において、血清ナトリウム値の有意な低下は選択的セロトニン再取り込み阻害薬(SSRI)投与群及びセロトニン・ノルアドレナリン再取り込み阻害薬(SNRI)投与群で認められた。一方で、ノルアドレナリン作動性・特異的セロトニン作動性抗うつ薬(NaSSA)投与群では血清ナトリウム値の低下が認められなかった。新世代抗うつ薬のクラス間で、血清ナトリウム値の変化度合いに有意な差が認められた。患者背景の違いを統計学的に調整しても、その差は依然として有意であった。

SSRIのクラス内比較において、血清ナトリウム値の有意な低下はパロキセチン投与群で認められた。セルトラリン投与群及びエスシタロプラム投与群においても血清ナトリウム値の低下傾向が認められた。一方で、フルボキサミン投与群では血清ナトリウム値の低下が認められなかった。SSRIに属する薬物間で、血清ナトリウム値の変化度合いに有意な差が認められた。患者背景の違いを統計学的に調整しても、その差は依然として有意であった。

SNRIのクラス内比較において、血清ナトリウム値の有意な低下はデュロキセチン投与群で認められた。一方で、ミルナシプラン投与群では血清ナトリウム値の低下が認められなかった。SNRIに属する薬物間で、血清ナトリウム値の変化度合いに有意な差が認められた。患者背景の違いを統計学的に調整しても、その差は依然として有意であった。

### 3. 2 不均衡解析

新世代抗うつ薬のクラス間比較において、低ナトリウム血症報告の有意な増加はSSRI使用群及びSNRI使用群で認められた一方、NaSSA使用群では認められなかった。SSRIのクラス内比較において、低ナトリウム血症報告の有意な増加はパロキセチン使用群、セルトラリン使用群及びエスシタロプラム使用群で認められた一方、フルボキサミン使用群では認められなかった。SNRIのクラス内比較において、低ナトリウム血症報告の有意な増加はデュロキセチン使用群で認められた一方、ミルナシプラン使用群では認められなかった。

### 3. 3 薬剤疫学・薬力学解析

後方視的コホート研究で得られた各抗うつ薬の血清ナトリウム値の変化は、セロトニントランスポーター(SERT)に対する結合親和性と有意に相関していた。一方で、他のトランスポーター・受容体に対する結合親和性との間に有意な相関は認められなかった。SERTの結合親和性に対する有意な相関は、不均衡分析における各抗うつ薬の低ナトリウム血症報告オッズ比を用いても依然として認められた。

## 4. 考察

新世代抗うつ薬のクラス間比較において、血清

ナトリウム値の低下は SSRI 投与群及び SNRI 投与群で認められた一方、NaSSA(ミルタザピン)投与群では認められなかった。低ナトリウム血症報告に関しても、ミルタザピン使用群においてのみ有意な上昇が認められなかったことから、ミルタザピンは他の新世代抗うつ薬に比べ低ナトリウム血症を起こしにくい可能性が示唆された。この知見は海外における複数の疫学研究結果に一致するものであり、低ナトリウム血症リスク低減におけるミルタザピンの有用性は日本人患者においても成立すると考えられた。

SSRI のクラス内比較ではフルボキサミン投与群において、SNRI のクラス内比較ではミルナシプラン投与群において血清ナトリウム値の低下が認められなかった。低ナトリウム血症報告に関しても、フルボキサミン使用群及びミルナシプラン使用群においてのみ有意な上昇が認められなかったことから、フルボキサミン及びミルナシプランは他の SSRI・SNRI に比べ低ナトリウム血症を起こしにくい可能性が示唆された。本研究で得られたこの新規知見に関しては、今後さらなる研究を重ね、その妥当性を検証する必要があると考えられた。

抗うつ薬誘発低ナトリウム血症の機序は未だ説明されていないが、複数の研究において抗利尿ホルモン (ADH) の過剰分泌の関与が示唆されている。

血中ナトリウム濃度の制御を担う ADH は中枢神経系で産生されることから、抗うつ薬が中枢作用を介して低ナトリウム血症を引き起こしている可能性は十分に考えられる。本研究において、血清ナトリウム値の低下と SERT に対する結合親和性との間に有意な相関が認められたことから、抗うつ薬誘発低ナトリウム血症は SERT の阻害を介している可能性が示唆された。同時に、ミルタザピン、フルボキサミン、ミルナシプランによる低ナトリウム血症の起こりにくさは、SERT に対する結合親和性が低いという薬力学的特性からも支持された。

### 5. 結語

ミルタザピン、フルボキサミン、ミルナシプランは、①血清ナトリウム値の低下が認められず、②低ナトリウム血症報告の有意な増加が認められず、③SERT に対する結合親和性が低いという①②を説明可能な薬力学的特性を有していることから、他の抗うつ薬に比べ低ナトリウム血症を起こしにくいと考えられた(図 1)。本研究は、現代の抗うつ薬療法を最適化する上で有用な臨床的知見を提供するのみならず、将来、より安全な抗うつ薬を開発する上で重要な薬理学的知見を提供するものである。

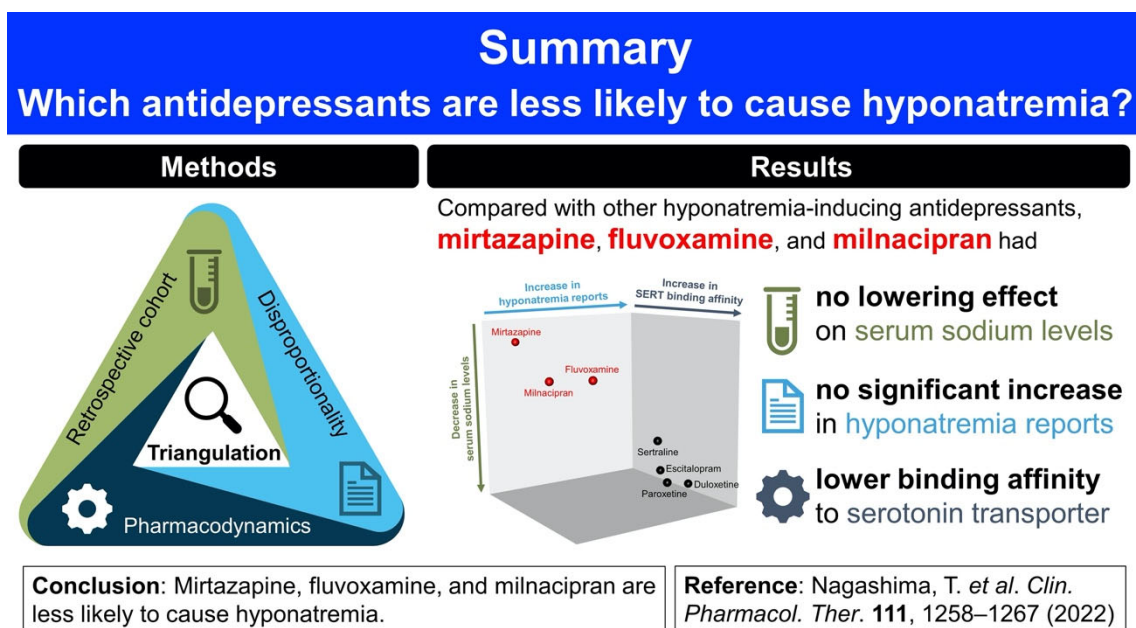


図 1 本研究の要約

## 6. 文献

- 1) Nagashima T, Hayakawa T, Akimoto H, et al. Identifying Antidepressants Less Likely to Cause Hyponatremia: Triangulation of Retrospective Cohort, Disproportionality, and Pharmacodynamic Studies. Clin Pharmacol Ther 2022; 111: 1258-1267.

## 日本大学医学部臨床情報統合データベースに基づき疾患リスク解析 を行う機械学習パイプラインの構築

早川隆<sup>1)2)</sup>, 長島卓也<sup>1)2)</sup>, 秋元勇人<sup>1)2)</sup>,  
皆川貴美乃<sup>2)</sup>, 高橋泰夫<sup>2)</sup>, 浅井聡<sup>1)2)</sup>

### A machine learning pipeline for risk analysis based on the integrative database of clinical information in NUSM

Takashi HAYAKAWA<sup>1)2)</sup>, Takuya NAGASHIMA<sup>1)2)</sup>, Hayato AKIMOTO<sup>1)2)</sup>,  
Kimino MINAGAWA<sup>2)</sup>, Yasuo TAKAHASHI<sup>2)</sup>, Satoshi ASAI<sup>1)2)</sup>

#### 要旨

日本大学医学部臨床情報統合データベースに蓄積されている薬剤処方・基礎疾患・検査値等の時系列情報から従来の疫学的手法よりも高精度に疾患リスクを見積もる機械学習パイプラインを構築した。データベースに記録されているイベント情報に基づき後ろ向きコホートを抽出し、機械学習アルゴリズムが適用できるように、観察期間の各月の薬剤の投与期間・総投与量、基礎疾患の罹患期間、検査値の推移などの値を時系列順に並べたデータテーブルを自動生成するプログラムコードを開発した。そしてマルチカーネル学習の GPU 上の高速実行コードを開発し、データテーブルから標的疾患発症の非線形リスクを高速に推定することを可能にした。

#### 1. 序論

我々のグループでは日本大学医学部の附属病院における 2004 年 1 月以降の薬剤処方・病名・検査値の全記録情報からなる臨床情報統合データベースを構築し、これを用いた疾患リスクの研究を行ってきた。データベースを用いることによって、各臨床研究ごとに一からデータ収集を行う場合と比べて効率的な研究が可能となった。近年、このような臨床情報データベースに機械学習の手法を適用して、従来の疫学的手法を超える精度で疾患リスクを評価しようとする試みが盛んになされている<sup>1)</sup>。しかしながら、このようなデータベースから機械学習によって高精度な情報を取り出すには、(1)機械学習を適用可能なデータフォーマットへの自動変換、(2)データの大規模性を余すところなく活用するための機械学習アルゴリズムの高効率実装、の 2 つが不可欠であり、これまで発表されてきた先行研究は上記 2 点において未だ発展途

上と言わざるを得ない。我々は上記 2 点を克服する第一歩として、データベースに含まれる臨床情報の時系列を従来よりも大規模かつ自動的に抽出し、機械学習アルゴリズムを適用可能な標準形式に変換し、そこから疾患リスクを高速に推定し統計的評価を与える機械学習プログラムのパイプラインを開発した。特に ATC 薬理作用・化学分類あるいは研究者の設定したカスタム分類ごとの薬剤の標準化投与量の算出を自動化したことで、投与量に基づく精密なリスク評価が可能となった。また、従来標準的に用いられてきた(一般化)Cox 回帰法よりも時間的情報を精密に抽出できるマルチカーネル法による最尤推定とその高効率実行コードを開発したことで、薬剤の投与期間や投与量などの説明変数の変化に伴う疾患リスクの非線形変化を系統的・効率的に評価することが可能となった。

<sup>1)</sup> 日本大学医学部生体機能医学系薬理学分野

<sup>2)</sup> 日本大学医学部臨床試験研究センター

早川隆 : hayakawa.takashi@nihon-u.ac.jp

## 2. 標準時系列データテーブルと疾患リスクハザードモデル

一般的に疫学研究では疾患発症リスクを記述する統計モデルとして、ある時刻に単位時間あたりどれだけの確率でその疾患を発症するか(ハザードレート)を、薬剤服用の有無・検査値・基礎疾患の有無などの説明変数の関数として表すハザードモデルが用いられ<sup>2)</sup>、本研究でもこのモデルを採用した。一般的な疫学研究のモデルと本研究の機械学習モデルの相違点は、臨床情報データベースに含まれる大量の説明変数候補から人為的に説明変数を絞り込む操作を最小限に抑えることができ、さらに説明変数に対して非線形に増加するリスク因子を複数推定できるようになる点である。

以下にモデルの概要を簡単に説明する。図1(a)左に示すように各患者ごとに観察期間中の各月において、ATC 薬理作用・化学分類ごとの薬剤の服用量・服用期間、(補完された)検査値の推移、ICD10 コードで表された疾患グループの罹患の有無・罹患期間などが説明変数値として与えられているものとする。本研究のハザードモデルでは、この説明変数のそれぞれに対して標的疾患の非線形リスク関数が設定されており[図1(a)中]、説明変数の値に応じて標的疾患の発症リスクが増加し、この各説明変数ごとのリスク増加の総和として当該月にその患者が標的疾患を発症するハザードレートが定まる。各リスク関数はデータから統計的に推定する。説明変数が少数でリスク関数が線形

の場合、このモデルは従来疫学研究で用いられてきたモデルと一致する。

上記のモデルでリスク関数を推定するためには、数万人単位の患者の数百~数千の説明変数の値を観察期間中の各月ごとに用意する必要がある。臨床情報データベースに格納されている情報は、概ね図1(b)のようなテキスト形式になっており、機械学習のアルゴリズムを適用可能なフォーマットへの変換が必要であった。薬剤処方情報の変換では各月の服用量を求めるために、ある薬剤カテゴリー(例えば ATC 化学分類 C08CA)の全ての薬剤の当該月の投与量を計算し、それらを薬剤添付文書に記載されている標準一日投与量を基準単位として合計する必要があった。基礎疾患については、個々の ICD10 小分類ごとではなく、慢性腎臓病(stage3以降): N18.3-N18.5 といった臨床的な意味付けの可能なグループ分けをおこなった上で、グループごとの罹患の有無・罹患期間を説明変数の値として算出する必要があった。検査値についても、HbA1c (JDS) 値と HbA1c (NGSP) 値の変換や、異なる検査会社の同じ検査項目をまとめるなどの前処理を必要とした。こうした多数のデータクリーニングのステップを自動化するプログラムを作成し、図1(b)の形式の生データから図1(a)左に示される説明変数値の時系列を表すデータテーブルを生成できるようにした。

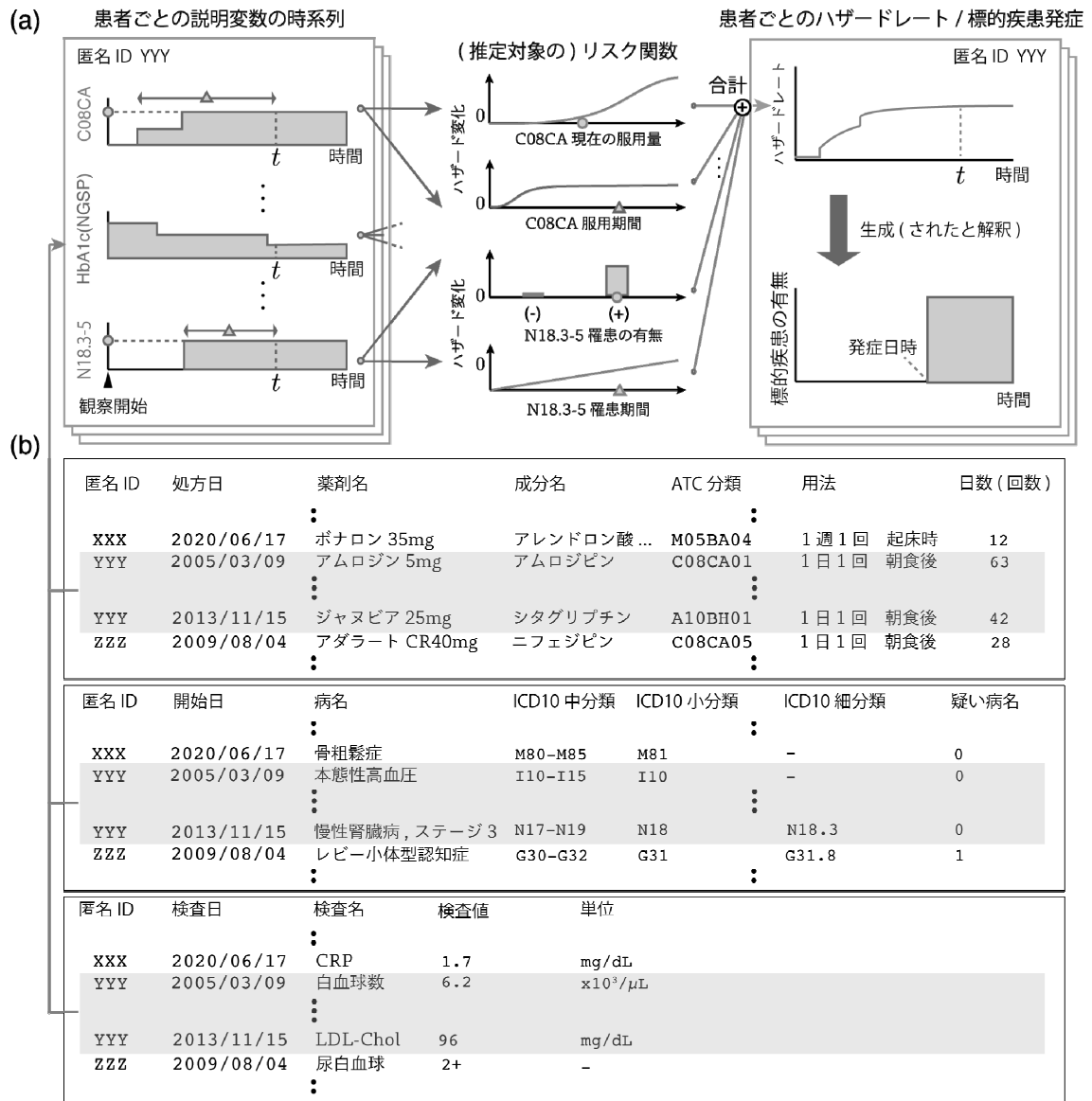


図 1. (a)標的疾患発症の非線形ハザードモデル (b) 臨床情報統合データベースのデータ形式の概要

### 3. マルチカーネル法に基づくスパース正則化付き最尤推定

本研究では、前節で述べた大量の説明変数それぞれの非線形リスクを推定する機械学習プログラムコードを開発した。非線形リスクの推定は従来の疫学研究においてもスプライン関数を用いた方法が用いられていたが、せいぜい1~2個の非線形関数しか推定対象に含めることができなかった。このため、薬剤の現在の投与量と過去の総投与量のどちらが疾患の主たるリスク要因なのか、あるいはその薬剤治療が適応となる基礎疾患の罹患期

間の方が主たるリスク要因なのか、といった基本的な疫学的な問いに答えることができなかった。他方で、2000年代に入ってから機械学習分野では多数の非線形関数を含むモデルの最尤推定とその統計的性質についての研究が急速に進んでおり、このような技術の一つであるマルチカーネル法を用いることで上記の問いに答えることができるようになっていた<sup>3)</sup>。大量の説明変数がありそれほどデータ量がない場合でも、少数の説明変数と非線形関数を用いて疾患リスクが表せるならば、ほとんどのリスク関数を零とするスパース正則化の



ついた推定を行うことで、第二種の統計的過誤を最小限に抑えた上で、重要な説明変数の非線形リスク関数を精度よく推定することができる。そして、薬剤投与量と基礎疾患のどちらが標的疾患の主たるリスク要因なのか、といった問いに対して、どちらのリスク関数が非零値と推定されるかを調べることで答えることができる(マルチカーネル法のオラクル性として知られる<sup>4)</sup>)。

また、従来の疫学研究ではハザードモデルのリスク推定にもっぱら(一般化)Cox 回帰法が用いられてきた。Cox 回帰法では、統計モデルとデータから自然に定まる尤度を最大化するのではなく、大幅に情報を省いて構成した部分尤度を最大化する。この部分尤度は、同時刻(あるいは同年齢)の患者のうちどの患者が標的疾患を発症したか、といった患者間比較に基づいてのみ算出されるため、時間的連続性に関する情報が大幅に省かれる。従来の疫学研究は、患者エントリーの後、定められたチェックポイントまで説明変数の情報を更新しないデザインが多く、この場合 Cox 回帰法で情報を省くことが一定の好ましい効果をもたらしていた。しかしながら、臨床情報データベースを用いる場合、観察期間中の当該患者の全ての説明変数の正確な時系列情報が容易に得られ、Cox 回帰法によって時間的情報を大幅に捨ててしまうことは好ましくない。

以上の考察に基づき、本研究ではマルチカーネル法を用いたハザードモデルのリスク関数の最尤推定、およびそれに付随して疫学的研究で必要とされる適合度検定や信頼区間推定といった統計処理を行うプログラムライブラリを、C++ / OpenACC 言語で開発した。

#### 4. プログラムライブラリの計算効率性・メモリ効率性

前節で述べたマルチカーネル法の最尤推定は、エントリー患者の各月ごとの説明変数の値を全て推定に用いるため、かなり大きな計算量が必要とされ、また計算中にこれらのデータを常時格納するためかなりのメモリ容量が必要となる。我々のグループの研究プロジェクトの一つとして認知症の発症リスクの解析を行なっているが、患者条件を厳しく絞込んだ場合でもデータ長は 431,698 [人・月]に及び、この一つ一つに対して数

十～数百個の説明変数が対応する。

このような大規模な解析における計算量とメモリ容量の問題を我々は、(i) 収束の速いアルゴリズムの使用、(ii) GPU アクセラレーターを用いた高速化、(iii) GPU 付属メモリへのデータの効率的な割り振り、の3点で工夫を凝らした計算プログラムを開発することで解決した。一般的に無料公開されている PyTorch・TensorFlow といった機械学習ライブラリはこれらの点で不十分であった。

まず、スパース正則化付きの最尤推定には拡張ラグランジアン法<sup>5)</sup>を用いた。スパース正則化付き尤度関数は微分不可能な点を含む凹関数であり、勾配法やニュートン法といった単純な方法は利用できない。非平滑凸最適化法を用いる必要があるが、これには大きく分けて二種類の方法がある。一つは Nesterov 加速法と呼ばれる方法とその派生法<sup>6)</sup>で、もう一つが拡張ラグランジアン法である。前者では尤度を最大化する際に、最大尤度からの誤差を大域的に定数×計算回数<sup>-2</sup>乗以下に抑えて収束させることが保証されており、後者では最大尤度周囲の局所において誤差を二次収束させる(定数×計算回数の指数関数の指数関数以下に抑える)ことが保証されている。実装の容易さのため前者の方がよく用いられるが、本研究では前者は実用に足らず後者の性質の方が重要であることが判明し、この方法を採用した。

次にこの拡張ラグランジアン法において計算用メモリを節約する近似法を用いる必要があった。マルチカーネル法ではデータ長  $T(=431,698[\text{人} \cdot \text{月}])$  に対して説明変数  $x_i$  ごとに  $T^2$  の要素を持つ行列  $G_i$  (Gram 行列という)を操作する必要があるが、このサイズのデータを計算機のメモリに載せることは不可能である。これを解決するいくつかの方法が知られており、我々は文献<sup>7)</sup>に従い、不完全 Cholesky 分解と呼ばれる  $G \approx LL^T$  ( $L$ は要素数  $TM$  個の行列、 $L^T$ は  $L$  の転置行列)なる近似を用いた。 $M$  を大きくしていき、 $M=T$  となる時に近似ではなく完全に正確な計算となる。 $M < T$  では、近似計算のため正確な結果を得られない恐れがあるが、実用上はこの近似は図 2 の例のように  $M=40$  程度で既に行列の Frobenius ノルムに関して 99.9%以上の近似精度を出せるため問題にならない。また、拡張ラグランジアン法の内部でニュートン法が用いられるが、ニュートン法もまたデータ長  $T$  に対

して  $O(T^2)$  のメモリ,  $O(T^3)$  の計算量を必要とする。我々は、ニュートン法をメモリ制限 BFGS 法<sup>8)</sup>で置き換えたアルゴリズムを用いることでメモリ・

計算量を大幅に減らすことに成功した(メモリ制限 BFGS 法は近似法ではないので結果の精度には影響しない)。

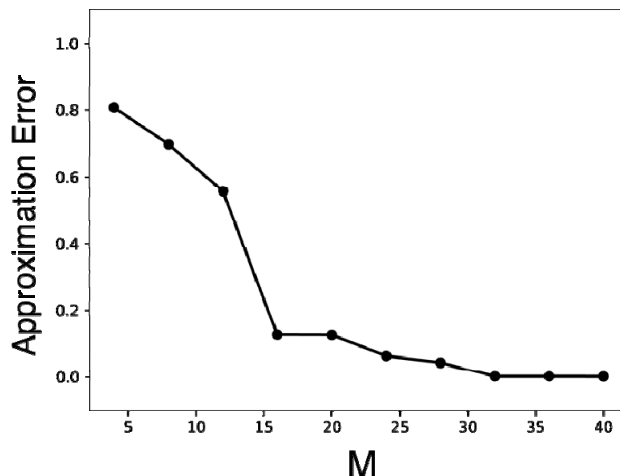


図 2. データ長  $T=431,698$  [人・月] の認知症発症リスクの推定であらわれる代表的な Gram 行列の不完全 Cholesky 分解  $G \approx LL^T$  ( $L$  は  $T \times M$  行列) の近似誤差。縦軸の近似誤差は  $\|G - LL^T\| / \|G\|$  で表した ( $\|\cdot\|$  は行列の Frobenius ノルム)。

最後に、これらのアルゴリズムは全て GPU アクセラレーター上で実行されるようにコードを改良した。近年では CPU の性能限界のため、大きな計算は全て GPU 上で実行することが慣例となってきた。この際、CPU-GPU 間の通信は通常極めて遅く、実行速度を上げるための最適な方法は、全ての計算を GPU に行わせ、CPU-GPU 間のデータ通信を極力無くすることとされる。我々は拡張ラグランジアン法のほぼ全ての計算を GPU 上で行うコードを開発し、4CPU コアの場合と 4CPU コア+4GPU カードの場合を比べて 30 倍近い大幅な計算速度の向上を確認した[表 1]。GPU 計算にはメインメモリは使用できないため、データを効率的に GPU 付属メモリに格納する必要があり、この点をうまく調整することで高速化が達成された。

## 5. 結語・展望

大規模臨床情報データベースを活用し従来の疫学的解析よりも高精度な疾患リスク解析を行うために、データベースに格納されている情報を適切なフォーマットに自動変換し、大規模データに適した数学的特性を持つ機械学習アルゴリズムを適用するプログラムパイプラインを作成した。開発したプログラムは大規模解析の自動化・計算効率性・メモリ効率性の高さといった点で有用であることを確認できた。このプログラムによって今後具体的に種々の疾患リスクの解析を進めていけるものと期待される。実際にいくつかの研究プロジェクトで解析が進行中であり、近日論文公開予定である。また、このプログラムをベースとしてベイズ推論を取り入れたさらに精緻な解析プログラムの開発も進行中であり、本研究結果がその基礎としての役割を果たしている。

表 1 認知症発症リスクの推定に要した時間(コホートサイズ 431,698 人・月)

計算環境	57 説明変数	661 説明変数
4×CPU(Intel Xeon E5-2680 V4)	2187.3±552.4 [秒]	(計算時間大のため 評価困難)
4×CPU(Intel Xeon E5-2680 V4) + 4 ×Tesla P100 16GB	78.3±14.4 [秒]	823.4±175.2 [秒]

## &lt;謝辞&gt;

本研究での計算プログラム開発の一部は、東京工業大学のスーパーコンピュータ TSUBAME3.0 を利用して実施した。

## 6. 文献

- 1) Shinozaki A. "Electronic medical records and machine learning in approaches to drug development," a peer-reviewed book chapter in "Artificial intelligence in oncology drug discovery and development." IntechOpen. 2020
- 2) Kleinbaum DS, Survival analysis: A self-learning text. 3rd Ed. Springer. 2011.
- 3) Lanckriet, NGR. Bartlett CP, Ghaoui LE, Jordan MI. Learning the kernel matrix with semidefinite programming. Journal of Machine Learning Research 2004; 5:27-72.
- 4) Bach FR. Consistency of the group lasso and multiple kernel learning. Journal of Machine Learning Research . 2008;9(6)
- 5) Suzuki T, Tomioka R. SpicyMKL: a fast algorithm for multiple kernel learning with thousands of kernels. Machine Learning. 2011; 85(1):77-108.
- 6) Nesterov Y. Gradient methods for minimizing composite functions. Universite catholique de Louvain, CORE Discussion Papers, 2007
- 7) Bach FR, Jordan MI. Kernel independent component analysis. Journal of Machine Learning Research ,2002;3:1-48.
- 8) Liu DC, Nocedal J. On the limited memory BFGS method for large scale optimization. Mathematical Programming, 1989;45:503-528.

## *DLGAP1* ノックアウトマウスの表現型解析

皆川貴美乃<sup>1)</sup>, 早川隆<sup>1)2)</sup>, 秋元勇人<sup>1)2)</sup>, 長島卓也<sup>1)2)</sup>,  
高橋泰夫<sup>1)</sup>, 浅井聡<sup>1)2)</sup>

### Behavioral characterization of *DLGAP1* knockout mice with SHIRPA primary screen

Kimino MINAGAWA<sup>1)</sup>, Takashi HAYAKAWA<sup>1)</sup>, Hayato AKIMITO<sup>1)2)</sup>,  
Takuya NAGASHIMA<sup>1)2)</sup>, Yasuo TAKAHASHI<sup>1)2)</sup>, Satoshi ASAI<sup>1)2)</sup>

#### 要旨

Discs large associated protein (DLGAP) は、シナプスを構成する足場タンパク質であり、精神・神経疾患への関与が報告されている。本研究では、*DLGAP1* ノックアウト (DLGAP1-KO) マウスの表現型解析として、網羅的行動テストバッテリーを行った。その結果、8~10 週齢における DLGAP1-KO マウスでは野生型マウスと比較して脱糞数の有意な増加が見られ、恐怖関連行動を引き起こす可能性が示唆された。

#### 1. はじめに

精神・神経疾患は、単一遺伝子の変異に基づく疾患とは異なり、環境因子とともに多因子性を示すことからその病態発症機序は複雑である。認知機能障害や反復行動など不均一な行動表現型を示す精神・神経疾患の原因として、シナプス構成成分のゲノム変異が多数報告されている<sup>1), 2)</sup>。

Discs large associated protein (DLGAP) は、シナプス後肥厚 (postsynaptic density: PSD) を構成する足場タンパク質であり、1~5 のアイソフォームが知られている。DLGAP1~4 は脳内で高発現し、これらの遺伝的変異は統合失調症や強迫性障害などに関与することが報告されている<sup>3)</sup>。また DLGAP1 では神経系以外の機能として、脂肪細胞の増殖と分化<sup>4)</sup>や癌細胞の浸潤<sup>5)</sup>に関わることが報告されている。本研究では DLGAP1 の機能を検討するために、ノックアウトマウスを用いて表現型解析を行った。

#### 2. 対象及び方法

#### 実験動物

C57BL/6N を遺伝的背景とする *DLGAP1* 遺伝子欠損 (DLGAP1-KO) マウスは、Cyagen 社から購入した。ヘテロ接合型マウス同士の交配で得られた、雌雄 8~10 週齢の DLGAP1 野生型、ヘテロ接合型、ホモ接合型マウス (各群 11 匹~12 匹) を使用した。実験動物の取り扱いは、日本大学医学部実験指針と日本大学動物実験マニュアルを遵守し、日本大学実験動物委員会により承認された実験課題 (承認番号: AP20MED067-1) として行った。日本大学医学部実験動物施設内において、恒温 23±2 度、恒湿 55±5%, 明暗サイクルを 12 時間 (8:00~20:00) とし、自由飲水摂食下にて飼育した。

#### 行動テストバッテリー

DLGAP1-KO マウスの表現型スクリーニングとして、理化学研究所マウス表現型解析開発チームで行われている RIKEN modified-SHIRPA 法を用いた<sup>6)</sup>。Viewing jar (高さ 30 cm×内径 10 cm のアクリル製シリンダー) 内にマウスを入れて 5 分間、姿勢、自発運動活性、呼吸の荒さ、震えの有無の

<sup>1)</sup> 日本大学医学部臨床試験研究センター

<sup>2)</sup> 日本大学医学部生体機能医学系薬理学分野

皆川貴美乃: minagawa.kimino@nihon-u.ac.jp

観察を行った。この時、排尿の有無と排便数を記録し、特異な動きは別に記録した。Arena（内径60cm×35 cm×20 cm アクリル製の箱）での行動観察は、約30 cmの高さから落としたマウスが動き出すまでの時間および30秒間の行動量と歩き方、

前進中の骨盤の高さおよび尾の高さについて観察を行った。30秒間の行動量は、11 cm×11 cmを1マスとして移動したマスを数えた。観察内容は、表1に示すようにスコアリングにより評価した。

表1 観察項目とスコア

観察内容	スコア
姿勢	0:完全に寝そべる, 1:横に寝る, 2:臥位, 3:四肢で立つ, 4:後肢で立つ, 5:ジャンプを繰り返す
自発運動活性	0:無動, 1:何気なく毛繕い, 2:活発に毛繕い, 3:活発に動く, 4:極めて活発に動く
呼吸の荒さ	0:不規則, 1:ゆっくり浅く, 2:普通, 3:亢進
震えの有無	0:無, 1:少し, 2:著しい
歩き方	0:普通, 1:流れているようだが普通ではない, 2:制限された動きのみ, 3:無動
骨盤の高さ	0:著しく引きずる, 1:わずかに床に触れる, 2:3 mm浮いている, 3:3 mm以上浮いている
尾の高さ	0:引きずる, 1:水平, 2:高い

### 統計解析

得られたデータは、平均値±標準誤差で示した。統計解析は one-way ANOVA と多重比較検定 (Tukey-Kramer test あるいは Steel-Dwass test) を行い、\* $p < 0.05$ , \*\* $p < 0.01$  を有意差ありとした。またスコアリングデータにおいては比率で示した。

### 3. 結果

Arena における観察の結果、動き出すまでの時

間や30秒間の行動量、歩き方について、野生型とホモ型およびヘテロ型マウスとの間に有意差はなかった (表2)。雌雄共に野生型と比較してホモ型マウスでは、尾の位置が高いマウスの割合が高い (雄:45% vs 82%, 雌:27% vs 73%, 図2上段)。また骨盤の高さについては、ホモ型およびヘテロ型マウスの方が野生型と比較して高い傾向にある一方で、ホモ型の雌マウスでは、わずかに床に触れる程度まで骨盤の位置が低いマウスが観察された (図2下段)。

表2 SHIRPA スクリーニングによる DLGAP1-KO 表現型解析の結果 (各群 n=12)

テスト	雄			雌		
	野生型	ヘテロ型	ホモ型	野生型	ヘテロ型	ホモ型
動き出すまでの時間 (秒)	3.51 ± 1.0	1.26 ± 0.3	1.59 ± 0.4	1.15 ± 0.3	1.30 ± 0.4	3.24 ± 1.3
30秒間の行動量 (マス)	17.91 ± 1.6	21.00 ± 1.5	18.57 ± 1.6	18.25 ± 2.3	17.00 ± 1.7	16.17 ± 1.8

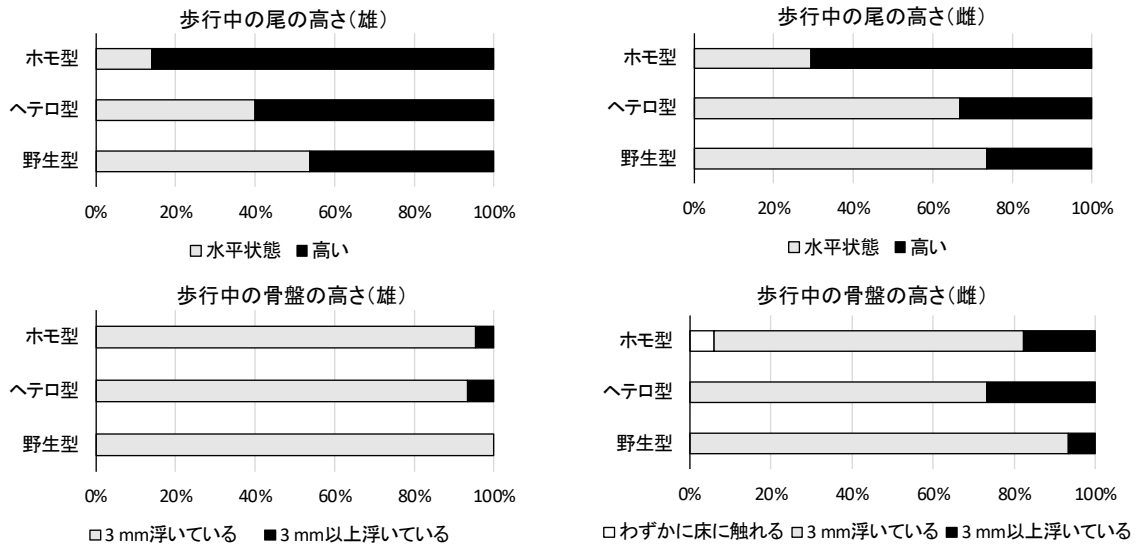


図1 歩行中の尾の長さ(上段)と骨盤の高さ(下段)の各群での比率

Viewing jar 内の観察では、各群のマウスにおいて、形態学的変化は見られず、姿勢や自発運動活性、呼吸の荒さ、震えの有無においても特異な変化は観察されなかった(データは示さない)。一方で、雌の脱糞数については、野生型と比較してのホモ型 ( $p = 0.0083$ ) とヘテロ型 ( $p = 0.0067$ )

で有意な増加が見られた(図2右下)。排尿の有無についても雌ではホモ型、ヘテロ型共に野生型と比較して排尿したマウスの割合が高い(図2右上)。雄では、ホモ型のマウスでのみ野生型と比較して脱糞数に有意な増加が認められた ( $p = 0.0245$ , 図2左下)。

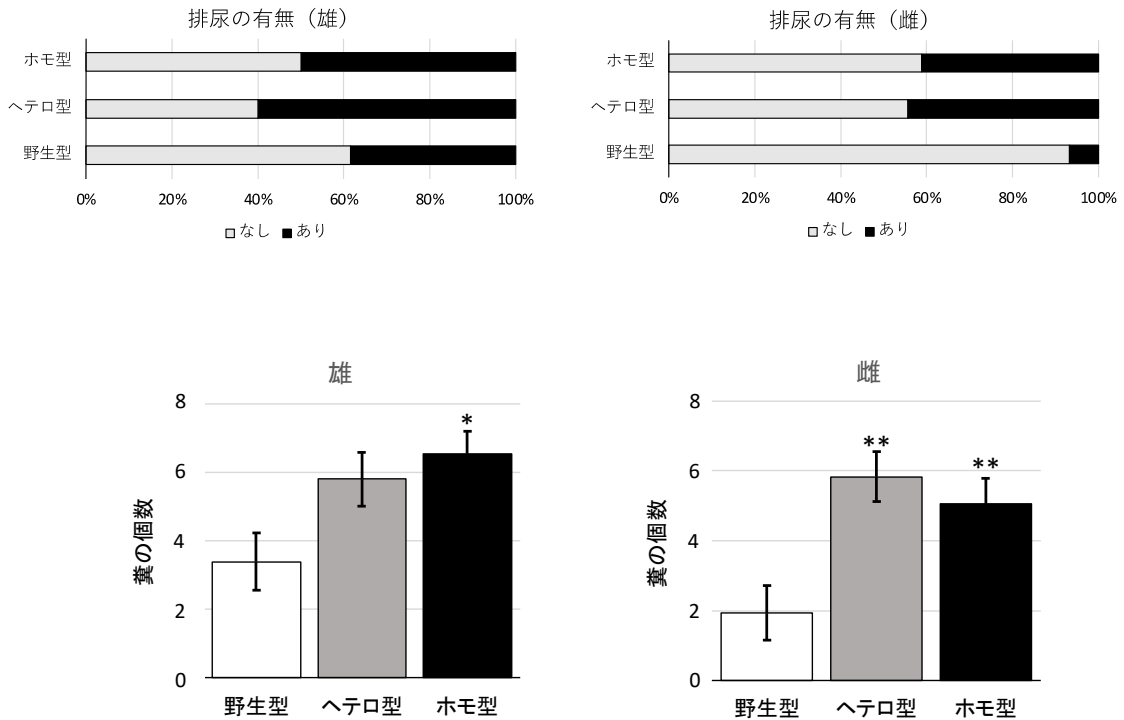


図2. Viewing jar 内での観察：排尿の有無と脱糞の個数(各群 n=11)

#### 4. 考察

本研究では、遺伝子改変マウスの網羅的な表現型解析方法として有用である RIKEN Modified-SHIPRA 法を用いて、DLGAP1-KO マウスの初期スクリーニングを行った。本検討では、8~10 週齢のマウスを使用して観察評価を行ったが乳幼児期の発達段階において、身体的特徴（耳介の開展、眼瞼開裂、毛生）や哺乳行動は、野生型と DLGAP1-KO マウスとの間に変化は観察されなかった。テスト期間中ではない4~6 週齢のホモ型マウスでは、自身の毛繕い行動による臀部の脱毛が雄 22 匹中 4 匹、雌 17 匹中 4 匹のマウスで観察された。この脱毛が続いたのは1 週間程度でその後脱毛は改善した。過度の毛繕い行動は、ストレス状態を示す指標となる。またホモ型の雌雄マウスにおいて、恐怖関連行動の指標の1つである脱糞数の有意な増加が見られたことから、DLGAP1-KO マウスは情動機能の異常が引き起こされる可能性が示唆された。今回の検討は、予備的行動観察であることから、今後は24 時間の継時的な行動観察や情動試験を行うことで総合的な評価ができると考えている。

#### 5. 結語

DLGAP1-KO マウスの表現型解析において、予備的行動観察により恐怖関連行動を示す可能性が示唆された。

#### 6. 文献

- 1) Daniel HG, Jonathan FL. Genetics and genomics of psychiatric disease. *Science*. 2015; 349(6255): 1489-94.
- 2) Myrrhe VS, Casper CH. Synapse pathology in psychiatric and neurologic disease. *Curr Neurol Neurosci Rep*. 2010;10: 207-214.
- 3) Rasmussen AH, Rasmussen HB, et al. The DLGAP family: neuronal expression, function and role in brain disorders. *Mol Brain*. 2017; 4:10(1):43.
- 4) Zhang J, Yang J, Yang N, et al. Dlgap1 negatively regulates browning of white fat cells through effects on cell proliferation and apoptosis. *Lipids Health Dis*. 2020;19(1):39
- 5) Leanne Li, Qiqun Zeng, et al. GKAP Acts as a Genetic Modulator of NMDAR Signaling to Govern Invasive Tumor Growth. *Cancer Cell*. 2018; 33(4): 736–751.
- 6) Masuya H, Inoue M, et al. Implementation of the modified-SHIPRA protocol for screening of dominant phenotypes in a large-scale ENU mutagenesis program. *Genome* 2005;16(11):829-37

日本大学医学部臨床試験研究センター研究紀要  
Bulletin of the Clinical Trials Research Center,  
Nihon University School of Medicine

編集委員会

委員長	高橋泰夫
委員	西山博孝
幹事	葛西玲広

Editorial Board

Editor-in-Chief	Yasuo TAKAHASHI
Associate Editors	Hiroataka NISHIYAMA
Managing Editors	Reo KASAI

---

発行者	日本大学医学部臨床試験研究センター長 浅井 聰
発行所	日本大学医学部 〒173-8610 東京都板橋区大谷口上町 30-1 Tel.03-3972-8111
Publisher	Satoshi ASAI Director, Clinical Trials Research Center, Nihon University School of Medicine 30-1, Ohyaguchi-kamicho, Itabashi-ku, Tokyo 173-8610 JAPAN Phone +81-3-3972-8111

---