

研究課題番号: 4-6

研究課題: データサイエンスと計算論研究の融合による脳病態研究の推進

主任研究者: 本田学 (NCNP 神経研究所・疾病研究第七部 部長)

総括研究報告

1. 研究目的

データサイエンス・AI 技術と脳計算理論を融合的に用いた精神・神経・筋疾患ならびに発達障害の脳病態解明・診断・治療・予防法の開発の推進、およびそのための基盤構築を行う。以下の小項目をそれぞれ具体的な目的とする。

① 脳病態 AI 活用研究推進のための体制・基盤構築

NCNP 病院や NCNP バイオバンク、各研究所に蓄積されたりサーチリソースを最大限に活用した先導的研究を推進するための基盤を整備する。

② AI を活用した脳病態臨床・基礎ビッグデータ解析手法の開発(データ駆動型研究)

遺伝子・分子情報、神経生理学的データ、臨床・疫学情報といった高次元ビッグデータをデータサイエンス技術によって解析し、各水準での特性を反映した特徴量抽出技術を開発するとともに、それらの水準間を架橋する解析技術・理論を開発する。

③ 脳計算理論・統計数理に基づく脳病態モデリング(理論駆動型研究)とデータ駆動・理論駆動の統合的アプローチの開発

神経回路モデルなどを用いて脳計算理論を計算機実装し、精神・神経疾患の病態モデリ

ングを行うことで、脳病態解明と最適な治療的介入手法開発に資する仮説の提案を行う(理論駆動型研究)。さらに提案する脳計算理論と実際の基礎・臨床データとを照合する融合的アプローチを用いることで、精神・神経疾患の新しい疾病概念・個別化医療技術の創出につなげる理論・技術の開発を行う。

2. 研究組織

主任研究者

・本田学 (NCNP 神経研究所・疾病研究第七部)

分担研究者

・星野幹雄 (NCNP 神経研究所・病態生化学研究部)

・西野一三 (NCNP 神経研究所・疾病研究第一部、メディカルゲノムセンター)

・間野達雄 (NCNP 神経研究所・疾病研究第四部)

・山下祐一 (NCNP 神経研究所・疾病研究第七部)

・三浦健一郎 (NCNP 精神保健研究所・精神疾患病態研究部)

・伊藤正哉 (NCNP 認知行動療法センター)

・大庭真梨 (NCNP 病院 臨床研究・教育研修部門 情報管理・解析部)

・宮下聡 (新潟大学脳研究所システム脳病態学分野(～R5.5)、NCNP 神経研究所・病態生化学研究部 (R5.5～))

- ・植木優夫（長崎大学 情報データ科学部）
- ・片平健太郎（産業技術総合研究所 人間情報インタラクション研究部門）

研究協力者

- ・高橋雄太（NCNP 神経研究所・疾病研究第七部）
- ・小島大樹（NCNP 神経研究所・疾病研究第七部）
- ・山口博行（NCNP 神経研究所・疾病研究第七部）
- ・出井勇人（NCNP 神経研究所・疾病研究第七部）
- ・宗田卓史（NCNP 神経研究所・疾病研究第七部）
- ・内田裕輝（NCNP 神経研究所・疾病研究第七部）
- ・小林般（NCNP 神経研究所・疾病研究第七部）

3. 研究成果

① 体制・基盤構築

- (1) 前年度に発足させた脳病態 AI センター (BAIC) 設立準備室（室長：NCNP 企画戦略局長、室長代理：本田）について、NCNP 内に区画を設置し、情報基盤整備の設計立案、必要な設備・人材の確保を行った（本田）。
- (2) センター内外の情報基盤を高速につなぐ 10G ネットワーク回線、および外部からの接続に関するセキュリティ環境を整備した（本田）。
- (3) 情報基盤構築の専門家である特任室長、科研費研究員、および事務助手を採用し、NCNP 病院・バイオバンク・各研究所に散在する大規模情報資産に関するヒアリングを

実施し、センター内情報資産のカatalog化を実現した（本田）。

(4) 世界最大規模の UK Biobank とデータ使用契約を締結し、国際的データ連携体制を整えた（本田）。

(5) NCNP と民間企業 (NTT) とのパートナーシップ協定締結を支援し、産学官連携体制を整えた（本田）。

(6) 隔月で Web セミナー（脳病態数理・データ科学セミナー：計 15 回）を、さらに年 1 回のハンズオンセミナー及びシンポジウム形式の研究会 (CPSY コース東京) を開催し、NCNP 内外のデータサイエンス・AI 研究者に対して、教育のみでなく、研究者同士をつなぐハブとしての機能を果たした（山下）。

② データ駆動型研究

(7) 低悪性度てんかん原生腫瘍検体のメチル化や全 RNA シークエンスを対象に大規模データ解析を実施し、新規サブグループや疾患に関わる分子メカニズムを示唆する結果を得た（星野）。

(8) 筋病理診断アルゴリズムを利用した筋病理鑑別アプリケーションの開発、筋疾患罹患筋の障害度を Mercuri スコアにより分類したデータセットの作成を進めた（西野）。

(9) 高感度ヒストン修飾解析系ならびにフローサイトメトリー解析を用いた細胞核分取技術をもとにパーキンソン病モデルマウスの解析を実施し、病理学的変化に関連した転写因子の変化を認めた（間野）。

(10) 統合失調症、双極性障害、うつ病、自閉スペクトラム症の脳構造画像の大規模コホートデータ解析から、皮質下構造特徴パターンと認知・社会機能のパフォーマンス

によって4つの脳構造バイオタイプに分類できることを示唆した(三浦)。

(11) 心理療法におけるAI技術応用可能性に関する文献調査、ヒアリング、情報交換を文理の枠組みを超えて広く実施した(伊藤)。

(12) 昨年度クリーニングを行った精神疾患レジストリデータのクラスタリング解析を行い、QOLなどのアウトカム指標との関連を検討した(大庭)。

③ 理論駆動型研究・統合的アプローチ

(13) 深層ニューラルネットワークの一種である多モダリティ変分自己符号化器の認知機能モデリングとしての適用可能性の検討を行った。さらに、デジタル脳モデルのプロトタイプとして、マカクザルのECoGデータを使用した脳波生成シミュレータを開発し、リアルタイムな脳状態の推測や脳波信号の生成、仮想的投薬シミュレーションに成功した(山下)。

(14) てんかん患者の手術脳検体を用いたトランスクリプトームデータの一次解析とデータベースへの登録作業、mTORシグナル経路の活性化状態の推定パイプラインの開発を行った(宮下)。

(15) 画像データとゲノムデータに適用できるデータサイエンス手法について理論面から手法開発を進め、関連する技術についてチュートリアル講演を行った(植木)。

(16) これまでに提案した所与のサブタイプ分類の予測の精度を測る指標 Cost of cluster-mean based prediction (CCMP)を、パーソナリティをもとに日常的な運動習慣を予測する問題に適用した。その結果、従来の統計的指標によるサブグループ分類が予測の観点からは最適でない可能性が示唆された

(片平)。

4. 研究成果刊行一覧

【論文】

計21件(うち査読付論文14件)

1. Dang Minh Tran, Nozomu Yoshioka, Norihisa Bizen, Yukiko Mori-Ochiai, Masato Yano, Shogo Yanai, Junya Hasegawa, Satoshi Miyashita, Mikio Hoshino, Junko Sasaki, Takehiko Sasaki, Hirohide Takebayashi, "Attenuated cerebellar phenotypes in Inpp4a truncation mutants with preserved phosphatase activity." *Disease Models & Mechanisms* 16.7 (2023). 査読有
2. Diana Esteller, Marianela Schiava, Rocío-Nur Villar-Quiles, Boris Dibowski, Nadia Venturelli, Pascal Laforet, Jorge Alonso-Pérez, Montse Olive, Cristina Domínguez-González, Carmen Paradas, Beatriz Vélez, Anna Kostera-Pruszczyk, Biruta Kierdaszuk, Carmelo Rodolico, Kristl Claeys, Endre Pál, Edoardo Malfatti, Sarah Souvannanorath, Alicia Alonso-Jiménez, Willem de Ridder, Eline De Smet, George Papadimas, Constantinos Papadopoulos, Sofia Xirou, Sushan Luo, Nuria Muelas, Juan J. Vilchez, Alba Ramos-Fransi, Mauro Monforte, Giorgio Tasca, Bjarne Udd, Johanna Palmio, Srtuhi Sri, Sabine Krause, Benedikt Schöser, Roberto Fernández-Torrón, Adolfo López de Munain, Elena Pegoraro, Maria Elena Farrugia, Mathias Vorgerd, Georgios Manousakis, Jean Baptiste Chanson, Aleksandra Nadaj-Pakleza, Hakan Cetin, Umesh Badrising, Jodi Warman-Chardon, Jorge Bevilacqua, Nicholas Earle, Mario Campero, Jorge Díaz, Chiseko Ikenaga, Thomas E. Lloyd, Ichizo Nishino, Yukako Nishimori, Yoshihiko Saito, Yasushi Oya, Yoshiaki Takahashi, Atsuko Nishikawa, Ryo Sasaki, Chiara Marini-Bettolo, Michela Guglieri, Volker Straub, Tanya Stojkovic, Robert Y. Carlier, Jordi Díaz-Manera, "Analysis of muscle magnetic resonance imaging of a large cohort of patient with VCP-mediated disease reveals characteristic features useful for diagnosis." *Journal of Neurology* 270.12 (2023): 5849-5865. 査

- 読有
3. Reito Nakamura, Ikumi Tomizawa, Atsushi Iwai, Tetsuo Ikeda, Kota Hirayama, Yung Wen Chiu, Takanobu Suzuki, Airi Tarutani, Tatsuo Mano, Atsushi Iwata, Tatsushi Toda, Youhei Sohma, Motomu Kanai, Yukiko Hori, Taisuke Tomita, "Photo - oxygenation of histidine residue inhibits α - synuclein aggregation." *The FASEB Journal* 37.12 (2023): e23311. 査読有
 4. Yuta Takahashi, Shingo Murata, Masao Ueki, Hiroaki Tomita, Yuichi Yamashita, "Interaction between Functional Connectivity and Neural Excitability in Autism: A Novel Framework for Computational Modeling and Application to Biological Data." *Computational Psychiatry* 7.1 (2023). 査読有
 5. Kentaro Katahira, "Evaluating the predictive performance of subtyping: A criterion for cluster mean - based prediction." *Statistics in Medicine* 42.7 (2023): 1045-1065. 査読有
3. エンス、久留米大学バイオ統計センター公開セミナー、福岡、2023年11月
 3. 三浦健一郎、精神疾患における眼球運動所見と臨床応用の可能性、第42回日本精神科診断学会、富山、2023年9月
 4. 伊藤正哉、大江美佐里、竹林由武、藤里紘子、牧野みゆき、山口慶子、高岸百合子、佐藤珠恵、金子響介、堀越勝、心的外傷後ストレス障害の症候学：オンライン調査からの示唆、第22回日本トラウマティック・ストレス学会、東京、2023年8月
 5. 塘由惟、佐藤船斗、大庭真梨、Numerical evaluation of a multiple imputation approach for clustering with missing data. 計算機統計学会第37回シンポジウム、宮崎、2023年11月

【図書】

計2件

1. ステファン・G・ホフマン, スティーブン・C・ヘイズ, デイビット・N・ロールシャイト (著) 菅原大地・榎原潤・伊藤正哉 (監訳) (2023) プロセス・ベースド・セラピーをまなぶ：「心の変化のプロセス」をターゲットとした統合的ビジョン 金剛出版
2. 伊藤正哉・中島俊・久我弘典・蟹江絢子・堀越勝(監修)・アブラモウイツ JS, ディーコン BJ, ホワイトサイド SPH (著) (2023) 不安へのエクスポージャー療法：原則と実践 創元社

【学会発表】

計21件

1. Yamashita Y, Altered hierarchical predictive processing: Exploring psychiatric and neurodevelopmental disorders through a neurorobotics approach, International Symposium on Predictive Brain and Cognitive Feelings, Tokyo, July, 2023. <招待講演>
2. 植木優夫、遺伝統計解析とデータサイ

分担研究報告書

マルチオミックスデータを用いた脳病態解明

星野 幹雄

国立精神・神経医療研究センター
神経研究所 病態生化学研究部

緒言

てんかんの発症・進展機構の理解は不足しており、現在は対症療法が治療の主流である。本研究では、NCNPのバイオリソースに蓄積されているてんかん手術検体から、ゲノム、トランスクリプトーム、プロテオーム情報を取得し、様々なビッグデータ AI 解析を行うことによって、てんかんの細胞内病態仮説を抽出する。さらにそれを、遺伝子改変マウス、遺伝子導入マウス等を用いて検証する。これにより、新たなてんかんの疾患概念の提唱と分類の提案をすることができる。また、細胞内情報伝達におけるハブ因子を同定することで、そのインヒビター、アクチベーターを用いた新たな治療法の開発も目指す。

方法

NCNP バイオバンクのてんかん手術脳検体 50 例余りを用いて total RNA-seq を行うとともに、十数例の検体については、単一核 (sn)RNA-seq を行う。得られたデータをバイオインフォーマティクス解析し、てんかんの病態仮説を抽出する。

結果

(1) LEAT (低悪性度脳腫瘍によるてんかん) について。

48 検体の total RNA-seq データの取得を完了し、大部分の検体について遺伝子変異 (VRAF や FGFR1 など) を同定した。そこか

ら病態に関連する可能性があるパスウェイを抽出し、論文として病態仮説を発表した (Iijima et al)。

(2) FCD (限局性皮質異形成症) について
コントロール 5 検体、FCD6 検体について snRNAseq を完了した。クラスタリングによりそれぞれの細胞種に分類し、その細胞種ごとの Differentially Expressed Genes (DEGs) を抽出した。更なる解析により、神経細胞、アストロサイト、オリゴデンドロサイトがそれぞれ病態に関連するという仮説を得つつある。

考察

LEAT に関しては一つ論文を出すことができた。脳腫瘍が原因であるてんかんに比べて、FCD はさらに病態が複雑と考えられる。得られたデータをさらに細かく解析することで、より有力な病態仮説を得るようにする。その上で、モデルマウスを用いてその仮説を検証する必要がある。

結論

NCNP バイオバンクの貴重な手術脳検体を用いて、近年の技術進歩によって可能となるマルチオミックス解析を行うことで、大きなブレークスルーが果たされるものと期待している。

参考文献

1. Dewa K, Arimura N, Kakegawa W, Itoh M, Adachi T, Miyashita S, Inoue YU, Hizawa K, Hori K, Honjaya N, Yagishita H, Taya S, Miyazaki T, Usui C, Tatsumoto S, Tsuzuki A, Uetake H, Sakai K, Yamakawa K, Sasaki T, Nagai J, Kawaguchi Y, Sone M, Inoue T, Go Y, Ichinohe N, Kaibuchi K, Watanabe M, Koizumi S, Yuzaki M, Hoshino M. Neuronal DSCAM regulates the peri-synaptic localization of GLAST in Bergmann glia for functional synapse

- formation. *Nat Commun.* 15(1):458, 2024
2. Iijima K, Komatsu K, Miyashita S, Suyama K, Murayama K, Hashizume K, Tabe NK, Miyata H, Iwasaki M, Taya S, Hoshino M. Transcriptional features of low-grade neuroepithelial tumors with the BRAF V600E mutation associated with epileptogenicity. *Genes Cells.* 2024 Mar;29(3):192-206.

AIを用いた筋疾患診断支援システム開発 分担研究者

西野 一三

国立精神・神経医療研究センター

神経研究所 疾病研究第一部

緒言

AIを活用する事で、専門家の少ない筋疾患診断支援ツールの開発を目指す。このようなツールの開発を行うことで、地球規模での筋疾患医療均てん化に寄与することを目指す。具体的には、1) 筋病理診断支援アルゴリズムおよび 2) 画像診断補助ツールの開発を目指す。1) についてはこれまでに開発したアルゴリズムを更に進化させるべく、他施設データを使用した場合の診断率低下の克服を目指す。2) については、各種筋疾患について罹患筋の障害度を **Mercuri score** により分類したデータセットを作成し、これに対して **random forest supervised machine learning** を用いたモデル作成を目指す。

我々は筋病理 HE 染色標本を CCD カメラで取り込んだ画像を用い、筋炎と遺伝性筋疾患を、さらには、筋炎を免疫介在性壊死性ミオパチー、皮膚筋炎、抗合成酵素諸症候群、封入体筋炎に、遺伝性筋疾患をジストロフィノパチー、ジスフェルリノパチー、カルバインパチー、先天性ミオパチーなどに高精度で分類できることを示した (Lab Invest 2022)。これは世界で初めての AI を活用した筋病理診断アルゴリズムである。筋病理を評価するアルゴリズムを開発したのは、世界で我々のグループのみである。本研究課題においては、この成果をさらに発展させて、より汎用性の高いアルゴリズムへと進化させることを目指す。また筋画像をも

ちいた診断アルゴリズムの開発についても検討を行う。これらのアルゴリズム開発には充実したデータセットが不可欠であるが、NCNP は世界最高峰の筋レポジトリを有しており、課題達成に極めて有利な条件を備えている。

方法

1) 筋病理診断支援アルゴリズム開発

これまでに開発したアルゴリズムを更に進化させるべく、構築を進めてきた筋炎判別アプリを AWS 上で稼働させ、他施設の筋病理医に利用してもらうことで、現在の筋炎判別モデルとそのアプリに対する評価を客観的に受け取り、改善する。また他施設でのデータを使用した場合の診断率低下の克服を目指す。

2) 画像診断補助ツール開発

各種筋疾患について罹患筋の障害度を **Mercuri score** により分類したデータセットを作成し、これに対して **random forest supervised machine learning** を用いたモデル作成を目指す。

3) 画像診断学習ツール開発

骨格筋画像診断のためには骨格筋の解剖学的知識が必須である。筋疾患診療に関わる医師が骨格筋の解剖を学習するためのアプリを奈良先端科学技術大学院大学との共同研究で開発することとした。

結果

1) 筋病理診断支援アルゴリズム開発

他施設で筋病理診断支援アルゴリズムを使用できる様、安全で信頼性の高い AWS 環境

を構築し、運用テストを行った。また本研究について、2023年1月倫理承認を得た。

2) 画像診断補助ツール開発

眼咽頭遠位型ミオパチー (OPDM) について罹患筋の障害度を Mercuri score により分類したデータセットを作成した。LRP 遺伝子、GIPC 遺伝子、NOTCH2NLC 遺伝子のいずれの遺伝子を原因遺伝子とした OPDM も同様の画像パターンを呈した。また眼咽頭筋ジストロフィー (OPMD) との鑑別には中殿筋、大殿筋、大内転筋が有用であることを明らかにし、論文として報告した。

3) 画像診断学習ツール開発

健常者の全身のマルチスライス CT データを用いて、個々の筋名をアノテーションしたアプリのプロトタイプが完成した。

考察・結論・今後の展望

- 1) AWS 環境は構築され、他施設での運用テストが可能な状況が整った。限定されたユーザーに公開、検証とデータ収集を進めるためのオリエンテーションを調整中である。
- 2) OPDM・OPMD 以外の筋疾患についても各筋の Mercuri score のデータ収集を行っており、それぞれについて論文投稿を行う予定である。そのデータを用いて random forest supervised machine learning を用いたモデル作成を目指す。
- 3) アプリ上の骨格筋分離とアノテーションはまだ不完全であることから、その修正を行うべく、奈良先端科学技術大学院大学側と協議を重ねる。完成後は疾病研究第一部からアプリにアクセスできる

ようにする。

参考文献

1. Eura N, Noguchi S, Ogasawara M, Kumutpongpanich T, Hayashi S, Nishino I; OPDM/OPMD Image Study Group: Characteristics of the muscle involvement along the disease progression in a large cohort of oculopharyngodistal myopathy compared to oculopharyngeal muscular dystrophy. *J Neurol.* 2023 Dec;270(12): 5988-5998.
2. 大久保真理子, 壁谷佳典, 西野一三: 「人工知能 (AI) による筋病理判読アルゴリズム」の開発. *Brain Nerve.* 74(8): 1019-1024, Aug, 2022
3. 西野一三: AI を用いた筋病理診断アルゴリズム. 第 63 回日本神経学会学術大会 (東京国際フォーラム), 千代田区, 5.20, 2022 (5.18-5.21)
4. Kabeya Y, Okubo M, Yonezawa S, Nakano H, Inoue M, Ogasawara M, Saito Y, Tanboon J, Indrawati LA, Kumutpongpanich T, Chen YL, Yoshioka W, Hayashi S, Iwamori T, Takeuchi Y, Tokumasu R, Takano A, Matsuda F, Nishino I: Deep convolutional neural network-based algorithm for muscle biopsy diagnosis. *Lab Invest.* 102(3):220-226. Mar, 2022
5. Kabeya Y, Iwamori T, Yonezawa S, Takeuchi Y, Nakano H, Nagisa Y, Okubo M, Inoue M, Tokumasu R, Ozawa I, Takano A, Nishino I: AI-based muscle histopathologist can differentiate major muscular dystrophies better than human. 24th International Congress of the World Muscle Society (Tivoli Concert Hall), Copenhagen, Denmark, 10.4, 2019(10.1-5)

神経変性過程における包括的エピゲノムアプローチ

間野 達雄

国立精神・神経医療研究センター

神経研究所 疾病研究第四部

緒言

パーキンソン病において血管リスクが発症リスクを減らすという疫学データがあり、その現象の分子メカニズムを理解するために、シヌクレインの線維状態へのインジェクションとマイクロコイルを用いた総頸動脈狭窄のマウスモデル実験を行った。前年度の研究では、シヌクレイン病理の改善を確認し、神経変性過程における神経細胞やグリア細胞のエピゲノム解析の重要性が示唆された。孤発性神経変性疾患は遺伝的背景と後天的要因が複雑に関連し合って病態を形成しているため、エピゲノム変化から病態を理解することが重要である。従来のエピゲノム解析は脳全体を用いた研究がなされてきたが、個々の細胞において異なる反応が起きているのに対して、全体を用いた解析は平均化してそれぞれの変化への感度が乏しい。そこで、本年度は、マイクロコイルを用いた総頸動脈狭窄の影響を調べるために、神経細胞特異的 ATAC-seq の解析を行い、ゲノムワイドに起きている現象を明らかにすることを目的とした。

方法

本研究では、シヌクレインの線維状態をマウスの脳内に注入し、マイクロコイルを用いて総頸動脈を狭窄させたマウスモデルを作成した。次に、マウス脳から神経細胞を単離し、ATAC-seq を行ってオープン

クロマチン領域を同定した。得られた ATAC-seq データを用いて Footprint 解析を行い、転写因子の結合部位を同定した。さらに、HIF1A の機能を検証するために、細胞モデルを用いて HIF1A 依存的オートファジーの評価を行った。

結果

神経細胞特異的 ATAC-seq 解析の結果、マイクロコイルを用いた総頸動脈狭窄により影響を受けた神経細胞のオープンクロマチン領域が同定された。Footprint 解析により、HIF1A が総頸動脈狭窄の影響を受けた神経細胞において重要な役割を果たしていることが明らかになった。さらに、細胞モデル実験の結果、HIF1A 依存的オートファジーがシヌクレイン病理の改善に関与していることが示された。

考察

本研究の結果から、マイクロコイルを用いた総頸動脈狭窄が HIF1A の活性化を介して神経細胞のエピゲノム状態に影響を与え、HIF1A 依存的オートファジーを誘導することでシヌクレイン病理を改善する可能性が示唆された。これは、血管リスクがパーキンソン病の発症リスクを減らすメカニズムの一端を説明するものと考えられる。本研究では、神経細胞特異的なエピゲノム解析を行うことで、従来の脳全体を用いた解析では見落とされていた細胞特異的な変化を捉えることができた。このアプローチは、孤発性神経変性疾患の病態理解に有用であると考えられる。

結論

神経細胞特異的 ATAC-seq 解析により、マイクロコイルを用いた総頸動脈狭窄が HIF1A を介してシヌクレイン病理を改善する可能性が示された。この結果は、血管リスクがパーキンソン病の発症リスクを減らすメカニズムの一端を説明するものであり、孤発性神経変性疾患の病態理解に役立つと考えられる。今後は、HIF1A 依存的オートファジーの詳細なメカニズムを解明し、パーキンソン病の新たな治療戦略の開発につなげていきたい。

参考文献

1. 間野達雄、岩田淳、医歯薬出版、医学のあゆみ、エピジェネティクスとアルツハイマー病、2023年12月、5ページ

脳計算理論を用いた精神疾患病態メカニズムの解明

山下 祐一

国立精神・神経医療研究センター

神経研究所 疾病研究第七部

緒言・背景

現行の精神障害カテゴリーは、生物学的基盤の非特異性と異種性のため、その生物学的妥当性に疑問が投げかけられている。近年、数理・データ科学を用いた研究手法が注目され、精神医学が直面するこの問題を解決する上で極めて強力な研究方略を提供することが期待されている。実際、精神医学に数理・データ科学を用いた研究報告も増加しているが、それらの多くが依然として現行の疾病カテゴリー分類に基づいて実施されている。特に、データ駆動型の研究では、症状、神経生理学、認知行動など、水準の相関を明らかにするものの、そのメカニズムになかなか迫れないという困難がある。これに対して、本研究では、疾患横断的・次元的アプローチに基づいて、データ駆動と理論駆動を用いた統合的アプローチを試みる。

研究の目的

本研究の目的は、精神障害に関する症状・神経生理・認知行動のビッグデータに対して、データ科学・機械学習技術を用いたデータ駆動的な手法と、脳の計算理論に基づく理論駆動的な手法を組み合わせることで、精神障害の統合的理解と新しい疾病概念を創出するための基盤技術を開発することである。

研究の方法

データサイエンス・機械学習技術を用いた「データ駆動」的手法として、具体的には、深層ニューラルネットワークを用いた、教師なし特徴表現学習手法（変分自己符号化器など）を用いた特徴量抽出法を探索的に検討する。同時に、「理論駆動」的手法として、予測情報処理の脳の計算理論に基づいて、行動や神経生理学的所見を説明する病態モデルを提案し、個別データの違いをモデルのパラメータ・特徴量として表現することを試みる。

得られた観察に基づいて、既存の精神障害カテゴリーにとらわれない精神障害の統合的理解と新しい疾病概念を創出することを目指す。

結果

データ駆動的な手法として、深層ニューラルネットワークの一種である、マルチモーダル変分自己符号化器を用いた特徴表現学習手法と、認知機能モデリングとしての適用可能性の検討を行った。マルチモーダル学習では、2つのモダリティ間の共通性を反映した、優れた潜在表現を獲得できることが明らかになった。さらに、シンボリック情報と、オブジェクト情報を用いたマルチモーダル学習は、順序構造やクラスタ構造など、数感覚に関連した潜在表現の獲得や、算術計算課題においても有益であった。このように、マルチモーダル学習は、情報の潜在表現の学習と、関連する下流タスクの成績に重要であることが明らかになった[1]。

理論駆動的な手法として、予測情報処理モデルを用いた、ニューラルネットワークモデリングと、実際にヒト被験者の行動データを用いた検証を行った。具体的には、予測

情報処理における“予測シグナルの遅延”が、統合失調症における意思作用感の障害に帰結する、との仮説をシミュレーションと実際の患者の行動データを用いて検証した[2]。

考察

本研究により、大規模なデータから深層学習技術を用いての潜在表現獲得、特徴量抽出に関する重要な知見を得た。また、理論駆動のモデリングが、神経生理・認知行動の知見に基づく仮説の検証に有効であることが示された。これらの成果は、データ科学・計算理論を用いた精神障害の新しい研究手法の開発に貢献し、従来の疾患カテゴリーに基づく精神医療では困難であった、患者個々人の特性に応じた病状評価・疾患経過や治療反応性予測といったオーダーメイド医療の進歩に貢献する可能性がある。

主な発表論文等

1. Wang C, Tachimori H, Yamaguchi H, Sekiguchi A, Li Y, Yamashita Y (2024) A multimodal deep learning approach for the prediction of cognitive decline and its effectiveness in clinical trials for Alzheimer's disease. *Translational Psychiatry*, 14(1), 105.
2. Noda K, Soda T, Yamashita Y (2024) Emergence of number sense through the integration of multimodal information: developmental learning insights from neural network models. *Front. Neurosci.* 18:1330512.
3. Okimura T, Maeda T, Mimura M, Yamashita Y (2023) Aberrant sense of

agency induced by delayed prediction signals in schizophrenia: a computational modeling study. *npj Schizophr* 9, 72.

4. Idei H, Yamashita Y (2024). Elucidating multifinal and equifinal pathways to developmental disorders by constructing real-world neurorobotic models. *Neural Networks*, 169, 57–74.

精神神経疾患の中間表現型情報による分類・層別化の研究

三浦 健一郎

国立精神・神経医療研究センター

精神保健研究所 精神疾患病態研究部

緒言

精神疾患の脳病態はまだ明らかになっておらず、そのためには臨床検査などで取得された中間表現型ビッグデータの解析が必須である。しかし、精神疾患の多次元データから診断や治療に有益な情報を得る試みは始まったばかりであり、そのための確立した方法はまだない。本研究では、精神神経疾患の脳病態研究を推進するための基盤整備に資するため、精神疾患の包括的なリサーチリソースデータベースに集積された中間表現型データを用いて、データサイエンスを活用した疾患の分類・層別化のための方法を開発する。

方法

ヒト脳表現系コンソーシアムデータベースに集積されたデータのクオリティチェックを行い、使用可能なデータを抽出する。次いで健常者および患者から収集された特徴量を探索・算出し、それを要素として含む多次元中間表現型データを構成する。その多次元データを用いて、多変量解析手法及び機械学習法を適用し、精神疾患の分類・層別化を試みる。

結果

R5年度は前年度に引き続き、中間表現型データのクオリティチェックを行い、データベースの整備を進めつつ、眼球運動及び

脳神経画像の解析を進め、下記の成果を得た。統合失調症、双極性障害、うつ病、自閉スペクトラム症の脳構造画像の大規模コホートデータを用いたデータ駆動型解析から、皮質下構造特徴パターンと認知・社会機能のパフォーマンスによって4つの脳構造バイオタイプに分類できることを示した[1]。大脳皮質の皮質厚、表面積を特徴量とした疾患横断的解析から、統合失調症、双極性障害、うつ病、自閉スペクトラム症の間に、皮質厚と表面積とで異なる類似性と相違点があることを明らかにした[2]。脳体積特徴により定義される統合失調症らしさの縦断的解析から脳体積の変化と認知機能の変化が関連することを示した[3]。統合失調症の眼球運動に視覚的サリエンス処理の異常があること [4]及び精神疾患間の比較において統合失調症が最も顕著であることを明らかにした。

考察

皮質下構造と認知・社会機能の組み合わせは精神疾患の客観的データに基づく分類・層別化に有用であることが示唆された。また、大脳皮質の構造、脳体積から得られる特徴量および眼球運動における視覚的サリエンス特徴などは、病態に基づく客観的な疾患分類に有用である可能性があり、その検討が今後の課題である。

結論

眼球運動、認知機能及び脳構造などの中間表現型情報は精神疾患の分類と層別化において有用な多次元特徴量なることが期待される。

参考文献

- [1] Okada N, Fukunaga M, Miura K, Nemoto K, Matsumoto J, Hashimoto N, Kiyota M, Morita K, Koshiyama D, Ohi K, Takahashi T, Koeda M, Yamamori H, Fujimoto M, Yasuda Y, Hasegawa N, Narita H, Yokoyama S, Mishima R, Kawashima T, Kobayashi Y, Sasabayashi D, Harada K, Yamamoto M, Hirano Y, Itahashi T, Nakataki M, Hashimoto RI, Tha KK, Koike S, Matsubara T, Okada G, van Erp TGM, Jahanshad N, Yoshimura R, Abe O, Onitsuka T, Watanabe Y, Matsuo K, Yamasue H, Okamoto Y, Suzuki M, Turner JA, Thompson PM, Ozaki N, Kasai K, Hashimoto R, Subcortical volumetric alterations in four major psychiatric disorders: A mega-analysis study of 5604 subjects and a volumetric data-driven approach for classification, *Mol Psychiatry*, 2023, doi:10.21203/rs.3.rs-2182255/v1
- [2] Matsumoto J, Fukunaga M, Miura K, Nemoto K, Okada N, Hashimoto N, Morita K, Koshiyama D, Ohi K, Takahashi T, Koeda M, Yamamori H, Fujimoto M, Yasuda Y, Ito S, Yamazaki R, Hasegawa N, Narita H, Yokoyama S, Mishima R, Miyata J, Kobayashi Y, Sasabayashi D, Harada K, Yamamoto M, Hirano Y, Itahashi T, Nakataki M, Hashimoto RI, Tha KK, Koike S, Matsubara T, Okada G, Yoshimura R, Abe O, van Erp TGM, Turner JA, Jahanshad N, Thompson PM, Onitsuka T, Watanabe Y, Matsuo K, Yamasue H, Okamoto Y, Suzuki M, Ozaki N, Kasai K, Hashimoto R, Cerebral cortical structural alteration patterns across four major psychiatric disorders in 5549 individuals, *Mol Psychiatry*, 2023, doi:10.1038/s41380-023-02224-7
- [3] Yamazaki R, Matsumoto J, Ito S, Nemoto K, Fukunaga M, Hashimoto N, Kodaka F, Takano H, Hasegawa N, Yasuda Y, Fujimoto M, Yamamori H, Watanabe Y, Miura K, Hashimoto R. Longitudinal reduction in brain volume in patients with schizophrenia and its association with cognitive function. *Neuropsychopharmacol Rep*. 2024;44(1):206-215. doi: 10.1002/npr2.12423
- [4] Yoshida M, Miura K, Fujimoto M, Yamamori H, Yasuda Y, Iwase M, Hashimoto R, Visual salience is affected in participants with schizophrenia during free-viewing., *Sci Rep*, 2024, 14, 4606, doi:10.1038/s41598-024-55359-0

認知行動療法に関する人工知能技術の応用可能性の検討

伊藤正哉¹、重枝裕子¹、矢部魁一¹、杉田創¹、久我弘典¹

¹ 国立精神・神経医療研究センター認知行動療法センター

諸言

心理療法の研究領域において、人工知能技術が積極的に活用されつつある (Koutsouleris, 2022)。当センター主導により精神疾患レジストリが構築されてきたが、その一方で、心理療法に関するビッグデータの構築については今後の開発が期待されている。また、心理療法中のコミュニケーションにおいて、マルチモーダルでのセンシングデータを収集する手法も提案されつつある。産業分野でも、メンタルヘルステックへの巨額の資本が投入され、デジタル技術を応用した心理ケアのサービス構築が進んでいる。前年度までの整理を踏まえ、本年度は新たに公表された文献を整理した。

方法

Pubmed 等の文献データベース、及びウェブ上で関連する記事や報告書を調査し、心理療法に対して人工知能技術を適用する上で参考となる研究のレビューを行った。

結果

文献の動向 2022年5月に Pubmed において "psychotherapy" AND "artificial intelligence" と検索した際には 141 件がヒットし、2016 年まで年間 5 本以下だったが、2022 年には 55 本と 10 倍に増えてい

た。2023 年 5 月に同条件で検索すると、343 件がヒットし、この 1 年で 200 本の論文が公表されている。

精神状態の識別や診断補助に関する AI の適用 メンタルヘルスに関連する研究において、データ駆動型の AI 技術を用いて、大量のデータから特定の感情状態や精神症状を識別させようと試みる研究が複数報告されている。例えば、Lalk らは、124 名の患者における 552 の心理療法トランスクリプトから、自己記入式尺度で回答された治療関係や症状の重症度を判別する試みが報告されている (Lalk et al., 2024)。この研究では、抽出されたトピックと症状 ($r=0.45$, 95%-CI 0.40, 0.51) や治療関係 ($r=0.20$, 95%-CI 0.16, 0.24) との相関を見出している。同様に、心理療法における自然言語から治療関係を測定する試みがなされており、第一人称の使用法が治療関係に関与することが示唆されている (Ryu et al., 2023)。

人間が関与しないという特性を活かして、すでに運用されているサービスで挙げている成果も報告されている (Habichit et al., 2024)。この研究では、イングランドの National Health Services における約 130000 名の患者を対象とした観察研究において、AI による自己リファラル Chatbot システムの有効性を検討している。その結果、対照したサービスでは 6% の自己リファラル上昇であったところ、AI を用いた介入群では 15% の上昇が観察された。特に、ノンバイナリーでは 179%、人種マイノリティにおいては 29% の上昇が認められたことから、人間が介在しないリファラルの有効性が示唆されている。

他に、身体の姿勢や動作から急性ストレスの状態を判別する研究が報告されている (Richer et al., 2024)。

こうした研究の他、サイコーシスの臨床診断補助として自然言語に基づく臨床意思決定支援システムを仮定し、その有用性や懸念を専門家らが議論したワークショップの要約が報告されている (Corona Hernández et al., 2023)。ここでは自然言語をマーカーとした検出、モニタリング、予測の手法を期待する議論がされる一方、その倫理面での問題も指摘されている。

AIを活用した介入 AIによるチャットボットを通して認知行動療法を提供する試みも進んでいる。2024年に公表されたメタ分析では、18のランダム化比較試験 ($n = 3477$)の結果、うつ ($g = -0.26, 95\% \text{ CI} = -0.34, -0.17$) と不安 ($g = -0.19, 95\% \text{ CI} = -0.29, -0.09$) に対する改善が報告されている (Zhong et al., 2024)。

また、空間コンピューティング、ヴァーチャル・リアリティ、大規模言語モデルを組み合わせ、VR上で心理療法を実施する Ceders-sinai 病院の試みも紹介されている (Spiegel et al., 2024)。同様のVRとAI補助の心理療法として、ドイツではマルチモーダルニューロフィードバック (fNIRS や EEG) を組み合わせたシステムの開発を紹介している (Klein et al., 2023)。

AIによる心理療法は様々なかたちを取る。その中で、人間が実施する心理療法をAIが代行する点で必要となる要件を整理した文献では、①人間が実施する心理療法のメカニズムが解明されていること、②人間が実施する心理療法では治療関係の構築が重要であること、③汎用AIが必要とさ

れうることを挙げている (Grodiewicz and Hohol, 2023)。

大規模言語モデルの有用性と懸念 Open AI社のGPT-4に代表される大規模言語モデルのヘルスケアへの活用が期待されている (Cheng et al., 2023)。行動ヘルスケアの分野での活用に関して、Stadeら (2024)が展望論文を公表しており、自動運転技術の比喩を用いて、心理療法補助AI、心理療法協働AI、完全自動心理療法AIの段階についてまとめている。心理療法補助AIとしては、感情状態の識別を高い精度で (人間の専門家に匹敵する精度で) ChatGPTが達成した報告がされている (Elyoseph et al., 2023)。

大規模言語モデルについては、それ以前のAIと同様に説明可能性に関する問題が懸念されており、トラストできるAIの開発や提供のあり方に関しての整理の必要性が指摘されている。同様に、生成AIによって誤ったメンタルヘルス関連の情報が広まることに関する懸念も示されている (Monteith et al., 2024)。そのユースケースとして、摂食症に対する治療ツールとしてAIサービスを想定した場合の可能性について議論されている (Abbade-Daga et al., 2023)。

結論

心理療法へのAI適用が進んでいるが、論文として報告されている知見はそのごく一部であると思われる。AIを活用した心理療法に関する Ethical, Legal, Social Issue に関する議論も十分とは言えず、今後は研究開発の中心を患者自身に置くことの重要性が指摘されており (Turchioe et al.,

2024)、さらなる研究が必要である

(Espejo et al., 2023)。

心理療法研究分野においては、実践に基づく研究 (practice-oriented research) というアプローチが提唱されている。こうしたアプローチが根付くことによって、大量の心理療法データを活用した AI 研究が前進していくと期待される (Atzil-Slonim et al., 2024; Castonguay et al., 2024)。AI による開発にはデータの収集や利活用上での規制のあり方や個人情報の保護の問題も重要であり、今後の検討の余地が大きく残されていると考えられる。

レビューされた文献

- Abbate-Daga, G., Taverna, A., & Martini, M. (2023). The oracle of Delphi 2.0: considering artificial intelligence as a challenging tool for the treatment of eating disorders. *Eating and weight disorders : EWD*, 28(1), 50.
<https://doi.org/10.1007/s40519-023-01579-8>
- Atzil-Slonim, D., Penedo, J. M. G., & Lutz, W. (2023). Leveraging Novel Technologies and Artificial Intelligence to Advance Practice-Oriented Research. *Administration and policy in mental health*, 10.1007/s10488-023-01309-3. Advance online publication.
<https://doi.org/10.1007/s10488-023-01309-3>
- Castonguay, L. G., Atzil-Slonim, D., de Jong, K., & Youn, S. J. (2024). Practice-Oriented Research: An Introduction to New Developments and Future Directions. *Administration and policy in mental health*, 10.1007/s10488-024-01369-z. Advance online publication.
<https://doi.org/10.1007/s10488-024-01369-z>
- Cheng, S. W., Chang, C. W., Chang, W. J., Wang, H. W., Liang, C. S., Kishimoto, T., Chang, J. P., Kuo, J. S., & Su, K. P. (2023). The now and future of ChatGPT and GPT in psychiatry. *Psychiatry and clinical neurosciences*, 77(11), 592–596.
<https://doi.org/10.1111/pcn.13588>
- Corona Hernández, H., Corcoran, C., Achim, A. M., de Boer, J. N., Boerma, T., Brederoo, S. G., Cecchi, G. A., Ciampelli, S., Elvevåg, B., Fusaroli, R., Giordano, S., Hauglid, M., van Hessen, A., Hinzen, W., Homan, P., de Kloet, S. F., Kooops, S., Kuperberg, G. R., Maheshwari, K., Mota, N. B., ... Palaniyappan, L. (2023). Natural Language Processing Markers for Psychosis and Other Psychiatric Disorders: Emerging Themes and Research Agenda From a Cross-Linguistic Workshop. *Schizophrenia bulletin*, 49(Suppl_2), S86–S92.
<https://doi.org/10.1093/schbul/sbac215>
- Elyoseph, Z., Hadar-Shoval, D., Asraf, K., & Lvovsky, M. (2023). ChatGPT outperforms humans in emotional awareness evaluations. *Frontiers in psychology*, 14, 1199058.
<https://doi.org/10.3389/fpsyg.2023.1199058>
- Espejo, G., Reiner, W., & Wenzinger, M. (2023). Exploring the Role of Artificial Intelligence in Mental Healthcare: Progress, Pitfalls, and Promises. *Cureus*, 15(9), e44748.
<https://doi.org/10.7759/cureus.44748>
- Grodniewicz, J. P., & Hohol, M. (2023). Waiting for a digital therapist: three challenges on the path to psychotherapy delivered by artificial intelligence. *Frontiers in psychiatry*, 14, 1190084.
<https://doi.org/10.3389/fpsyg.2023.1190084>
- Habicht, J., Viswanathan, S., Carrington, B., Hauser, T. U., Harper, R., & Rollwage, M. (2024). Closing the accessibility gap to mental health treatment with a personalized self-referral chatbot. *Nature medicine*, 30(2), 595–602.
<https://doi.org/10.1038/s41591-023-02766-x>
- Klein, F., Müller-Von Aschwege, F., Elfert, P., Råker, J., Philipsen, A., Braun, N., Selaskowski, B., Wiebe, A., Guth, M., Spallek, J., Seuss, S., Storey, B., Geppert, L. N., Lück, I., & Hein, A. (2023). Developing Advanced AI Ecosystems to Enhance Diagnosis and Care for Patients with Depression. *Studies in health technology and informatics*, 309, 18–22.
<https://doi.org/10.3233/SHTI230731>
- Lalk, C., Steinbrenner, T., Kania, W., Popko, A., Wester, R., Schaffrath, J., Eberhardt, S., Schwartz, B., Lutz, W., & Rubel, J. (2024). Measuring Alliance and Symptom Severity in Psychotherapy Transcripts Using Bert Topic Modeling. *Administration and policy in mental health*, 10.1007/s10488-024-01356-4. Advance online publication.

- <https://doi.org/10.1007/s10488-024-01356-4>
Monteith, S., Glenn, T., Geddes, J. R., Whybrow, P. C., Achtyes, E., & Bauer, M. (2024). Artificial intelligence and increasing misinformation. *The British journal of psychiatry : the journal of mental science*, *224*(2), 33–35.
<https://doi.org/10.1192/bjp.2023.136>
- Richer, R., Koch, V., Abel, L., Hauck, F., Kurz, M., Ringgold, V., Müller, V., Küderle, A., Schindler-Gmelch, L., Eskofier, B. M., & Rohleder, N. (2024). Machine learning-based detection of acute psychosocial stress from body posture and movements. *Scientific reports*, *14*(1), 8251.
<https://doi.org/10.1038/s41598-024-59043-1>
- Ryu, J., Heisig, S., McLaughlin, C., Katz, M., Mayberg, H. S., & Gu, X. (2023). A natural language processing approach reveals first-person pronoun usage and non-fluency as markers of therapeutic alliance in psychotherapy. *iScience*, *26*(6), 106860.
<https://doi.org/10.1016/j.isci.2023.106860>
- Spiegel, B. M. R., Liran, O., Clark, A., Samaan, J. S., Khalil, C., Chernoff, R., Reddy, K., & Mehra, M. (2024). Feasibility of combining spatial computing and AI for mental health support in anxiety and depression. *NPJ digital medicine*, *7*(1), 22.
<https://doi.org/10.1038/s41746-024-01011-0>
- Turchioe, M. R., Hermann, A., & Benda, N. C. (2024). Recentering responsible and explainable artificial intelligence research on patients: implications in perinatal psychiatry. *Frontiers in psychiatry*, *14*, 1321265.
<https://doi.org/10.3389/fpsy.2023.1321265>
- Zhong, W., Luo, J., & Zhang, H. (2024). The therapeutic effectiveness of artificial intelligence-based chatbots in alleviation of depressive and anxiety symptoms in short-course treatments: A systematic review and meta-analysis. *Journal of affective disorders*, *356*, 459–469. Advance online publication.
<https://doi.org/10.1016/j.jad.2024.04.057>

レジストリに蓄積される臨床ビッグデータを活用した予測モデル開発と精度評価

大庭 真梨

国立精神・神経医療研究センター

病院 臨床研究・教育研修部門

緒言

本研究では、臨床症状を中心に経時的に収集したレジストリビッグデータに対してデータサイエンス手法を適用し、特性を共有する集団の探索やイベント予測モデルの開発、およびその評価を行う。予測モデルとして統計モデル等を用い、精神・神経疾患、発達障害の病態モデリングを行うことで、脳病態解明や予測、介入手法開発に資する解析手法の提案を行うことを目標とする。

3年間で (1) 時系列データ解析手法と層化・予測モデル方法の情報収集（文献的検討）、(2) 精神疾患レジストリ等の実データへの適用 (3) 成果発表を目指す。

方法

前年度の医学研究における予測モデル構築の考え方の整理、予測モデルの文献的検討を踏まえ、精神疾患レジストリデータに対してクラスタリングを行った。具体的には、心理評価尺度の欠測を統計手法により補い、それらの多変量データに基づいてアンサンブルクラスタリングを行い、QOL等のアウトカム指標との関連を検討した。ベースラインデータの欠測や追跡不能の発生状況を整理し、結果に与える影響の評価を開始した。ランダムな欠測を仮定したクラスタリングについてシミュレーション実験を実施中である。

結果

精神疾患レジストリの同意取得数は2023年末の時点で2300件、約1200人の情報が入力済みで追跡割合は6割であった。欠測は殆どの対象者に存在し、欠測補完を行わない解析では解析データ数が4割に減少した。

欠測に対して連鎖方程式による多重代入法を用いた補完を行った。文献調査並びにシミュレーション実験を通じて、単一クラスタリングよりもアンサンブルクラスタリングの方が、性能が良いことが確認された。既存のアンサンブルクラスタリングについてシミュレーション実験による性能比較を行い、最も真のクラスターとの一致割合が高い手法を特定した。

考察

レジストリデータの欠測に対する多重補完法とアンサンブルクラスタリングは相性がよく、欠測を補いながら対象集団を分類し、解釈可能な結果を得ることができた。

今後はクラスタリング結果と、他の情報や経時データの関連を精査し、予後予測に資するクラスタリングか否かを検討する。

ランダムな欠測を仮定した多重補完の妥当性、欠測の生じ方によるクラスタリングの性能についてはまだ確認すべき点が多く、今後もレジストリデータにより近いシミュレーションデータを発生させ性能評価実験、あるいは手法上の拡張研究を継続する。

結論

欠測を有するデータを補完するとともに、クラスタリングの実用と手法に関する

性能評価を行った。今後は性能評価を継続するとともに成果発表を目指す。また、経時変化に対する予測についても検討を継続する。

学会発表

塘由惟, 佐藤船斗, 大庭真梨. Numerical evaluation of a multiple imputation approach for clustering with missing data. 計算機統計学会第 37 回シンポジウム、宮崎、2023 年 11 月 11 日、口頭発表

機械学習を用いたオミクスビッグデータの解析

宮下 聡

新潟大学脳研究所 システム脳病態学分野
(～R5.5)

国立精神・神経医療研究センター

神経研究所 病態生化学研究部 (R5.5～)

緒言

シングルセル RNA シーケンシング技術に代表される、近年のシーケンス技術の飛躍的な技術革新の結果、今や医学・生物学の研究手法は大きく変容した。一方で、シーケンシングの結果得られるビッグデータから、真に生物学的に重要な知見を引き出すためには、ビッグデータ解析技術のさらなる向上が必要である。特に、シングルセル RNA シーケンシングに特有の問題である dropout や sparse 性といった技術的な問題によって、正確性の高い 2 群間の遺伝子発現の違いや希少細胞クラスターの同定は現状では困難であり、さらなる技術開発が求められている。こうした諸問題に対して、近年は、機械学習を適用して解決を図ろうとする試みが多くなされてきている。本研究では、特に、シングルセル RNA シーケンシングから得られるビッグデータを用いて、オートエンコーダーや深層学習のアルゴリズムで学習を行うことにより、従来の手法では同定されなかった新規細胞集団や、精神・神経疾患の原因となる遺伝子の抽出を行う。この技術開発により、生物学ビッグデータ解析手法の新しい解析手法の提案を目指す。R5 年度は、国立精神・神経医療研究センター(NCNP)で取得されたてんかん手術脳検体のデータを用いて、てん

かん患者の脳内で発現が上昇するシグナル経路の探索手法の開発と同定を試みた。

方法

NCNP バイオバンクから提供を受けた限局性皮質異形成(FCD)患者の手術脳検体 8 症例(type II 6 例、type III 2 例)から取得した単一細胞核 RNA-seq(snRNA-seq)データを用いた。海馬硬化症患者から取得した側頭葉部で正常な脳構造を示した症例をコントロールとして用いた。取得した snRNA-seq データは、SCPA(Bibby 2022)を用いて各シグナル経路の上昇・減少を評価した。

結果

FCD 患者由来のデータとコントロールを比較することによって、有意に上昇・減少したシグナル経路を同定することに成功した。興味深いことに、FCD 患者群では mTOR シグナル経路の上昇を同定することができた。加えて、本研究で用いた手法では、50 を超えるシグナル経路を網羅的に探索することができ、これまでに関連が報告されていないシグナル経路の増減も見出すことができた。

考察・結論

FCD は mTOR シグナルの上昇を特徴とする疾患である。mTOR シグナルは、本来タンパクのリン酸化などの情報をもとにシグナルの増減を評価することが多い。本研究手法では、タンパク情報ではなく、転写産物の情報を利用して mTOR 活性を評価することができた点が極めて重要な成果である。さらに、これまでに関連が報告されていなかったシグナル経路についても網羅的

に同定することができたため、今後のてんかん治療標的候補の探索につながることを期待される。

参考文献(業績)

1. Dewa K, Arimura N, Kakegawa W, Itoh M, Adachi T, **Miyashita S**, Inoue YU, Hizawa K, Hori K, Honjaya N, Yagishita H, Taya S, Miyazaki T, Usui C, Tatsumoto S, Tsuzuki A, Uetake H, Sakai K, Yamakawa K, Sasaki T, Nagai J, Kawaguchi Y, Sone M, Inoue T, Go Y, Ichinohe N, Kaibuchi K, Watanabe M, Koizumi S, Yuzaki M, Hoshino M. Neuronal DSCAM regulates the peri-synaptic localization of GLAST in Bergmann glia for functional synapse formation. Nat Commun. 15(1):458, 2024
2. Iijima K*, Komatsu K*, **Miyashita S***, Suyama K, Murayama K, Hashizume K, Tabe NK, Miyata H, Iwasaki M, Taya S, Hoshino M. Transcriptional features of low-grade neuroepithelial tumors with the BRAF V600E mutation associated with epileptogenicity. Genes Cells. 2024 Mar;29(3):192-206.

データサイエンスによる疾患予測モデリング

植木 優夫

長崎大学

情報データ科学部

緒言

近年のデータサイエンス技術の急速な発展に伴い、医学・生物学分野においても、統計的機械学習などの先端技術の活用が進んでいる。本研究では、脳病態研究領域における神経生理学的データおよび臨床・疫学情報といった高次元ビッグデータに対して、疾患リスク予測を目標とした統計・データサイエンス手法の研究を通じて、各水準での特性を反映した特徴量抽出の技術開発を目指す。一般に、高次元ビッグデータを既存技術により統一的に扱うことは困難である。したがって、データ解析手法自体が重要な研究開発対象となる。本課題では、疾患リスク予測を目標としたテーラーメイドな統計モデリング手法の研究を行い、効果的な特徴量抽出法並びに疾患予測モデルの開発を試みる。

方法

今年度は、データサイエンス技術の先行研究の調査と手法の理論的検討および解析用プログラム開発を行う。その後、開発した技術をシミュレーションにより評価し、同時に、実データを用いて動作検証を行う。実データとしては公開されている医学・生物学分野のオープンデータや ADNI(米国 Alzheimer's Disease Neuroimaging Initiative)の遺伝子データを利用する。

結果

Lasso や Elastic net などの特徴選択が可能なデータ適応的な回帰手法を仮説検定に組

み込むことのできる柳井の一般化決定係数を用いた仮説検定手法[参考文献 1]について、プログラム実装を行い効果的な特徴量抽出を行うための手法研究に取り組んだ。この柳井の一般化決定係数による仮説検定手法を、単変量回帰や飽和回帰モデルなどの旧来のモデルに基づく仮説検定手法、それに加えて SKAT 検定、burden 検定、SKAT-O 検定などのグループ検定に対し、ゲノムデータを想定したシミュレーション実験を通じて比較し性能評価を行った。また、ADNI データに適用し、実データにおける挙動を確かめた。これらの比較実験により、柳井の一般化決定係数による仮説検定手法により、既存手法を上回る検出力が得られることが見出された。得られた知見を日本分類学会 2023 年度シンポジウムにおいて発表した（「データ適応的検定を用いたゲノムワイド関連解析」、2023 年 12 月 2 日、長崎）。さらに、柳井の一般化決定係数を fused lasso に適用し、ゲノムデータへの応用を想定したシミュレーションにより性能検証を行い、結果を論文にまとめた。

考察

シミュレーション研究により、仮説検定にデータに応じた特徴選択を用いることで既存の仮説検定手法を上回る検出力が得られた。従来の仮説検定は、帰無仮説のモデルを固定してデータとのあてはまりの程度を評価するために柔軟さに欠けるが、データに応じた特徴選択を加えることで性能が向上したと考える。

結論

特徴選択を組み入れた柳井の一般化決定係数による仮説検定手法[参考文献 1]が、既存の仮説検定手法を上回る性能を示す可能性が示唆された。以降、今回得られた知見を活かして疾患予測モデリングの研究を進める。

参考文献

1. Ueki M. Testing conditional mean through regression model sequence using Yanai's generalized coefficient of determination. Comput Stat Data Anal 158, 107168, 2021.

予測精度に注目した統計的クラスタリングに基づく個人分類の評価

片平 健太郎

産業技術総合研究所

人間情報インタラクション研究部門

緒言

医学や社会科学等の人間を対象とする諸分野においては、集団の異質性に対処するために対象者の特徴量に基づいてクラスタリングを行い、個人をサブグループに分類するアプローチがしばしば用いられる。そのような個人分類は個別化された治療・介入や予後の予測に役立つと考えられる一方で、連続的な量を離散化することによる情報の損失の問題がある。このような情報の損失が現実的には許容されるものであるのか、結果の予測にどのような影響を及ぼすのかについては、体系的に調査されていない。

方法

本研究では、パーソナリティ等の心理測定から身体活動（日常的な運動の習慣）を予測するという問題において、クラスタリングに基づく個人分類の性質を検討した (Katahira et al., 2023)。具体的には約2万人から得た身体活動とパーソナリティ特性を含む心理的特性に関する調査データを分析することで、サブグループに基づくアプローチの予測性能を、変数を離散化せずに予測に用いる次元的方法と比較しながら検討した。K-means 法等の基本的なクラスタリング手法を用い、心理的特徴の次元データに基づいて個人を分類し、いくつかの異なるクラスタ解を得た。クラスタ

数は Gap 統計量等に基づき決定した。次に、これらのクラスタを用いて身体活動レベルを予測し、その説明力を R^2 で評価した。その結果を、各特徴量を連続変数として重回帰モデルに投入した予測（次元アプローチ）と比較した。また、クラスタに基づく予測の精度を、前年度に開発した統計指標である Cost of cluster-mean based prediction (CCMP, Katahira, 2023) を用いて評価した。

結果

K-means 法を用いたパーソナリティに基づくクラスタリングにおいては、4つのクラスタからなる分類が支持された。その分類に基づく身体活動の予測は、連続変数に基づく回帰モデルの60%程度の予測力しか得られなかった。同等の予測精度を達成するためには、クラスタ数を少なくとも20に増やす必要があった。また、CCMPも20以上のクラスタ数を支持した。

考察・結論

本研究の結果は、クラスタリングに基づく分類を外部結果の予測に使用する場合、従来の統計的基準によって提案されたクラスタ解が最適ではない可能性があることを示唆している。

これまでのサブグループ分類においては、特徴量の分布のみに注目してクラスタ数やクラスタ手法などが選択されていた。本研究は、特徴量の分布のモデルとしては適切ではないとされるモデルが、外的な変数の予測のために有用である場合があることを示唆している。適切な個別化医療のための個人分類は、予測や解釈性の観点も多面的に踏まえて検討していく必要があるだ

ろう。

参考文献

Katahira, K., Takano, K., Oba, T., Kimura., K. (2023). Evaluating the performance of personality-based profiling in predicting physical activity as an external outcome (under review)

<https://osf.io/preprints/psyarxiv/45x8u>

Katahira, K. (2023). Evaluating the predictive performance of subtyping: A criterion for cluster mean-based prediction. *Statistics in Medicine*, 42(7), 1045-1065.

Brain Pathophysiology AI Research:
Integration of Data-driven and Theory-driven Approach
Manabu Honda, M.D., Ph.D.

Department of Information Medicine, National Institute of Neuroscience, NCNP

1. Research Objectives

The aim is to promote the understanding, diagnosis, treatment, and prevention of mental, neurological, and muscular diseases as well as developmental disorders by integrating data science, AI technology, and computational neuroscience. The specific objectives are as follows:

- Establishing a framework and infrastructure for promoting brain pathology research using AI: Develop infrastructure to leverage research resources accumulated at NCNP Hospital, NCNP Biobank, and various research institutes.
- Developing AI-based clinical and basic big data analysis methods for brain pathologies (data-driven research): Analyze high-dimensional big data such as genetic, molecular, neurophysiological, clinical, and epidemiological data using data science techniques to extract features reflecting characteristics at each level and develop cross-level analysis techniques and theories.
- Modeling brain pathologies based on brain computation theory and statistical mathematics (theory-driven research) and developing integrated data-driven and theory-driven approaches: Implement brain computation theory using neural circuit models to propose hypotheses for brain pathology understanding and optimal therapeutic interventions. Develop theories and techniques for new disease concepts and personalized medical technologies by correlating proposed brain computation theories with actual basic and clinical data.

2. Research Outcomes

Organization and infrastructure building

- (1) Established the Brain Pathology AI Center (BAIC) Preparation Office within NCNP, designed the information infrastructure, and secured necessary equipment and personnel.
- (2) Set up a 10G network line for high-speed connection between internal and external information infrastructures of the center and improved security environments for external access.
- (3) Employed a special appointment director, research assistants, and administrative assistants with expertise in information infrastructure construction, conducted hearings on large-scale information assets scattered across NCNP Hospital, the Biobank, and various research institutes, and cataloged the information assets within the center.
- (4) Concluded a data usage agreement with the UK's UK Biobank, establishing an international data collaboration system.

- (5) Supported the signing of a partnership agreement between NCNP and a company (NTT), establishing an industry-academia-government collaboration system.
- (6) Conducted bimonthly web seminars and an annual hands-on seminar and symposium (CPSY Course Tokyo) to not only educate but also serve as a hub for connecting researchers within and outside NCNP.

Data-Driven Research

- (7) Conducted large-scale data analysis on methylation and whole RNA sequencing of low-grade epilepsy-associated tumor specimens, obtaining results suggesting new subgroups and molecular mechanisms related to the disease.
- (8) Developed an application for muscle pathology differential diagnosis using muscle pathology diagnostic algorithms and created a dataset classified by Mercuri scores for the degree of muscle involvement in muscle diseases.
- (9) Analyzed Parkinson's disease model mice using a highly sensitive histone modification analysis system and flow cytometry-based nuclear fractionation technique, identifying transcription factor changes associated with pathological changes.
- (10) Analyzed large cohort data of brain structural images in schizophrenia, bipolar disorder, depression, and autism spectrum disorder, suggesting the classification of four brain structure biotypes based on subcortical structural feature patterns and cognitive and social function performance.
- (11) Conducted literature reviews, hearings, and information exchanges on the applicability of AI technology in psychotherapy across various disciplines.
- (12) Conducted clustering analysis of mental disorder registry data cleaned last year and examined the association with outcome indicators such as QOL.

Theory-Driven Research and Integrated Approaches

- (13) Examined the applicability of multimodal variational autoencoders, a type of deep neural network, as cognitive function modeling. Additionally, developed a biosignal generation simulator using ECoG data from macaques as a prototype digital brain model, successfully simulating real-time brain state estimation, ECoG signal generation, and virtual drug administration.
- (14) Conducted primary analysis of transcriptome data from surgical brain specimens of epilepsy patients and registered the data in a database, developing a pipeline for estimating the activation state of the mTOR signaling pathway.
- (15) Developed data science methods applicable to image and genome data from a theoretical perspective and conducted tutorial lectures on related technologies.
- (16) Applied the proposed Cost of Cluster-Mean Based Prediction (CCMP) metric for subtype

classification prediction accuracy to the problem of predicting daily exercise habits based on personality. The results suggested that conventional statistical subgroup classification might not be optimal from a prediction perspective.