

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リ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）
- ・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) 脳病態 AI センター(BAIC)設立準備室（室長：永田充生 NCNP 企画戦略局長、室長代理：本田）を発足させ、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) NCNP 内に設立準備室の区画を設置し、NCNP 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) 知識情報基盤構築の専門家を新規に

BAIC 設立準備室特任室長に採用するため、複数の候補者とヒアリングを行った。2023年度4月より、慶應義塾大学医学部 立森教授を特任室長（兼任）として採用した（本田）。

(4) 世界最大規模の UK Biobank とデータ使用契約を締結し、データを活用できる体制を整えた（本田）。

(5) 脳病態数理・データ科学セミナーを隔月（年度内計6回）開催し、データベースやAI活用のための教育を行った（山下）。

#### ② データ駆動型研究

(6) 低悪性度てんかん原生腫瘍などの検体について、total RNA シークエンシングを行い、検体の特徴を表出するクラスタリングに成功した（星野）。

(7) 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を進めるとともに、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（西野）。

(8) 高感度ヒストン修飾解析系、および、Flow cytometry を用いた細胞核の分取技術の立ち上げをおこなった（間野）。

(9) 眼球運動と認知機能データを組み合わせ、健常者と統合失調症を高い精度で判別できるモデルを作成した（三浦）。

(10) 心理療法における AI 技術応用可能性に関する文献調査とヒアリングを実施した（伊藤）。

(11) 精神疾患レジストリ解析の研究計画について審査と承認を得て、データのクリーニング・解析に着手した（大庭）。

#### ③ 理論駆動型研究・統合的アプローチ

(12) 多モダリティ変分自己符号化器を用いた特徴表現学習手法(データ駆動的手法)の開発に加え、強化学習・予測情報処理モデルを用いた、意思決定行動・知覚認知プロセスのモデル化(理論駆動的手法)を行った。さらに、それらを統合する試みとして予測情報処理モデルにおけるパラメータ間相互作用について安静時 fMRI データを用いた検証を行った(山下)。

(13) Variational Autoencoder を適用することで細胞遺伝子発現の高精度な再現に成功し、ヒトてんかん患者のシングルセル RNA シーケンシングデータの解析を開始した(宮下)。

(14) 非線形ランダム効果とベータ負の二項分布を用いたデータサイエンス手法を用いた時空間的モデリングによって、既存モデルの課題を緩和できることを示した(植木)。

(15) 所与のサブタイプ分類の予測の精度を測る指標である Cost of cluster-mean based prediction (CCMP) を提案し、シミュレーションとオープンデータへの適用を通してその有用性を明らかにした(片平)。

#### 4. 研究成果刊行一覧

##### 【論文】

計 13 件(うち査読付論文 9 件)

1. Sano, T., Kawazoe, T., Shioya, A., Mori-Yoshimura, M., Oya, Y., Maruo, K., Nishino, I., Hoshino, M., Murayama, S., & Saito, Y. (2022). Unique Lewy pathology in myotonic dystrophy type 1. *Neuropathology: official journal of the Japanese Society of*

*Neuropathology*, 42(2), 104–116.

<https://doi.org/10.1111/neup.12790> (査読有)

2. Sato, K., Niimi, Y., Mano, T., Iwata, A., & Iwatsubo, T. (2022). Automated Evaluation of Conventional Clock-Drawing Test Using Deep Neural Network: Potential as a Mass Screening Tool to Detect Individuals With Cognitive Decline. *Frontiers in neurology*, 13, 896403.  
<https://doi.org/10.3389/fneur.2022.896403> (査読有)
3. Uchida, Y., Hikida, T., & Yamashita, Y. (2022). Computational Mechanisms of Osmoregulation: A Reinforcement Learning Model for Sodium Appetite. *Frontiers in neuroscience*, 16, 857009.  
<https://doi.org/10.3389/fnins.2022.857009> (査読有)
4. Ito, S., Miura, K., Miyayama, M., Matsumoto, J., Fukunaga, M., Ishimaru, K., Fujimoto, M., Yasuda, Y., Watanabe, Y., & Hashimoto, R. (2022). Association between globus pallidus volume and positive symptoms in schizophrenia. *Psychiatry and clinical neurosciences*, 76(11), 602–603.  
<https://doi.org/10.1111/pcn.13465> (査読有)
5. Nagamitsu, S., Kanie, A., Sakashita, K., Sakuta, R., Okada, A., Matsuura, K., Ito, M., Katayanagi, A., Katayama, T., Otani, R., Kitajima, T., Matsubara, N., Inoue, T., Tanaka, C., Fujii, C., Shigeyasu, Y., Ishii,

R., Sakai, S., Matsuoka, M., Kakuma, T., Horikoshi, M. (2022). Adolescent Health Promotion Interventions Using Well-Care Visits and a Smartphone Cognitive Behavioral Therapy App: Randomized Controlled Trial. *JMIR mHealth and uHealth*, 10(5), e34154.  
<https://doi.org/10.2196/34154> (査読有)

#### 【図書】

計 4 件

1. 間野達雄, 岩田淳 (2022 年 7 月), 加齢脳とエピジェネティクス異常, 老年精神医学雑誌, ワールドプランニング, pp. 715-720.
2. 間野達雄, 岩田淳 (2022 年 7 月), 【Endosome、Lysosome、Exosome-小胞とは?】 Exosome と疾患: アルツハイマー病, *Clinical Neuroscience*, 中外医学社, Exosome と疾患, pp.862-864.
3. Tatsuo Mano, Atsushi Iwata (2022 年 8 月), [Section VII: Autophagy and other disorders causing dementia] Autophagy and Huntington's disease, *Autophagy dysfunction in Alzheimer's disease and dementia*, Academic press, pp.245-262.
4. 山下祐一 (2022 年 11 月), 計算論的精神医学の視点からみた統合失調症, 統合失調症という問い, 日本評論社, pp.163-186.

#### 【学会発表】

計 15 件

1. Hoshino M, Research on brain

development and disorders by utilizing spontaneous mutant animals, NEURO2022, July 2022, Okinawa, Japan (招待講演).

2. Yamashita Y, Psychiatric and neurodevelopmental disorders as failures in hierarchical predictive process: neurorobotics approach, International Symposium on Artificial Intelligence and Brain Science (AIBS2022), July 2022, Okinawa, Japan (招待講演).
3. 山下祐一, 計算論的精神医学: 数理・データ科学を用いて精神疾患を理解する, 日本神経科学会・教育講演, 2022 年 6-7 月, 沖縄 (招待講演).
4. 三浦健一郎, 臨床現場で使えるタブレットを用いた統合失調症の補助診断法の開発, 第 41 回日本精神科診断学会, 2022 年 9 月, オンライン.
5. 宮下聡, 発生期小脳顆粒細胞における一過的な神経前駆細胞の増殖メカニズム, NEURO2022, 2022 年 7 月, 沖縄.

## 分担研究報告書

(課題名) マルチオミックスデータを用いた脳病態解明

(所属) 国立精神・神経医療研究センター  
病態生化学研究部

(氏名) 星野 幹雄

### 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

今年度は、NCNP バイオバンクのてんかん手術脳検体を用いて、total RNA-seq を行う。また、一部の検体については、単一核(sn)RNA-seq を行う。得られたデータをバイオインフォマティクス解析し、てんかん病態を理解するための基盤となる情報を集める。

### 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た。また LEAT の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。遺伝子発現に基づくクラスタリングにより、同じ遺伝子変異を持つ LEAT の中でもサブクラスターの存在が示唆され、それぞれが異なる病態を反映している可能性が示唆された。さらに、3 検体については、snRNAseq のデータ取得を完了したので、これから解析を行う。

(2)FCD (限局性皮質異形成症) についてほぼ全ての検体について、mTOR パネルによる遺伝子変異の検出を完了している。また、10 検体の total RNAseq データを取得したが、この検体すうは今後も増やす予定である。それに先行して、mTOR 遺伝子そのものに遺伝子変異を持つ 6 検体については、先行して snRNAseq を行った。そのデータをコントロール (5 検体) のデータと比較し、各細胞種ごとで異なる発現をする遺伝子群(DEG)を抽出した。詳細な解析はこれからである。また、一部の検体については、mTOR 系の活性をウエスタンブロットで調べた。

### 考察

まだ全てのデータを取得しているわけではないので、次年度もデータ取得に努める。もちろん、次年度からはデータ解析に本格的に取り組む。病態仮説が描出されたら、モデルマウスを用いてその仮説を検証することになる。

### 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. Fujiyama T, Takenaka H, Asano F, Miyanishi K, Hotta-Hirashima N, Ishikawa Y, Kanno S, Seoane-Collazo P, Miwa H, Hoshino M, Yanagisawa M, Funato H. Mice Lacking Cerebellar Cortex and Related Structures Show a Decrease in Slow-Wave Activity With Normal Non-REM Sleep Amount and Sleep Homeostasis. *Front Behav Neurosci.* 2022, 16:910461.
2. Hashimoto Y, Kuniishi H, Sakai K, Fukushima Y, Du X, Yamashiro K, Hori K, Imamura M, Hoshino M, Yamada M, Araki T, Sakagami H, Takeda S, Itaka K, Ichinohe N, Muntoni F, Sekiguchi M, Aoki Y. Brain Dp140 alters glutamatergic transmission and social behaviour in the mdx52 mouse model of Duchenne muscular dystrophy. *Prog Neurobiol.* 2022, 216:102288.
3. Sano T, Kawazoe T, Shioya A, Mori-Yoshimura M, Oya Y, Maruo K, Nishino I, Hoshino M, Murayama S, Saito Y. Unique Lewy pathology in myotonic dystrophy type 1. *Neuropathology.* 2022, 42(2):104-116
4. Inoue, Y.U., Miwa, H., Hori, K., Kaneko, R., Morimoto Y., Koike, E.,

Asami, J., Kamijo, S., Yamada, M, Hoshino, M. and Inoue, T.: Targeting Neurons with Functional Oxytocin Receptors: A Novel Set of Simple Knock-In Mouse Lines for Oxytocin Receptor Visualization and Manipulation. *eNeuro.* 9(1): 0423-21, Feb, 2022

## 分担研究報告書

(課題名) AI を用いた筋疾患診断支援システム開発

(所属) 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

(氏名) 西野 一三

## 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、カルパインパチー、先天性ミオパチーなどに高確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。本研究課題に小居ては、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。

## 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

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 OPDM との鑑別には中殿筋、大殿筋、大内転筋が有用であることが分かった。

## 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 の解析結果については論文投稿中である。他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、random forest supervised machine learning を用いたモデル作成を目指す。

## 参考文献

1. 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
2. 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)
3. 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

4. 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)



## 分担研究報告書

(課題名) 神経変性過程における包括的エピゲノムアプローチ

(所属) 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

(氏名) 間野 達雄

### 緒言

孤発性神経変性疾患は遺伝性背景と後天的要因が複雑に関連しあって病態を形成している。このような病態理解において死後脳を出発点とした網羅的解析が重要であり、本研究では神経変性疾患のエピゲノム変化から病態を理解する。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ているが、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。従って、本研究では特に、神経変性過程における神経細胞、グリア細胞（アストロサイト、オリゴデンドロサイト、ミクログリア）のエピゲノム解析を行い、ゲノムワイドに起きている現象を明らかにする。

### 方法

脳組織における主要な細胞群について、それぞれの核および細胞質マーカーの検索を行う。その上で、フローサイトメトリーに利用可能な抗体について蛍光標識を行い、脳組織から分離した核の染色を行い、それぞれの細胞種を適切に染色できていることを確認する。これらの抗体を用いて、各種細胞核をフローサイトメトリーにより分離し、CUT&Tag法によるヒストン修飾解析を行った。

### 結果

神経細胞およびアストロサイトについては、細胞質マーカー(NeuN および GFAP)を用いて分離することが可能であった。オリゴデンドロサイトおよびミクログリアについては、核マーカー(Olig2 および PU.1)を用いることにより分離可能であった。ヒストン修飾解析については、HH3K4me3 および H3K27ac について行い、5000 個の核からゲノムワイドのヒストンプロファイルデータを得ることができた。Technical replicate ではいずれも相関係数 0.9 以上と、少ない細胞数から良好な再現性のあるデータを得ることができた。Gene ontology 解析ではそれぞれの細胞種の機能を反映した遺伝子領域のピーク検出ができていることを確認した。

### 考察

細胞種特異的な解析を行うにあたって、(1)少数の核から効率の良い方法でライブラリ調整をする、(2)細胞種に応じたマーカーを選択する、という2点が課題であり、本検討から解決することができた。本検討で確立した手法を疾患モデル、および患者脳に適用することで細胞種ごとの病態をエピゲノム解析からアプローチすることが可能となる。

### 結論

細胞種特異的ヒストン修飾解析の手法を確立した。実験間の再現性も高く、疾患脳およびモデル系に適用する。

#### 参考文献

1. 間野達雄、岩田淳、ワールドプランニング、老年精神医学雑誌 加齢脳とエピジェネティクス異常、2022年7月、6ページ

## 分担研究報告書

(課題名) 脳計算理論を用いた精神疾患病態メカニズムの解明

(所属) 国立精神・神経医療研究センター  
疾病研究第七部

(氏名) 山下祐一

### 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。

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、ベイズ推論・強化学習などの脳の計算理論に基づいて、各水準間のデータ生成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。

得られた特徴量に基づいて、既存の精神障害カテゴリーにとらわれない疾患横断的・次元的アプローチに基づいて、疾患亜型を探索的に検討し、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試みる。

### 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、多モダリティ変分自己符号化器を用いた特徴表現学習手法の開発と、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。理論駆動的な手法として、強化学習モデル、予測情報処理モデルを用いた、意思決定行動・知覚認知プロセスのモデル化を行った。さらに、これらのアプローチを統合する試みとして、予測情報処理モデルによって予測されるパラメータ間の相互作用を、実際の精神疾患における安静時 fMRI データを用いて検証した。

今後は、これらの成果に基づいて、データ

駆動的手法、モデル駆動的手法によって得られた特徴量に基づいて、既存の精神障害カテゴリーにとらわれない疾患横断的・次元的方法に基づいて、疾患亜型を探索的に検討し、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試みる。

#### 考察

本研究により、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、患者個人の人々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. Idei H, Ohata W, Yamashita Y, Ogata T, Tani J. Emergence of sensory attenuation based upon the free-energy principle. *Sci Rep* 12, 14542, <https://doi.org/10.1038/s41598-022-18207-7>, 2022Aug, 査読有
2. Uchida Y, Hikida T and Yamashita Y, Computational Mechanisms of Osmoregulation: A Reinforcement Learning Model for Sodium Appetite. *Front. Neurosci.* 16:857009, 2022May. doi: 10.3389/fnins.2022.857009, 査読有
3. Takahashi Y, Murata S, Ueki M, Tomita H, & Yamashita Y (2022) Interaction between functional connectivity and neural excitability in autism: A novel framework for computational modeling and biological validation. *PsyArXiv*. 28 Apr. 2022, <https://doi.org/10.31234/osf.io/ut39r>

, 査読無

4. 山下 祐一, 計算論的精神医学：精神医学における計算理論と知能モデルの役割と展望, *日本神経回路学会誌* 29 巻, p. 37-40, 2022Jun. <https://doi.org/10.3902/jnns.29.37>, 査読無
5. Noda K, Soda T, & Yamashita Y, Emergence of Number Sense in Deep Multi-modal Neural Networks. *PsyArXiv*. 2022Aug. doi:10.31234/osf.io/4bfam, 査読無.

## 分担研究報告書

(課題名) 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

(所属) 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

(氏名) 三浦健一郎

### 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。

R4年度は、眼球運動、認知機能、脳神経画像をはじめとする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

備を進めつつ中間表現型と症状などとの関連についての基礎的知見を得た[1]。また、以下に記すように、眼球運動と認知機能の組合せによる統合失調症と健常者の分類を試みた[2]。

注視、追跡運動、自由視などの種々の課題を用いて得られる眼球運動の各種特徴量及びウェクスラー式知能検査(WAIS-III)及び記憶検査(WMS-R)によって定量化される各検査のスコアからなる認知機能データを組み合わせて特徴量ベクトルを構成し、L1正則化及びL2正則化に基づくロジスティック回帰分類器を構成し、その性能を評価した。さらに眼球運動及び認知機能の特徴量からそれぞれ一つずつの特徴量のペアを用いて分類器を構成し、その正診率を評価した。

尚、本研究は「人を対象とする生命科学・医学系研究に関する倫理指針」を遵守した研究計画書を作成し、研究施設での倫理委員会の承認を得て参加者からは書面による同意を得て実施した。

### 結果

眼球運動 16項目と認知機能 25項目の計 41項目の特徴量を用いた L2 正則化に基づくロジスティック回帰分類器は、トレーニングセットとテストセットに対してそれぞれ 92.0%と 75.8%の正診率を示した。また、41項目を用いた L1 正則化に基づくロジスティック回帰分類器では、トレーニングセットとテストセットでそれぞれ 93.2%と 81.6%の正診率を示した。L1 正則化に基づく分類器には、眼球運動の自由視課題におけるスキャンパスの長さ及び WAIS-III の符号、算数、記号探し、理解の 5つ

の検査のスコアの回帰係数が非ゼロであった。眼球運動及び認知機能の特徴量からそれぞれ一つずつの特徴量の組合せを用いた解析では、WAIS-III の眼球運動の特徴量と符号あるいは記号探索の組合せにおいて高く、ロバストな正診率（トレーニングセットとテストセットの両方に対し>80%）を示した。

### 考察

眼球運動は視覚機能を維持する脳の生理的な感覚運動機能である。一方、WAIS-III や WMS-R は、脳の高次統合機能を評価するものである。これらは脳病態の異なる側面を反映したものと考えられる。本研究の結果は、中間表現型情報を組み合わせた多次元情報を持ちることにより高い正診率を持つ分類器を構成できることを示唆する。統合失調症を客観的に検査する試みがいくつか行われている中で、眼球運動や認知機能の検査は、タブレットなどの小型の情報機器を用いて非侵襲的に行うことができるという利点がある。また、本研究の結果は、短時間で行うことができるわずか数個の検査で 80%以上の高い正診率が得られることを示している。さらに、特別な訓練を受けたスタッフ無しに遠隔で実施できるように構成できる可能性もあり、近年発展してきているデジタルヘルスとしての実用性が高いと考えられる。

精神科医は、ICD-11 や DSM-IV/5 などの診断基準を用いて統合失調症を診断し、PANSS などを用いて症状を評価する。しかし、症状が曖昧な場合には、統合失調症の有無の鑑別が困難なこともある。本知見を応用した医療機器は、そのような場合

に、より正確な臨床診断に寄与する可能性がある。また、治療にあたっては、客観的情報の提示により、医師と患者の間の shared decision making の促進にも寄与すると考えられる。

### 結論

眼球運動と認知機能の中間表現型特徴量を組み合わせることで、健常者と統合失調症を高い正診率で判別することができる。これらの中間表現型は、統合失調症の臨床診断を補助する際に有効である。本結果は、より正確な診断及び、医師間の診断の一貫性の向上、患者と医師の間の shared decision making を助ける客観的な補助診断法の開発に寄与する。

### 参考文献

- [1] Ito S, Miura K, Miyayama M, Matsumoto J, Fukunaga M, Ishimaru K, Fujimoto M, Yasuda Y, Watanabe Y, Hashimoto R. Association between globus pallidus volume and positive symptoms in schizophrenia. *Psychiatry Clin Neurosci.*, 76, 602-603, 2022.
- [2] Okazaki K, Miura K, Matsumoto J, Hasegawa N, Fujimoto M, Yamamori H, Yasuda Y, Makinodan M, Hashimoto R. Discrimination in the clinical diagnosis between patients with schizophrenia and healthy controls using eye movement and cognitive functions. *Psychiatry Clin Neurosci.* 2023, in press.

## 分担研究報告書

(課題名) 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

(所属) 国立精神・神経医療研究センター  
認知行動療法センター

(氏名) 伊藤正哉

### 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。本年度は、様々な側面で発展を続ける心理療法分野における人工知能技術の適用について整理した。

### 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** Pubmed において "psychotherapy" AND "artificial intelligence" と検索すると 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。心理療法の研究においても、心理療法中の対話データに対して自然言語処理を適用し、セッションの評価や患者の精神状態を推論させる研究が報告されている (Graham et al., 2019)。モダリティーとしては、自然言語だけでなく、音声や顔認識技術を用いて、患者の感情を解析する試みもなされている (Poria, Cambria, Bajpai, & Hussain, 2017)。こうした技術は、患者が自身の感情を適切に表現できない場合や、非言語的な感情表現を解析するのに特に有用である (Calvo, D'Mello, Gratch, & Kappas, 2015)。また、統合失調症や精神科疾患の診断補助として自然言語マーカー (Corona Hernández et al., 2023) や音声マーカー (Cummins et al., 2015) を用いる試みや、自殺リスクの予測に応用する研究 (Lejeune et al., 2022) も生まれつつある。

**心理療法後の予後予測** プロセス、遺伝、デモグラフィック、臨床データといったマルチモーダルな 69 変数を用いて、インターネット認知行動療法後の寛解を予測する試みが報告されている (Wallert et al., 2022)。

**AI を活用した介入** AI による完全自動型の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。その先駆けとなる Woebot 社のアプリについて報告した Fitzpatrick et al (2017) は 1294 回引用され

ており、同アプリは物質使用障害などにも拡大されている。興味深い知見として、36070 ユーザーに調査した結果、人間のセラピストに感じるのと同等の“絆”の感覚を、デジタル会話エージェントにも抱かれることが報告されている (Darcy et al., 2021)。

会話エージェントによる介入の 32 研究についてのレビューでは、17 研究がウェルビーイングの促進に、21 研究が精神症状の治療として報告されていた (Jabir et al., 2023)。ただし、AI によるメンタルヘルスアプリとして有効性が検証されているのは、Fitzpatrick et al (2017) のほか、2 研究に限られている (Milne-Ives et al., 2022)。

**心理療法の補助** チャットボットや会話エージェントについての文献レビューでは、これらがセラピストと患者との間のコミュニケーションを促進したり、心理教育やアドヒアランスの補助となりうることを示している (Vaidyam, Wisniewski, Halamka, Kashavan, & Torous, 2019)。

**アクセス性とコスト効率性** 必ずしも AI が搭載されている必要はないが、デジタルプラットフォームは、リモート地域に住む患者や、コストや時間の制約から対面の治療にアクセスできない患者にとって、心理療法を受けるための手段を提供する (Kretzschmar, Tyroll, Pavarini, Manzini, & Singh, 2019)。これにより、心理療法へのアクセスが向上し、全体的なメンタルヘルスケアの質が向上する可能性がある (Torous, Myrick, Rauseo-Ricupero, & Firth, 2020)。また、チャットボットによって、心理療法を擬似的に体験することにより、心理療法活用の心理的抵抗を減らす

可能性も考えられる。

**AI の活用に関するリスク** 心理療法プロセスを支援するために機械学習を用いる危険性について指摘する論文もある。例えば、データの二次利用についての法規制が整備されていない現状にあって、個人情報が無断で第三者に売却される危険性が指摘されている (Lewanowicz et al., 2022)。また、AI による誤った判断や有害事象が適切に報告されるためのシステム整備や、そのような事態が発生した場合の責任の所在も明確化されておらず (Luxton, 2014)、多く課題が残されている。

## 結論

心理療法への AI 適用が進んでいるが、その多くはチャットボットによる介入が多い。自然言語以外では、音声や表情等のモダリティーを扱う研究も散見される。しかし、心理療法セッションにおける様々なモダリティーのデータを扱うような研究はまだかなり限定的である。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、さらなる研究が必要である。

## 文献

- Calvo, R. A., D'Mello, S., Gratch, J., & Kappas, A. (2015). *The Oxford handbook of affective computing*. Oxford University Press.
- 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., Koops, 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>
- Cummins, N., Scherer, S., Krajewski, J., Schnieder, S., Epps, J., & Quatieri, T. F. (2015). A review of depression and suicide risk assessment using speech analysis. *Speech communication*, 71, 10-49.
- Darcy, A., Daniels, J., Salinger, D., Wicks, P., & Robinson, A. (2021). Evidence of Human-Level Bonds Established With a Digital Conversational Agent: Cross-sectional, Retrospective Observational Study. *JMIR formative research*, 5(5), e27868. <https://doi.org/10.2196/27868>
- Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). "Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial". *JMIR Mental Health*, 4(2), e19.
- Graham S, Depp C, Lee EE, Nebeker C, Tu X, Kim HC, Jeste DV. Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. *Curr Psychiatry Rep*. 2019 Nov 7;21(11):116. doi: 10.1007/s11920-019-1094-0. PMID: 31701320; PMCID: PMC7274446.
- Jabir, A. I., Martinengo, L., Lin, X., Torous, J., Subramaniam, M., & Tudor Car, L. (2023). Evaluating Conversational Agents for Mental Health: Scoping Review of Outcomes and Outcome Measurement Instruments. *Journal of medical Internet research*, 25, e44548. <https://doi.org/10.2196/44548>
- Koutsouleris, N., Hauser, T. U., Skvortsova, V., & De Choudhury, M. (2022). From promise to practice: towards the realisation of AI-informed mental health care. *The Lancet. Digital health*, 4(11), e829–e840. [https://doi.org/10.1016/S2589-7500\(22\)00153-4](https://doi.org/10.1016/S2589-7500(22)00153-4)
- Lejeune, A., Le Glaz, A., Perron, P. A., Sebti, J., Baca-Garcia, E., Walter, M., ... & Berrouiguet, S. (2022). Artificial intelligence and suicide prevention: a systematic review. *European psychiatry*, 65(1), e19.
- Lewanowicz, A., Wiśniewski, M., & Oronowicz-Jaśkowiak, W. (2022). The use of machine learning to support the therapeutic process - strengths and weaknesses. *Postepy psychiatrii neurologii*, 31(4), 167–173. <https://doi.org/10.5114/ppn.2022.125050>
- Luxton, D. D. (2014). Artificial intelligence

- in psychological practice: Current and future applications and implications. *Professional Psychology: Research and Practice*, 45(5), 332.
- Milne-Ives, M., Selby, E., Inkster, B., Lam, C., & Meinert, E. (2022). Artificial intelligence and machine learning in mobile apps for mental health: A scoping review. *PLOS digital health*, 1(8), e0000079. <https://doi.org/10.1371/journal.pdig.0000079>
- Kretzschmar, K., Tyroll, H., Pavarini, G., Manzini, A., & Singh, I. (2019). "Can your phone be your therapist? Young people's ethical perspectives on the use of fully automated conversational agents (chatbots) in mental health support". *Biomedical Informatics Insights*, 11, 1178222619829083.
- Poria, S., Cambria, E., Bajpai, R., & Hussain, A. (2017). "A review of affective computing: From unimodal analysis to multimodal fusion". *Information Fusion*, 37, 98-125.
- Torous, J., Myrick, K. J., Rauseo-Ricupero, N., & Firth, J. (2020). "Digital mental health and COVID-19: Using technology today to accelerate the curve on access and quality tomorrow". *JMIR Mental Health*, 7(3), e18848.
- Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., & Torous, J. B. (2019). "Chatbots and conversational agents in mental health: A review of the psychiatric landscape". *The Canadian Journal of Psychiatry*, 64(7), 456-464.
- Wallert, J., Boberg, J., Kaldo, V., Mataix-Cols, D., Flygare, O., Crowley, J. J., Halvorsen, M., Ben Abdesslem, F., Boman, M., Andersson, E., Hentati Isacsson, N., Ivanova, E., & Rück, C. (2022). Predicting remission after internet-delivered psychotherapy in patients with depression using machine learning and multi-modal data. *Translational psychiatry*, 12(1), 357. <https://doi.org/10.1038/s41398-022-02133-3>

## 分担研究報告書

(課題名) 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

(所属) 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

(氏名) 大庭真梨

## 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

本年は医学研究における予測モデル構築の考え方について整理した。また、予測モデルおよび時系列データ解析に関する勉強会、文献的検討を行った。並行して、精神疾患レジストリデータの利活用申請、取得、取得後のクリーニング、テキスト情報のコード化、および基本的な集計を行った。

## 結果

医学研究における予測モデルの構築に関するガイドライン Transparent Reporting of a multivariable prediction model of Individual Prognosis Or Diagnosis (TRIPOD) statement、TRIPOD Explanation and Elaboration (E&E)、Protocol for development of a reporting guideline (TRIPOD-AI)をレビューし要点をまとめた。

・ TRIPOD statement および TRIPOD E&E<sup>1</sup>

医学研究における予測モデルは時間の概念に基づいて診断予測モデルと予後予測モデルに大別され、前者は因子が測定された時点に有病である可能性、後者は将来までのある期間でアウトカムが発生する可能性を推定する。いずれも高い予測能を目指すもので、予測能評価には較正と識別と過剰適合の定量化が必要である。予測能評価は内部検証として、ブートストラップ法やクロスバリデーションにより行う。さらに外部検証も強く推奨される。外部検証データには、時間差や地域差のあるデータを使用されることが多く、予測と観測値に系統的なズレがあればモデルは検証データに基づき再較正される。

予測モデルの発表には以下の要素が必要である：データの詳細 (対象集団、サンプル数とイベント数、予測因子とアウトカムおよび測定方法、時期と追跡期間、欠測の有無と対処)、モデルのタイプと選択理由と構築過程、検証方法、検証に使用したデータ詳細等、および資金源である。これらは再現性や一般化可能性、バイアスリスク、

使い分けを理解するのに役立つ。線形モデルにおける変数選択に関しては、単変量解析による予測因子の選択は推奨されない。情報量基準や変数増加法や減少法がE&Eには紹介されていたが、変数の無数の組み合わせを評価する代わりにLasso法の採用も有用である。

・TRIPOD-AI, Prediction model Risk Of as ASsessment Tool(PROBAST)-AI<sup>2,3</sup>  
作成中のTRIPOD-AI,PROBAST-AIについて、作成プロトコルのみが公開されている。大規模データ、多くの因子を用いた機械学習やAIの予測モデルは実践研究の増加が著しい。一方、医学研究では方法由来の欠点、非透明性、再現性の低さの懸念もあり、公正な評価と既存モデルとの中立的な比較が推奨されている。教師あり学習全般を対象にTRIPODと同様の手順で検討が進んでいること、全医学研究およびがん研究を関連論文をレビューし、バイアスリスクを評価し、報告内容について検討されていることが報告されている。<sup>3</sup>  
機械学習を用いた予測モデルに関するレビュー研究<sup>4</sup>によると、出版されレビューされた152件の論文うち3分の1は欠測について記載がなく、3分の2はその処理方法の記載はあったが、そのメカニズムや理由への言及は稀であった。3分の2に含まれる処理方法は多くが削除による完全データの解析であり、そのほか多重補完、surrogate splitsなどの方法が挙げられていたと報告されており、依然として完全データを用いた解析が多く、欠測考慮による予測や分類の影響は不明であった。

精神疾患レジストリは同意取得数は2022年末の時点で1424件、600人に近い基本情報が収集されており、クリーニングや再コーディング、欠測状況の確認などを行った。

## 考察

リスク因子とアウトカムの関連解析では、データの欠測メカニズムを、ランダムに発生したとみなせる欠測と、欠測理由が欠測値自体にある欠測、に分けてアプローチすることが多い。前者に対して多重補完

(Rubin, 1976)や混合モデルの当てはめが多用される。あるいはいずれにも様々な感度解析手法が適用される。特に多重補完は予測モデルでも実用例がみられ、今後検討する価値があると考えられた。

実データにおいて欠測理由を知ること及び観測データの特徴を知ることが重要であるものの、欠測と予測能の関係はデータに依存するため、現在利活用中の精神疾患レジストリデータの状況に合わせた方法の探索が必要であると考えられた。

## 結論

予測モデル構築に関するガイドンスとその発展をまとめ、解析方法と欠測への処理方法について情報収集を行った。欠測対応の重要性は指摘されているものの既存の予測研究ではその対処法は定まっておらず、データに合わせた方法の検討の必要性が感じられた。

## 文献

1. Moone KGM, et al. Transparent Reporting of a multivariable prediction

model for Individual Prognosis Or  
Diagnosis (TRIPOD): Explanation  
and Elaboration. *J Annals of Internal  
Medicine*. 2015; 162; w1-73.

2. Collins GS and Moons KGM.  
Reporting of artificial intelligence  
prediction models. *Lancet*. 2019; 393:  
1577-9.
3. Collins GS, et al. Protocol for  
development of a reporting guideline  
(TRIPOD-AI) and risk of bias tool  
(PROBAST-AI) for diagnostic and  
prognostic prediction model studies  
based on artificial intelligence. *BMJ  
Open* 2021; 11: e0480008.
4. Nijman SWJ, et al. Missing data is  
poorly handled and reported in  
prediction model studies using  
machine learning: a literature review.  
*JCE*. 2022;142:218-29.

## 分担研究報告書

(課題名) 機械学習を用いたオミクスビッグデータの解析

(所属) 新潟大学脳研究所システム脳病態学分野

(氏名) 宮下 聡

### 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。

### 方法

シミュレーションデータや公共の成体マウスの single cell RNAseq(scRNAseq)データ、公共の自閉症患者の scRNAseq データを取得した。それらのデータの遺伝子発現量を Seurat によって用いて標準化したもの、もしくは、MAGIC によって処理をし、発現量を normalize したものをを用いた。Normalize したデータを training 用と validation 用に分割し、Variational Autoencoder(VAE)モデルによって学習を行わせた。

### 結果

本年度は研究計画に従い、VAE を用いたシングルセル RNA シーケンシングデータの解析を行なった。はじめに、マウス小脳顆粒細胞を用いて、我々の VAE モデルが、既知の細胞種の分類・特徴量抽出に適しているかを、検証した。学習後の VAE の潜在空間を UMAP によって、次元圧縮し、既知のマーカーを用いてそれぞれのクラスターで発現している分子を調べた。その結果、我々の VAE モデルは、神経細胞とグリア細胞といった異なる細胞種を適切に分類できただけではなく、神経細胞の中でも興奮性、抑制性、神経前駆細胞を異なるクラスターとして分類することができた。

次に、自閉症患者の scRNAseq データを用いて解析を行ったところ、自閉症患者に固有のアストロサイト群があることを見出した。

### 考察・結論

以上の結果は、我々の VAE モデルが脳サンプルから得られた scRNAseq データの

細胞分類に対して効果的であることを示している。今後は他のデータに対する適応可能性を検討するとともに、細胞分類に寄与している分子の同定を行い、解釈性の向上を目指す。

#### 参考文献(業績)

1. Nakamura Y, **Miyashita S**, et al.  
Cerebrospinal fluid-contacting neuron tracing reveals structural and functional connectivity for locomotion in the mouse spinal cord. *Elife*. 2023
2. Nakata S, **Miyashita S**, et al., Epigenetic upregulation of *Schlafen11* renders WNT- and SHH-activated medulloblastomas sensitive to cisplatin. *Neuro Oncol*. 2022

## 分担研究報告書

(課題名) データサイエンスによる疾患予測モデリング

(所属) 長崎大学 情報データ科学部

(氏名) 植木 優夫

### 緒言

近年の統計的機械学習に代表されるデータサイエンスの目覚ましい発展に伴い、医学・生物学分野においてもその活用が進められている。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、疾患予測を目標とした統計モデリング・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では、内在する本質的な複雑さによって統一的な取り扱いが難しく、解析手法の開発自体が最先端の研究対象となる。本課題では、疾患予測を目標としたテラ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の開発を目指す。

### 方法

今年度は、データサイエンス技術の調査と理論的検討を行い、並行して技術開発と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動作確認し、同時に、実データを用いて妥当性を検証する。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

(1) 機械学習手法による医学・生物学データ解析について

データ解析手法のプログラム実装を行い、統計的機械学習手法による疾患予測モデリングの研究に取り組んだ。様々なモデリングによる予測性能をシミュレーションによって調べるとともに実際のデータにおいて動作を検証した。得られた知見を国際シンポジウム JOINT SYMPOSIUM Co-hosted by SixERS & USTP, Okayama, Japan (2022.12.9)において発表した。また、計算論的精神医学における機械学習モデルを用いた研究を行った[参考文献 1]。

(2) 時空間疫学データのモデリング手法の研究について

時空間データをモデリングする際には、時間方向の依存関係と空間方向の依存関係を同時に考慮する必要がある、注意深い統計モデリングが求められる。今年度は、新型コロナウイルス感染症の感染者数の時空間データに対して、ベータ負の二項分布を用いた新たな時空間データモデリング手法を開発し、統計学の専門誌に論文を発表した[参考文献 2]。

### 考察

今年度に得られた研究成果を発展させて、今後は、脳病態に関する臨床ビッグデータを用いた疾患予測のための統計モデリング手法の開発を推進する。機械学習手法による疾患予測性能の検証から、データの特性による手法の振る舞いを把握することができた。また、時空間データモデリング研究



で得られた知見は、空間依存性を有する脳画像の解析にも活用できるものと期待する。次年度は、臨床データと大規模遺伝子データを用いた疾患予測手法開発に取り組む。

#### 結論

脳病態研究領域におけるデータに対して、データの性質を加味したテーラーメイドな統計モデリングを行うことで、既存手法を上回る予測性能が得られることを期待している。

#### 参考文献

1. Takahashi Y, Murata S, Ueki M, Tomita H, Yamashita Y. Interaction between functional connectivity and neural excitability in Autism: a novel framework for computational modeling and application to biological data. *Computat Psychiatry*, 2023, 7:14-29.
2. Ueki M. Beta-negative binomial nonlinear spatio-temporal random effects modeling of COVID-19 case counts in Japan. *J Appl Stat*, 2022  
DOI:  
10.1080/02664763.2022.2064439

## 分担研究報告書

(課題名) 個別化医療に向けた高次元データモデリングの方法と理論

(所属) 産業技術総合研究所

(氏名) 片平健太郎

### 緒言

同じ診断がついた患者の間でもその病態や治療反応性に異質性があることが認識されはじめ、個人の特性に応じた治療を施す個別化医療への関心が高まっている。疾患の異質性を扱う方法として、クラスター分析を使用して特徴の類似した患者をグループ化する、サブグループ化が用いられている。精神疾患に関してもサブグループ化研究はこれまで多くなされている

(Marquand et al., 2016)。しかしながら、サブグループ化の方法やその妥当性を評価するため指標はまだ十分に検討されていない。本研究では、アウトカムの予測精度に着目し、計算機シミュレーションを通してその理論的性質を検討した (Katahira, 2023)。

### 方法

**シミュレーションによる検討** 患者の特徴量の分布やそれに対するアウトカムを生成する統計モデルを構築し、シミュレーションにより生成したデータをもとに、K-means や階層クラスター分析、混合正規分布等の代表的なクラスターリング手法を評価した。

**サブグループ分類の評価指標の提案** サブグループ分類に基づく予測精度を評価する指標である Cost of cluster-mean based

prediction (CCMP) を提案し、シミュレーションによりその妥当性を評価した。また、CCMP をもとに各種のクラスターリング手法を評価した。

**実データでの検討** 糖尿病に関するオープンデータを用いて、各種のクラスターリング手法、および異なるクラスター数の結果を評価した。

### 結果

シミュレーションの結果、予測する対象であるアウトカムが予測に用いる特徴量の線形関数である場合は、K-means 法が最適な予測を与えることが、シミュレーションおよび理論解析により示された。また、提案した指標である CCMP はサブグループ分類に基づく予測の精度をバイアスなく評価することが可能であり、最適なサブグループ分類を選択できることが示された。また、糖尿病のデータセットへの CCMP の適用の結果、従来のクラスター数の決定指標に基づいたクラスター数は必ずしも最適なサブグループ分類とはならず、K-means 法によるクラスターリングが最適な予測精度を与えることが示された。

### 考察・結論

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ー数やクラスター手法などが選択されていた。本研究で得られた CCMP は、アウトカムの予測の観点からサブグループ分類を評価するものであり、個別化医療のための疾患の分類に有用な指標となりうる。本研究は、特徴量の分布のモデルとしては適切ではないモデルが、アウトカムの予測のため

に有用である場合があることを示唆している。本研究の成果は、個別化医療のための疾患のサブグループ分類に有益な示唆をもたらすことが期待される。

#### 参考文献

- Marquand, A. F., Wolfers, T., Mennes, M., Buitelaar, J., & Beckmann, C. F. (2016). Beyond lumping and splitting: a review of computational approaches for stratifying psychiatric disorders. *Biological psychiatry: cognitive neuroscience and neuroimaging*, 1(5), 433-447.
- 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**

Our project studies on psychiatry, neurology, myopathy, and developmental disorders for comprehensive outcomes including finding brain pathophysiology, diagnoses, treatments, and preventions. Especially, we focus on the computational approaches: data science, AI methods, and brain computation. Our project also involves infrastructure building for the abovementioned objectives.

Our project consists of following three subjects:

- I. Building organization and infrastructure for researches on computational brain pathophysiology.
- II. Developing data-driven and AI-driven methods for clinical and fundamental brain pathophysiology.
- III. Integrating brain pathophysiological modeling and data-driven approaches.

**2. Research Outcomes**

**I. Organization and infrastructure building**

- a) We built the preparation office for Brain pathophysiology AI research center (BAIC) in NCNP designed the infrastructure. We also acquire human resources for BAIC.
- b) We located the preparation office for BAIC in NCNP and build the network and security

environment.

- c) We conducted preliminary surveys for seeking a project section chief of the preparation office for BAIC. Finally, we hired Tachimori Hisateru, Ph.D., from April 2023.
- d) We signed a contract with UK Biobank.
- e) We hosted the Seminar Series on Computational Theory and Data Science in Brain Pathophysiology. We held 6 seminars in FY 2022.

**II. Developing data-driven and AI-driven methods**

- f) We successfully clustered total RNA sequencing of specimens, including low-grade epileptogenic tumors, to express specimen characteristics.
- g) We proceeded the development of a muscle pathology identification application using a muscle pathology diagnostic algorithm, and created a dataset classifying the degree of disability of affected muscles according to the Mercuri score.
- h) We have launched development of a highly sensitive histone modification analysis system and a cell nucleus preparative technique using flow cytometry.
- i) By Combining eye movement and cognitive function data, we create a model that can

discriminate normal subjects and schizophrenia accurately.

- j) We surveyed and interviewed about the potential application of AI-driven methods in psychotherapy.
- k) We got the review and approval of the research plan for the psychiatric disease registry analysis and began data cleaning and analysis.

### III. Integrating brain pathophysiological modeling and data-driven approaches

- l) We developed a feature representation learning method using a multimodal variational autoencoder and a model of decision-making behavior and perceptual-cognitive processes using reinforcement learning and predictive information processing models. Furthermore, as an attempt to integrate these methods, we examined the interaction between parameters in the predictive information processing model using resting-state fMRI data.
- m) By applying Variational Autoencoder, we succeeded in reproducing cellular gene expression with high accuracy and started analyzing single-cell RNA sequencing data from human epilepsy patients
- n) We showed that spatio-temporal modeling using data science methods with nonlinear random effects and beta-negative binomial distribution can relaxes the issues of existing models.
- o) We proposed Cost of Cluster-Mean based Prediction (CCMP), a measure of the

accuracy of prediction of a given subtype classification, and its effectiveness is demonstrated through simulation and application to open data.