

特集 2 脳画像研究により精神疾患はどこまでわかったか、そして今何をやるべきなのか

3. 精神疾患のサブタイプや再定義をめざした階層的教師有・教師無機械学習法 ～調和された多施設多疾患安静時 fMRI データを用いた研究～

山下 歩^{1,2)}

抄録：深層学習や機械学習などの情報技術の発展が目覚ましく、精神疾患にかかわる脳画像研究においてもその有効性が示されつつある。例えば、安静時に撮像した機能的核磁気共鳴画像法 (functional magnetic resonance imaging : 安静時 fMRI) データに教師有学習を適用した生物学的基盤に基づいた精神疾患の判別や、教師無学習を適用した精神疾患のサブタイピングなどが行われている。しかしながら、安静時脳活動と行動指標の関係に関する効果量が小さい、脳画像データの施設間差、DSM 診断に基づく技術開発などの問題があった。本稿では、これらの問題点および脳画像研究により精神疾患はどこまでわかってきているのか、そして今何をやるべきなのかについて、筆者らが取り組んできた大規模多施設多疾患データセットの構築、脳画像データにおける施設間差を調和する harmonization 技術の開発、脳回路に基づく精神疾患のサブタイプ調査に関する取り組みを紹介し、精神疾患における脳画像研究の現状と課題、解決策について概説する。

日本生物学的精神医学会誌 34 (1) : 24-29, 2023

Key words : fMRI, harmonization, machine learning, big data

近年、情報技術の進展が目覚ましく、さまざまな分野で成果をあげている。例えば、2017年にはDeepMind社のAlphaGoが囲碁の世界チャンピオンに勝利し¹⁾、2021年にはAlphaFoldがタンパク質の構造をきわめて正確に予測できることを示した⁷⁾。また、2022年に開発されたAlphaTensorはこれまでヒトが発見できなかった効率的な行列演算アルゴリズムを発見することに成功した²⁾。おそらく10年以内に精神疾患の早期診断や個別化医療を可能とする情報技術が医療分野のゲームチェンジャーになるのではないかと筆者は考えている。一方で、現状の精神疾患研究領域において、実应用到に近づいている脳画像を用いた情報技術が大まかに二つ存在する。一つ目は、脳画像から推定した個人の機能的脳回路パターンから、個人が精神疾患であるかを予測する精神疾患の脳回路マーカー技術である。脳は安静にしているときにも自発的に活動しており (安静

時脳活動)、機能的核磁気共鳴画像法 (functional magnetic resonance imaging : fMRI) で計測することができる (安静時 fMRI)³⁾。安静時脳活動における脳領域間の時間相関により機能的結合を定義することができ (安静時機能的結合)、全脳間で機能的結合を計算したものが機能的脳回路パターンである (図1A)。安静時 fMRI は特殊な装置などが不要なく、MRIの中で眼を開けて10分間じっとしているだけの課題であるため、実験協力者への負担も小さく、一般診療として行われている器質因除のためのMRI撮像に加えて撮像を行うことが可能である。この脳回路パターンには個人差があることが知られており、健常群と精神疾患群で異なることが明らかとなってきた。脳回路マーカー技術は、機械学習技術などを用いることで脳回路パターンから個人の精神疾患度を算出し、疾患度が高ければ精神疾患群、低ければ健常群であると予測する技術である (図

Hierarchical supervised/unsupervised approach for subtype and redefine psychiatric disorders using resting state functional magnetic resonance imaging

1) 国際電気通信基礎技術研究所 (〒619-0288 京都府相楽郡精華町光台2丁目2番地2) Ayumu Yamashita : Advanced telecommunications research institute international (ATR). 2-2-2, Hikaridai, Seika-cho, Sorakugun, Kyoto 619-0288, Japan

2) 東京大学大学院情報理工学系研究科 システム情報学専攻物理情報学講座 (〒113-8656 東京都文京区本郷7-3-1) Ayumu Yamashita : Graduate School of Information Science and Technology the University of Tokyo. 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

【山下 歩 E-mail : ayumu722@g.ecc.u-tokyo.ac.jp】

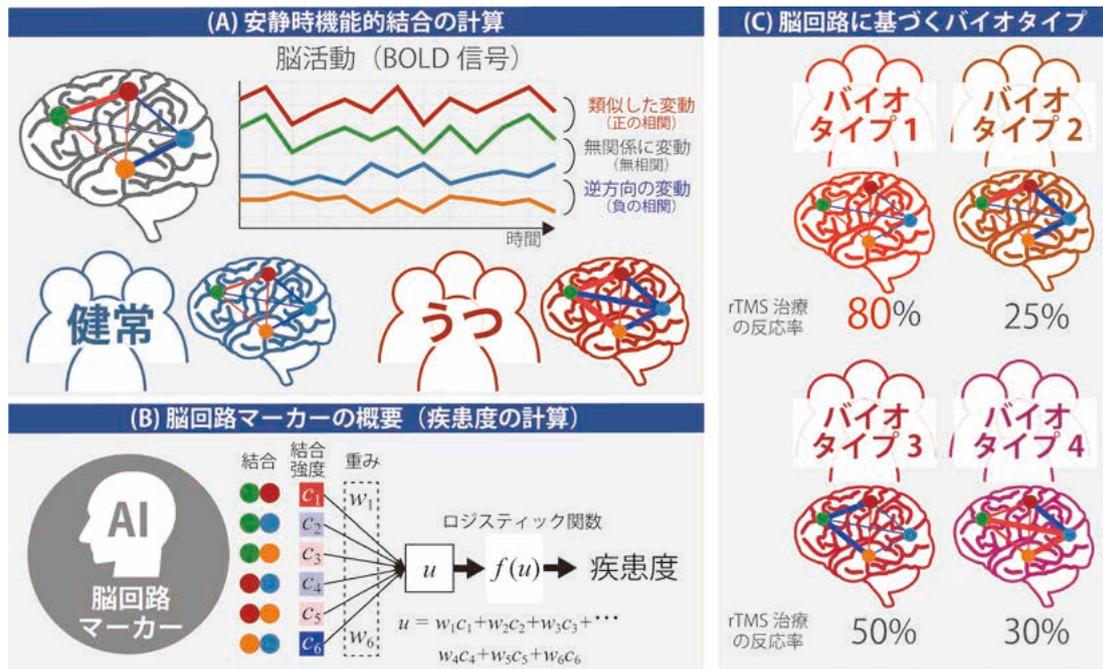


図1 安静時fMRIを用いた脳回路技術

(A) ここでは脳を4つの領域に分け、各脳領域の自発的な活動を表す。これらの活動は類似した変動をする領域間 (赤と緑) や、まったく独立した変動をする領域間 (緑と青)、逆方向に変動をする領域間 (青と黄色) がある。同期の程度を時間的な相関として定量化したものが機能的結合である。全脳で機能的結合を計算することで、個人の脳回路パターンを推定できる。(B) 各結合の強度に重みを掛け合わせ、足し合わせをロジスティック関数に入ると疾患度を出力する脳回路マーカの例。(C) 脳回路パターンに基づく大うつ病のバイオタイプ。バイオタイプごとにrTMSの反応率が異なる。

1B)。実際、筆者らの研究グループでは自閉症¹⁵⁾、大うつ病^{5, 16)}、統合失調症¹⁸⁾、強迫性障害¹³⁾などで脳回路マーカ技術の開発に成功している。2022年には、脳回路マーカの実応用に向けた取り組みとして臨床研究が開始されている。二つ目は、脳回路パターンに基づく疾患のサブタイピング技術である。同じ精神疾患群の中でも脳回路パターンが少しずつ異なっており、脳回路パターンをクラスタリングすることで精神疾患をサブタイプに分類することができる。例えば、Drysdaleらの研究では大うつ病群約300人を含む約1,000人の安静時fMRIデータを使用し、脳回路パターンに基づいてクラスタリングを行うことで、大うつ病が4つのバイオタイプに分けられることを示した¹⁾。このバイオタイプごとに反復経頭蓋磁気刺激療法 (repetitive transcranial magnetic stimulation: rTMS) の反応率が異なることが示されているため、治療法選択時にバイオタイプがわかれば個人ごとに適切な治療法の選択が可能となるだろう (図1C)。これらの結果は、fMRIのような脳画像から推定された脳回路パターンが、精神疾患の診断補助や個別化医療に役立つ可能性を示している。

しかしながら、脳画像ベースの情報技術を精神疾患医療に応用するためには、まださまざまな問題点が存在している。一つ目は単一機能的結合と行動指標の関係に関する効果量が小さい点である (問題点①)。先行研究では、数万人を超える大規模安静時fMRIデータを用いて機能的結合とさまざまな認知課題で計測された認知機能との相関関係を調査したところ、相関の大きな結合においても相関係数として0.1程度であることが明らかになっている。この程度の小さな効果量を再現性高く発見するためには数千人規模のデータが必要であると言われている⁹⁾。ただし、複数の機能的結合 (脳回路パターン) に基づいて行動との関連を調べることで数千人規模のデータは必要ないという主張も多くみられていることには注意が必要である^{10, 12)}。いずれにせよ、信頼性の高い結果を得るためには千人規模のデータは必要であるため、複数施設が協力してデータを収集することが必要となるだろう。しかし、多施設から取得した大規模安静時fMRIデータを解析してみると、撮像した脳画像データの施設間の違いが健常群と患者群の違いよりも大きいことが明らかとなってきた¹⁷⁾。これが二つ目の問題点、脳画像データの施

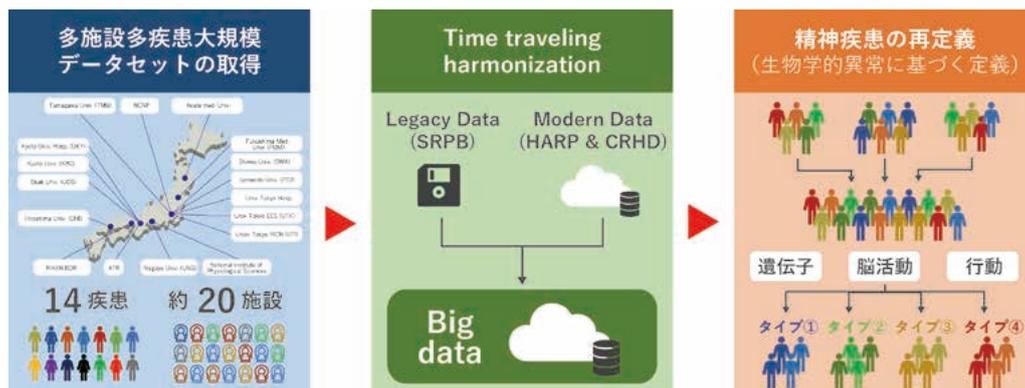


図2 現在の取り組み

設間差である(問題点②)。これは、撮像施設ごとのMRI装置や撮像方法などの計測の違いや、疾患や性別などの被験者の違いに起因することが明らかとなっている¹⁷⁾。例えば、この施設間差の問題により、一つの施設で得られた百人規模のデータを使用して構築された脳回路マーカーが、他施設で取得されたデータに汎化しないという問題が報告されている⁴⁾。臨床応用の観点から、多施設での妥当性や信頼性は必須なため、施設間差を減らす調和技術(harmonization技術)の開発が重要である。最後の問題点は、現在の多くの情報技術がDSMの診断に基づいて開発されている点である(問題点③)。DSMに基づく診断は診断一致率の低さや疾患の異質性などの問題が示唆されてきており、診断基準を見直す動きが活発化している。例えば、アメリカ国立精神衛生研究所は研究領域基準(Research domain criteria: RDoC)という精神疾患を理解するための枠組みを提案している⁶⁾。これは、精神疾患を確立された診断カテゴリーとしてではなく、人間の基本的な神経行動学的機能に基づいて理解しようという取り組みである。

このような問題点を克服するために、国立研究開発法人日本医療研究開発機構(Japan agency for medical research and development: AMED)脳科学研究戦略推進プログラム(脳プロ)では世界に先立ち2008年から統一撮像プロトコル(Strategic research program for brain [SRPB]プロトコル)を用いた多施設多疾患の安静時fMRIデータ取得を開始し、2020年までに全国14施設から5疾患(大うつ病、自閉症、強迫症、統合失調症、慢性疼痛)を含む1,627例の安静時fMRIデータを取得し、一般公開している¹⁴⁾。また、現在行われている戦略的国際脳科学研究推進プログラム(国際脳)では2018年から現在までに約20施設から14疾患(大うつ病、自閉症、強迫症、統合失調症、慢性疼痛、注意欠陥・

多動性障害、双極性障害、認知症、パーキンソン病、摂食障害、てんかん、軽度認知障害、精神病発症危険状態、不安症)を含む3,874例のデータを取得している⁸⁾。このデータセットは、米国で行われているHuman connectome project (HCP)やAdolescent brain cognitive development (ABCD)を参考にし、高時空間解像度の撮像プロトコル(Harmonized protocols [HARP]とConnectome related to human disease [CRHD])を用いており、安静時fMRIだけでなく拡散強調画像や課題時fMRIデータなども取得した多角的なデータセットになっている(図2)。このような統一的多施設多疾患データセットは世界的にも他に類を見ず、今後このデータセットを使用して、より高性能で信頼性の高い精神疾患の脳回路マーカー構築や、多疾患の関係性が明らかになることが期待される。

次に、問題②を解決するために、筆者らの研究グループでは同じ被験者が複数の施設に実際に赴き安静時fMRIデータを計測する旅行被験者データを取得している。施設間差は、MRIの違いや撮像方法の違いである計測バイアスと疾患や性別などの生物学的な違いであるサンプリングバイアスに分けられる。計測バイアスは測定に基づく差異であり小さいほうがよい一方で、サンプリングバイアスには重要な生物学的な効果も含まれるためなるべく保持したい情報であるが、単一データセットだけでは二つのバイアスを分離することは困難である。旅行被験者データを用いることでサンプリングバイアスを統制できるため、純粋な計測バイアスを推定することが可能となる。筆者らは旅行被験者を用いて推定した純粋な計測バイアスのみを補正するharmonization技術の開発に成功している¹⁷⁾。一方で、脳プロから国際脳のように長期的にデータ取得を行う場合、技術の向上に伴う撮像方法の改良は避けて通ることができない。例えば、脳プロではできるだけ多くの

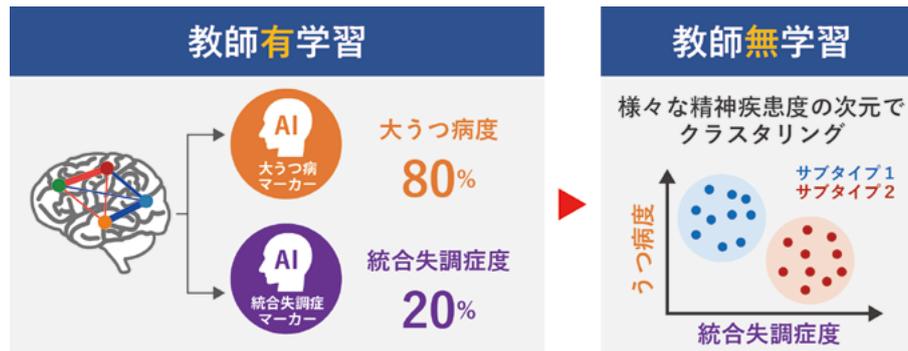


図3 脳回路に基づく精神疾患の再定義のための階層的教師有・教師無学習

MRI装置で計測することを目的とした比較的計測が容易な撮像プロトコル (SRPB) を用いたが、国際脳では国際連携を見据えた高時空間解像度の撮像プロトコル (HARPとCRHD) を用いている。そのため、これまで撮り貯めてきたデータをすべて有効活用するためには、施設間差だけでなく撮像プロトコル差を考慮する必要がある。例えば、HARPやCRHDプロトコルで撮像されたデータを用いて構築した脳回路マーカーがSRPBプロトコルで撮像されたデータに汎化するのかなどは検証すべき重要な事項である。また、従来のSRPBプロトコルで撮像されたデータと最先端のHARP・CRHDプロトコルで撮像されたデータを統合できれば、より大規模なデータセットを構築でき、より信頼性の高い脳回路マーカーの構築や多疾患の関係性の理解に繋がることが期待できる (図2)。国際脳における旅行被験者データセットでは、このようなtime traveling harmonizationも見据えて、同じ旅行被験者が3種類のプロトコル (SRPB, HARP, CRHD) で撮像を行っており、世代を超えたharmonization技術 (time traveling harmonization技術) の開発に現在取り組んでいるところである。

最後に、問題③を解決するために、筆者らが現在取り組んでいるRDoCの考え方に基づいた大規模多施設多疾患データセットを用いた研究に関して紹介する。前述の様な多施設の多大な協力の末に集められた大規模多施設多疾患データを用いることで、脳画像のような生物学的異常に基づいて精神疾患を見直す試みに着手することが初めて可能となっている。例えば、この多疾患データに対して脳回路パターンに基づいたクラスタリングのような教師無学習を行うことで、精神疾患の再定義を行うことが可能性として考えられる (図2)。しかしながら、一般的に脳回路には10,000から100,000結合が含まれ、データ数よりも大きいため、機械学習によって過学習が容易に引き起こされる危険性がある。ま

た、脳回路全体のパターンのクラスタリングでは、年齢に関するクラスターや (高齢群と若年群) 性別に関するクラスター (男性群と女性群) など、精神疾患にかかわらない群分けが生じる可能性も高い。我々の最終目標は、過学習を回避しまったく異なる撮像施設での汎化性能を確保しながら、精神疾患を生物学的次元で理解することである。そこで、筆者らの研究グループでは、階層的教師有・教師無学習法を用いて脳回路パターンに基づく精神疾患のサブタイピングを提唱している (図3)。この方法ではまず教師有学習を用いて、健常群と精神疾患群を信頼性高く区別する比較的少数の機能的結合を同定した。この段階での目的は、精神疾患と潜在的に関連する生物学的次元をみつけることである。したがって、大うつ病のようなDSMの診断を使用することは、関連する生物学的次元を探索するための有用な情報を提供する可能性がある。ここでは、少数の機能的結合を組み合わせた脳回路マーカーの出力を疾患度として、生物学的に意味のある次元として使用した。大うつ病、統合失調症、自閉症の3疾患で脳回路マーカーを構築し、大うつ病度、統合失調症度、自閉症度を精神疾患にかかわる三次元空間として定義した。異なる精神疾患をこの次元に位置づけることで、疾患間の関係を明らかにできる可能性がある。次に、精神疾患のサブタイプを発見するために、この低い生物学的次元に教師無学習である階層的クラスタリングを適用した。その結果、大うつ病度と自閉症度が高い群、大うつ病度だけが低い群、すべての疾患度が低い群、すべての疾患度が高い群、自閉症度のみが高い群などに分類されることが明らかとなった。今後は、これらの群間でどのような症状が異なるのか、どのような治療法が有効であるのかなどを検証し、これらのサブタイプの意義を検証していく予定である。

このように、長い時間をかけて多施設共同で取得してきたデータを適切なharmonization技術を用い

て調和した大規模多施設多疾患脳画像データに、機械学習法を適用することで、より信頼性の高い脳回路マーカーの構築や精神疾患のサブタイプの発見に繋がることを期待される。脳回路マーカーの構築や精神疾患のサブタイプの発見が、精神疾患の診断補助技術や個別化医療に役立つことを期待している。本論文に記載した筆者らの研究に関してすべて倫理的配慮を行っている。

謝辞：本文で紹介した研究は、AMED の課題番号 JP22dm0307002 の支援を受けて実施された。ご協力・ご支援をいただきました皆様に感謝申し上げます。

利益相反：開示すべき利益相反は存在しない

文 献

- 1) Drysdale AT, Grosenick L, Downar J, et al (2017) Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat Med*, 23 : 28-38.
- 2) Fawzi A, Balog M, Huang A, et al (2022) Discovering faster matrix multiplication algorithms with reinforcement learning. *Nature*, 610 : 47-53.
- 3) Fox MD and Raichle ME (2007) Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci*, 8 : 700-711.
- 4) He Y, Byrge L and Kennedy DP (2020) Nonreplication of functional connectivity differences in autism spectrum disorder across multiple sites and denoising strategies. *Hum Brain Mapp*, 41 : 1334-1350.
- 5) Ichikawa N, Lisi G, Yahata N, et al (2020) Primary functional brain connections associated with melancholic major depressive disorder and modulation by antidepressants. *Sci Rep*, 10 : 3542.
- 6) Insel TR and Cuthbert BN (2015) Medicine. Brain disorders? Precisely. *Science*, 348 : 499-500.
- 7) Jumper J, Evans R, Pritzel A, et al (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, 596 : 583-589.
- 8) Koike S, Tanaka SC, Okada T, et al (2021) Brain/MINDS beyond human brain MRI project : A protocol for multi-level harmonization across brain disorders throughout the lifespan. *Neuroimage Clin*, 30 : 102600.
- 9) Marek S, Tervo-Clemmens B, Calabro FJ, et al (2022) Reproducible brain-wide association studies require thousands of individuals. *Nature*, 603 : 654-660.
- 10) Rosenberg MD and Finn ES (2022) How to establish robust brain-behavior relationships without thousands of individuals. *Nat Neurosci*, 25 : 835-837.
- 11) Silver D, Schrittwieser J, Simonyan K, et al (2017) Mastering the game of go without human knowledge. *Nature*, 550 : 354-359.
- 12) Spisak T, Bingel U and Wager T (2022) Replicable multivariate BWAS with moderate sample sizes. *bioRxiv*. doi : 10.1101/2022.06.22.497072.
- 13) Takagi Y, Sakai Y, Lisi G, et al (2017) A neural marker of obsessive-compulsive disorder from whole-brain functional connectivity. *Sci Rep*, 7 : 7538.
- 14) Tanaka SC, Yamashita A, Yahata N, et al (2021) A multi-site, multi-disorder resting-state magnetic resonance image database. *Sci Data*, 8 : 227.
- 15) Yahata N, Morimoto J, Hashimoto R, et al (2016) A small number of abnormal brain connections predicts adult autism spectrum disorder. *Nat Commun*, 7 : 11254.
- 16) Yamashita A, Sakai Y, Yamada T, et al (2020) Generalizable brain network markers of major depressive disorder across multiple imaging sites. *PLoS Biol*, 18 : e3000966.
- 17) Yamashita A, Yahata N, Itahashi T, et al (2019) Harmonization of resting-state functional MRI data across multiple imaging sites via the separation of site differences into sampling bias and measurement bias. *PLoS Biol*, 17 : e3000042.
- 18) Yoshihara Y, Lisi G, Yahata N, et al (2018) Overlapping but asymmetrical relationships between schizophrenia and autism revealed by brain connectivity. *bioRxiv*. doi : 10.1101/403212.

■ ABSTRACT

Hierarchical supervised/unsupervised approach for subtype and redefine psychiatric disorders using resting state functional magnetic resonance imagingAyumu Yamashita^{1,2)}

- 1) *Advanced telecommunications research institute international (ATR)*
- 2) *Graduate School of Information Science and Technology*

Information technologies such as deep learning and machine learning have made remarkable progress, and their effectiveness has been demonstrated in brain imaging research related to psychiatric disorders. For example, the application of supervised learning to resting-state functional magnetic resonance imaging (fMRI) data has been used to identify psychiatric disorders based on their biological basis, and unsupervised learning has been used for subtyping of psychiatric disorders. However, there have been problems such as small effect sizes on the relationship between resting brain activity and cognitive functions, inter-imaging-site differences in brain imaging data, and development of the technologies based on DSM diagnoses. In this paper, I discuss these problems, the extent to which brain imaging studies have revealed psychiatric disorders, and what needs to be done now. I introduce our efforts to overcome these problems, constructing a large multi-imaging-site, multi-disorder dataset, developing a novel harmonization technique to mitigate inter-site differences in brain imaging data, and investigating subtypes of psychiatric disorders based on the brain circuit.

(Japanese Journal of Biological Psychiatry 34 (1) : 24-29, 2023)
