

平成29・30年度厚生労働科学研究費補助金(政策科学総合研究事業)
分担研究報告書

脳 MRI データとAI解析によるうつ病診断の汎用性向上とデータ駆動型サブタイプ分類に関する研究

研究代表者 岡本泰昌 広島大学医歯薬保健学研究科 精神神経医科学 教授
研究分担者 吉本潤一郎 奈良先端科学技術大学院大学先端科学技術研究科情報科学領域 准教授
研究協力者 徳田智磯 奈良先端科学技術大学院大学先端科学技術研究科情報科学領域 博士研究員
研究分担者 山脇成人 広島大学大学院医歯薬保健学研究科特任教授
研究分担者 池田和隆 東京都医学総合研究所分子精神医学分野長
研究分担者 橋本 亮太 大阪大学大学院大阪大学・金沢大学・浜松医科大学・千葉大学・福井大学連合
小児発達学研究科附属子どものこころの分子統御機構研究センター 准教授

研究要旨

脳機能画像および脳構造画像の比較的大規模なデータセットを、複数の AI アルゴリズムを用いて解析することによって、各データセットに対する最適な AI アルゴリズムを特定するとともに、それぞれのデータセットに対する測定バイアスについても検討した。また、安静時脳活動データから AI 技術を活用したうつ病診断の実用化と医師の診断を用いたうつ病判別技術の確立および診断を用いないデータ駆動的なうつ病関連予測技術の確立を目指して、解析パイプラインのためのパラメータ最適化、教師あり学習によるうつ病判別法および教師なしクラスタリングによるうつ病サブタイプ分類法の開発を行った。教師あり学習に関しては、うつ病患者群 93 名と健常者群 93 名の研究参加者から収集したデータを対象に、うつ病の判別器の作成を行い、うつ病と健常対照者を判別することができ、独立した外部データにおいても汎化性能も確認できた。教師なし学習に関しては、うつ病患者群 67 名と健常者群 67 の研究参加者から収集したデータを対象に、多重ベイズ共クラスタリングを適用し、データ駆動的なうつ病サブタイプ分類を試みた。その結果、抗うつ剤(セロトニン再取り込み阻害剤; SSRI)に対する治療反応性の良し悪しと対応付けられる 3 つのクラスタ(サブタイプ)が発見され、これらのクラスタは、直角回を中心とした 12 個の安静時機能結合性と幼児期トラウマ経験の大小によって特徴づけられていることを見出した。

A. 研究目的

うつ病は抑うつ気分と意欲低下に特徴づけられる精神疾患である。世界的に見ても全人口の 4.4%にあたる 3 億人以上がうつ病に苦しんでいるという高い罹患率に加えて、自殺の要因ともなりうることから¹⁾、うつ病に対する適切な診断と治療が不可欠である。しかしながら、がんや脳卒中などでは、血液検査や画像診断などの客観的な診断法が確立されているのに対して、うつ病では、客観的な診断法がまだ確立されておらず、面談や質問紙の内容を DSM²⁾に代表される診断基準や医師の経験と照らし合わせて診断しているのが現状である。

これまでにわれわれは、AI を用いた診断、治療反応予測に関して、数十例のうつ病と健常者の臨床データと課題遂行時や安静時の脳活動といったバイオデータを組み合わせて探索的な検討を行ってきた(Shimizu et al, 2015; Yoshida et al, 2017; Ichikawa et al, 2017)。しかし、脳機能画像以外のバイオデータは検討していないこと、サンプル数が十分でないこと、一部は外部テストデータでの汎化性能は検証していないなどの課題が残存している。

また、客観的な診断指標を確立しようと、我々はこ

れまで課題遂行時や安静時の fMRI 画像から、人工知能の一分野である機械学習法を活用して、DSM に基づく医師の診断結果を予測するアルゴリズムの探索的検討を行ってきた^{3,4)}。その結果、言語流暢性課題時の脳活動を用いたうつ病診断については、90%を超える予測精度を、安静時脳活動を用いたうつ病診断については、機能結合性を特徴量としたアンサンブル学習法の導入によって約 70%の予測精度を達成することができた。一方で、fMRI の原画像から特徴量となる機能結合性を算出し、診断モデルを構築する過程では、全脳をどのように領域分割するか、多施設でデータ取得する際に生じる測定バイアスをどのように除去するかなどのパラメータ設定によって、汎化性能が大きく異なる。このパラメータ設定を最適化することは、診断のための解析パイプラインを確立する上で、重要な問題であるものの、これまで網羅的なパラメータ探索は行ってこなかった。

それに加え、うつ病は診断内の異質性が非常に高い。例えば、うつ病治療の臨床現場での第一選択肢は抗うつ剤による薬物療法であるが、薬物療法のみで 2-4 ヶ月の急性期に寛解に達する者は 50%未満であるという現状がある⁵⁾。したがって、単純に診断だけ

にとどまらず、うつ病内の異質性(サブタイプ)を分類し、治療効果予測につながる特徴量がデータ駆動的に発見できるのが理想である。

以上の問題解決に向けて、0)各データセットに対する最適な AI アルゴリズムの特定と、それぞれのデータセットに対する測定バイアスの検討、1)解析パイプラインの最適化に向けたパラメータの網羅的探索、2)医師の診断情報を用いた教師あり学習法に基づくうつ病分類法の開発と外部汎化性能の検証および、3)医師の診断情報を用いない教師なし学習法に基づくデータ駆動的なうつ病サブタイプ分類法の開発に取り組み、その有用性を検証した。

B. 研究方法

【検討1:各 MRI データセットに対する最適な AI アルゴリズムの特定と測定バイアスの検討】

安静時脳機能画像を用いたうつ病患者と健常者の判別(診断)に関しては、高次元データを低次元の部分空間に射影することで、分類精度を向上させることが期待されたため、いくつかの分類方法について検討した。すなわち、ナイーブな線形判別分析法(Linear Discriminant Analysis ; LDA)、サポートベクターマシン(Support Vector Machine; SVM)、最小二乗回帰法(ordinary least squares regression; OLS)、部分最小二乗回帰法(partial least squares; PLS)、さらに2次多項式カーネル(KPLS-Poly(2))、3次多項式カーネル(KPLS-Poly(3))、Gaussianカーネル(KPLS-Gauss)を用いた部分最小二乗回帰法などの、精度、感度、特異度を比較検討した。また9名のトラベリングサブジェクトを用いて異なる施設のMRIスキャナーで測定がおこなない、施設間の測定バイアスとうつ病と健常対照者の差異と比較した。

脳構造画像に関しては、広島大データにVoxel-based morphometry (VBM)解析を行い、サポートベクターマシンを用いて、うつ病患者と健常者の判別の感度と特異度とともに、測定を行ったスキャナーの寄与度もあわせて検討した。

【検討2:解析パイプライン用パラメータの網羅的探索】

広島大学および連携医療機関を受診した148名のうつ病患者群、および、広島大学で募集した269名の対象健常者群から、MRIを用いて安静時脳活動データを収集した。

データ収集は、異なる場所に設置された、撮像条件も異なる4つのfMRIスキャナを用いて実施した(表1)。各研究参加者のデータはこのうちの1つのfMRIスキャナを用いて収集されたものである。MRI撮像時、研究参加者には、眠らざりリラックスした状態で、できる限り何も考えないように教示した。合わせて、セッションごとに開眼、閉眼のいずれかの状態であるよう教示した。

Parameter	Site1	Site2	Site3	Site4
Participants (Patients / HC)	67 / 68	14 / 49	23 / 34	44 / 118
MRI scanner	GE Signa HDxt	GE Signa HDxt	Siemens Magnetom	Siemens Verio
Magnetic field (T)	3.0	3.0	3.0	3.0
Field of view (mm)	256	256	192	212
Matrix	64x64	64x64	64x64	64x64
Number of slices	32	32	38	40
Number of volumes	150	150	112	244
In-plane resolution (mm)	4.0 x 4.0	4.0 x 4.0	3.0 x 3.0	3.3 x 3.3
Slice thickness (mm)	4.0	4.0	3.0	3.2
Slice gap (mm)	0	0	0	0
TR (ms)	2,000	2,000	2,700	2,500
TE (ms)	27	27	31	30
Total scan time (mm:ss)	5:00	5:00	5:03	10:10
Flip angle (deg)	90	90	90	80
Slice acquisition order	Ascending (Interleaved)	Ascending (Interleaved)	Ascending (Interleaved)	Ascending (Interleaved)

表 1: 安静時脳活動データ収集に用いた4種類のfMRI撮像条件

撮像された各fMRI画像に対して、まず、スライスタイミング補正、体動補正、標準脳への位置合わせ、空間フィルタによる平滑化などの一般的な前処理を行うことによって、各ボクセル単位のBOLD信号時系列を算出した。その後、後述するいずれかの脳領域分割法によって定義される各脳領域内の空間平均を取ることで、各脳領域の活動時系列を得て、2領域間の時間変動に対する相互相関係数をその2領域間の安静時機能結合性として計算した(例えば、全脳を90脳領域に分割する場合には、 $90 \times 89 / 2 = 4005$ 組の安静時機能結合性が得られる)。その後、安静時機能結合性を入力データ、DSM基準に基づく医師の診断結果(うつ病罹患の有無)を出力教師データとして、アンサンブル学習法の一つであるRandom Forestにより診断モデルを構築した。

以上の手続きにおいて、探索したパラメータ種類とそれぞれの設定値とは以下の通りである。

- 計測時の開眼・閉眼条件
 - ① 開眼優先: 開眼時のセッションのみのデータのみを利用する
 - ② 閉眼優先: 閉眼時のセッションのみのデータのみを利用する
- 脳領域分割法
 - ① BAL: 文献⁶⁾と同じ方法を用いる
 - ② Stanfordx90: Stanford大学が公開している90領域に分割した機能脳領域^{7,8)}を用いる
 - ③ Stanfordx499: Stanford大学が公開している499領域に分割した機能脳領域⁷⁾を用いる
- 攪乱変数除去
 - ① SCCA(スパース正準相関分析法): 文献⁶⁾と同じ方法を用いる
 - ② 一般線形モデル: 最小二乗線形回帰法により攪乱変数による影響を除去する
- 変数選択法
 - ① SCCA: 文献⁶⁾と同じ方法を用いる
 - ② 二群検定法: 各特徴量ごとに学習データ

内の健常者群とうつ病患者群の間に統計的有意差が認められたもののみを診断モデルの入力変数として用いる

以上のパラメータ設定の各組合せに対して、診断モデルの予測性能は、leave-one-site-out 交差検証法を利用して、精度 (accuracy) と感度 (sensitivity) により評価した。

【検討3: SLR+SCAA 法によるうつ病判別と外部データへの汎化性能の検証】

安静時の脳機能結合を用いたバイオマーカー作成に関して、広島大学 4 施設より集められた MRI データの内、躁病、薬物依存、アルコール依存、精神病性障害、パーソナリティ障害等の併存疾患を除外した後の症例 93 名と年齢性別を合わせた健常者 93 名の合計 186 名が解析対象とされた。被験者は、広島市周辺地域のクリニック (うつ病患者) と地方紙による広告 (健常者) から集め、Yahata et al (2016) の判別器作成用のアルゴリズムと同様な方法を用いて、うつ病バイオマーカーの作成を行った。具体的には、脳を 140 個の小領域に分割し、1 人 1 人について各領域における機能的 MRI 信号の時間波形を取り出し、それらが任意の 2 領域間でどの程度似ているか相関係数として数値化した。140 個の小領域の全てのペア (9, 730 個) について相関係数 (機能的結合) を計算することで、個人の脳全体の回路を定量できて、全脳の回路図 (=9, 730 個の数値からなるベクトル) が作成される。これを研究参加者全員分について求め、開発した人工知能技術 (SCAA+SLR) を適用した。外部独立したデータセットは放医研のデータを利用した。

【検討4: ベイズ共クラスタリング法によるデータ駆動的うつ病サブタイプ推定】

前節で述べた研究参加者の部分集合にあたるうつ病患者群 67 名と健常者群 67 名については、安静時 fMRI データに加えて、うつ病の重症度を評価する臨床指標 (HRSD, BDI) や幼児期トラウマ体験を指標化した CATS (Child Abuse Trauma Scale)、血液サンプルより測定した遺伝子多型や BDNF メチル化レベルなどの生理指標も計測・取得し、参加者数 134 人 × 特徴量 2948 次元のデータ行列を構成した (安静時 MRI データ以外の特徴量については表 2 を参照)。

Numerical features	
age	age when first depressive symptoms show up.
	The number of days elapsed for current episode.
	BAS* (Behavioral Activation Scale).
	BDNF (Quantity of brain-derived neurotrophic factor in blood).
	BDI* (Beck Depression Inventory).
	BIS* (Behavioral Inhibition Scale).
	CATS (Child Abuse and Trauma Scale).
	CATS:total for all items
	CATS:N for items on neglect
	CATS:S for items on sexual abuse
	CATS:P for items on punishment
	CATS:E for items on emotion
	Cortisol (Quantity of cortisol in blood).
	CpG., S., SU., SD. (Methylation probability).
	FC1-2701 (Functional connectivity).
	GAF* (Global Assessment of Functioning).
	PHQ9* (Patient Health Questionnaire).
	HRSD17* (17-item Hamilton Rating Scale for Depression).
	HRSD21* (21-item Hamilton Rating Scale for Depression).
	HRSDchange (Increment rate of HRSD17 scores defined as (HRSD17 _{6w} - HRSD17)/HRSD17).
	JART* (Adult reading test).
	LES (Life Experiences Survey).
	LES:total for all events
	LES:P for positive events
	LES:N for negative events
	PANASP* (Positive Affect Schedule).
	PANASN* (Negative Affect Schedule).
	SHAPS* (Snaith-Hamilton Pleasure Scale).
	STAI* (State-Trait Anxiety Inventory).
	N*, E*, O*, A*, C* (Five factors in revised NEO Personality Inventory)
Categorical features	
	BDI*_i1_ (Items of BDI initially).
	BDI*_i2_ (Items of BDI after six weeks of treatment).
	drug* (states of dosing of lexapro).
	HRSD*_i1_ (items of HRSD initially)*.
	HRSD*_i2_ (items of HRSD after six weeks of treatment)*.
	Melancholic* (Melancholic depression or not).
	MINI* (Mini-International Neuropsychiatric Interview): the numbering corresponds to the following psychiatric symptoms.
	Major depressive disorder (1).
	Dysthymia (2), Suicide risk (3).
	Mania (4), Panic disorder (5), Agoraphobia (6).
	Social phobia (7).
	Obsessive compulsive disorder (8).
	PTSD (9), Alcohol dependence and abuse (10).
	Drug dependence and abuse (11).
	Psychotic disorder (12), Anorexia (13), Bulimia (14).
	Generalized anxiety disorder (15).
	Antisocial personality disorder (16).
	Recurrent* (Recurrent depression or not).
	Response* (whether there is response to the treatment based on HRSD17).
	Remission* (whether a patient is remitted after the tremens).
	Sex
	SNPs* 1-8: Single Nucleotide Polymorphisms that are located in the following genome sites, respectively. (in parenthesis are the relevant gene functions)
	rs1187323 (NTRK2), rs34118353 (5HT1a receptor), rs3756318 (NTRK2), rs3813929 (5HT2c receptor), rs4554739 (NTRK2), rs56384968 (SLC6A4), rs6265 (BDNF), rs6294 (5HT1a receptor)
Integer features	
	Episode (the number of past experiences of depression).
	RecNum (the number of times of recurrent depression)

表 2: ベイズ共クラスタリング解析に用いた安静時 fMRI データ以外の特徴量のリスト

このデータ行列に対して、ベイズ推定の原理を用いて、参加者と特徴量を同時にクラスタリングし、クラスタ数も自動決定できる多重ベイズ共クラスタリング⁹⁾を適用した。クラスタリング後、患者群と健常者群の区別が最も良くできるクラスタに注目し、そのクラスタに含まれる安静時機能結合性を同定した。また、各クラスタと抗うつ剤 (セロトニン再取り込み阻害剤; SSRI) に対する治療反応性の関係性について統計的に解析した。

(倫理面への配慮)

本研究は、広島大学研究倫理委員会の承認を得て実施し、すべての参加者から書面による同意を得ている。

C. 研究成果

【検討1: 各データセットに対する最適な AI アルゴリズムの特定と測定バイアスの検討】

安静時脳機能画像を用いていくつかの分類方法を組み合わせて、精度、感度、特異度を用いて、その仮説を検証した (図1)。

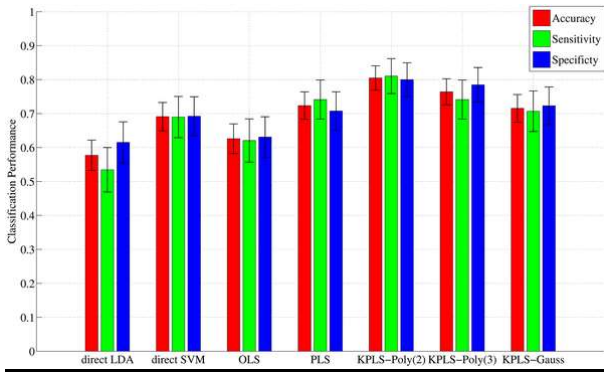


図1 うつ病患者と健常者の判別成績

ナイーブな線形判別分析法 (Linear Discriminant Analysis ; LDA) (精度 57.7%、感度 53.4%、特異度 61.5%) およびサポートベクターマシン (Support Vector Machine; SVM) (精度 69.1%、感度 69.0%、特異度 69.2%) と比べて、2 次多項式カーネルを用いた部分最小二乗回帰法 (Kernel Partial Least Squares regression with the 2nd order Polynomial; KPLS-Poly (2)) は有意に優れた精度を示した (精度 80.5%、感度 81.0%、特異度 80.0%)。この結果は、分類モデルを構築するために臨床測定の予測モデルを利用することの有用性を示唆している。さらに、KPLS-Poly (2) は、通常の最小二乗回帰法 (Ordinary Least Squares Regression; OLS) (精度 62.6%、感度 62.1% および特異度 63.1%) と比べて有意に優れた精度を示し、回帰モデルの潜在空間を考慮することの有用性が示唆された。

安静時脳機能画像の測定バイアスに関しては、異なる施設スキャナーで安静時脳機能画像データを収集した場合、施設間の測定バイアスはうつ病などの精神疾患と健常者の違いと同程度かさらに大きかった。

脳構造画像を用いたうつ病患者と健常者の判別 (診断) に関しては、予備的な検討ではあるが、精度 75.9% (感度 78.1%、特異度 72.9%) で、異なる施設の スキャナー (Scan10,21,22, 23, 30,40,41) で測定した影響は小さかった (図2)。

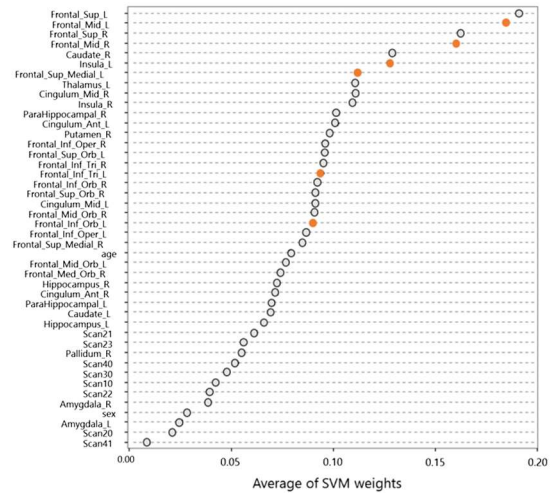


図2 SVM を用いたうつ病診断への寄与度

【検討2; 解析パイプライン用パラメータの網羅的探索】

B.節で述べたあらゆるパラメータ設定の組合せに対して診断モデルの予測精度を評価した。その結果、開眼状態での安静時 fMRI データに対して、BAL⁶⁾ の脳領域分割法による機能結合性を定義したものを特徴量として、診断モデルへの入力前に一般線形モデルによる攪乱変数除去と有意水準 1% の二群検定法による特徴量選択の前処理を施した時が最適となり、その時の精度は 0.75、感度は 61% となった (表3)。

開眼/閉眼	脳領域分割	変数選択	攪乱変数除去	精度	感度
開眼	BAL	二群検定法 ($\alpha=0.01$)	一般線形モデル	0.7483	0.6081
閉眼	BAL	SCCA	SCCA	0.7196	0.6028
開眼	Stanfordx90	二群検定法 ($\alpha=0.01$)	一般線形モデル	0.6962	0.4918
閉眼	BAL	SCCA+二群検定法 ($\alpha=0.03$)	SCCA	0.6869	0.5354
開眼	Stanfordx90	SCCA+二群検定法 ($\alpha=0.05$)	SCCA	0.6801	0.4621
閉眼	Stanfordx90	SCCA	SCCA	0.6747	0.4874
開眼	BAL	二群検定法 ($\alpha=0.03$)	一般線形モデル	0.6737	0.5068
閉眼	BAL	SCCA+二群検定法 ($\alpha=0.01$)	SCCA	0.6729	0.5181
開眼	BAL	SCCA+二群検定法 ($\alpha=0.01$)	SCCA	0.6643	0.4589
閉眼	BAL	SCCA+二群検定法 ($\alpha=0.05$)	SCCA	0.6589	0.5272
開眼	BAL	SCCA+二群検定法 ($\alpha=0.05$)	SCCA	0.6573	0.4797
閉眼	BAL	二群検定法 ($\alpha=0.03$)	一般線形モデル	0.6542	0.5068
閉眼	Stanfordx90	SCCA	SCCA	0.6434	0.4394
閉眼	Stanfordx499	SCCA+二群検定法 ($\alpha=0.05$)	SCCA	0.641	0.4058
閉眼	Stanfordx90	SCCA+二群検定法 ($\alpha=0.05$)	SCCA	0.6381	0.4347
閉眼	Stanfordx90	二群検定法 ($\alpha=0.05$)	一般線形モデル	0.6381	0.459
閉眼	BAL	SCCA+二群検定法 ($\alpha=0.03$)	SCCA	0.627	0.4604
閉眼	Stanfordx499	SCCA+二群検定法 ($\alpha=0.05$)	SCCA	0.6247	0.438
閉眼	Stanfordx499	SCCA	SCCA	0.62	0.4072
開眼	Stanfordx499	二群検定法 ($\alpha=0.05$)	一般線形モデル	0.6084	0.4392

表3: 解析パイプラインのための網羅的パラメータ探索結果 (精度に関して上位 20 位までを抜粋)

【検討3; SLR+SCAA 法によるうつ病判別と外部データへの汎化性能の検証】

全ての対象では判別率 51%、AUC 0.52 にとどまったため、うつ病の中でも生物学的要因の影響が大きいとされるサブタイプ、メランコリー型の特徴を有する一群に絞った検討を行った。メランコリー型うつ病 66 名と同数の健常者を合わせた計 132 名の安静時 fMRI データを用いて判別器の作成を行い、一個抜き交差検証 (LOOCV) を行った

結果、判別率 70%、AUC:0.77 まで成績が向上した。このバイオマーカーを、完全に独立な施設のデータセットに適用した結果、判別率は 65%(AUC 0.62)であり、汎化性能が確認された。判別器として抽出された 12 の脳機能結合の中には、うつ病の先行研究において重要であることが示唆されている脳領域が多数含まれていた。さらに、このバイオマーカーの臨床的意義や有用性を明らかにするために、判別器の値 (Weighted Linear Sum: WLS) とうつ病重症度との関連について検討した結果、この値は抑うつ症状の質問紙得点 (BDI) と相関を示し、6 週間の抗うつ薬治療により健常方向に変化したことから、うつ状態を反映するバイオマーカーである可能性が示唆された。

【検討4: ベイズ共クラスタリング法によるデータ駆動的うつ病サブタイプ推定】

B.節で述べたデータ行列に対して、多重ベイズ共クラスタリングを適用した結果、15 種類の共クラスタ構造が得られた(図 3)。

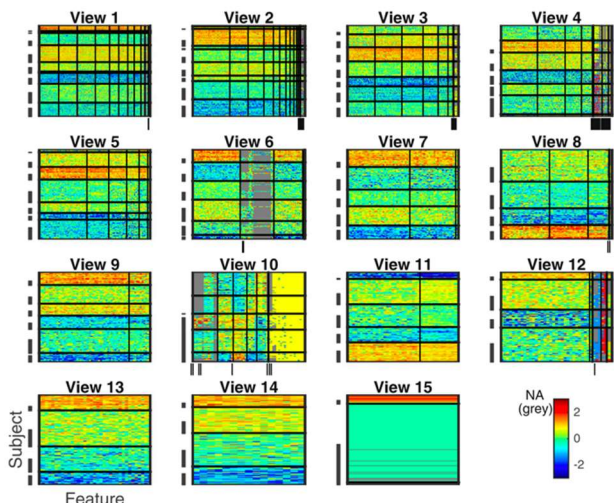


図 3: 多重ベイズ共クラスタリングの適用結果(縦軸は参加者、横軸は特徴量を表しており、参加者軸に付随したダッシュ記号はうつ病患者を表している。ヒートマップ中の太線はクラスタの境界を示している。)

このうち、View 10 で示されるクラスタ構造は、他の View に比べても特にうつ病患者群と健常者群の分類とも一致しており、最初の 2 つのクラスタは健常者群のみで構成されている一方で、残りの 3 つのクラスタに属する参加者のほとんどがうつ病患者となった。

そこで、View 10 で示されるクラスタ構造について、より詳細に調べた。まず、このクラスタ構造に含まれる安静時機能結合性を抽出したところ、右角回を中心とするスポークアンドハブ構造の機能ネットワーク(図 4)がこのクラスタを特徴づけていることが分かった。

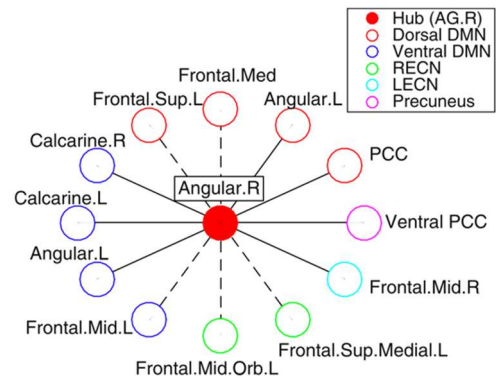


図 4: 図 3 の View 10 のクラスタ構造に含まれる安静時機能結合ネットワーク

また、ほとんどがうつ病患者から構成される 3 つのクラスタの違いを調べたところ、BDI や HRSD によって評価されたうつ病重症度の初診時と SSRI 投薬開始後 6 週間(または 6 ヶ月)後の差、すなわち、SSRI に対する治療反応性がこれらのクラスタの特徴を反映していることが分かった。また、これら 3 つのクラスタの違いには、幼児期トラウマ指標である CATS も関連していることがわかった。そこで、View 10 に含まれる機能結合性の第 1 主成分と CATS の 2 次元でデータの分布を可視化したところ、図 5 のようになった。すなわち、右角回を中心とした機能ネットワークにより、D3 クラスタ(治療反応性が良いクラスタ)への割当てが決まり、残りの 2 つのクラスタについては、CATS が小さければ D2 クラスタ(治療反応性が良いクラスタ)へ、CATS が大きければ D1 クラスタ(治療反応性が悪いクラスタ)へと割当てが決まる構造があることが分かった。

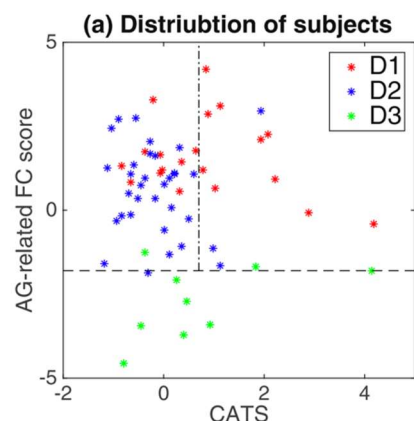


図 5: 右角回を中心とした機能ネットワークの第 1 主成分と幼児期トラウマ指標 (CATS) の散布図(各点の色の違いは、View 10 内で所属するクラスタの違いを表している)

D. 考察

各データセットに対する最適な AI アルゴリズムの特定と測定バイアスの検討、解析パイプラインの最適化

に向けたパラメータの網羅的探索、および、医師の診断情報を用いない教師なし学習法に基づくデータ駆動的なうつ病サブタイプ分類法の開発に取り組んだ。

検討1の最適なAIアルゴリズムの特定と測定バイアスの検討の結果から、安静時脳機能画像に関しては、臨床測定の予測モデルを媒介すること、低次元の特徴空間を考慮することで、比較的小サンプルでも分類性能の向上が期待できることを明らかにした。また、安静時脳機能画像は測定バイアスの影響が大きく、スパース推定による特徴選択、施設効果の線形回帰による regression-out、アンサンブル学習、traveling subjects を用いた測定バイアスの分離と推定といった数理統計手法の利用が測定バイアスの補正に必要であると考えられた。一方、構造画像に関しては、SVM に適用した結果からは異なる施設のスキナーで測定した影響(測定バイアス)は小さいことが明らかになった。

検討2のパラメータの網羅的探索の結果から、安静時 fMRI 撮像時では開眼状態の方が、最適な予測精度が実現できるという意味では望ましいことが示唆される。しかしながら、他のパラメータに依存して閉眼時の方が予測精度が良くなる場合もあり、実用上、開眼・閉眼のどちらが良いかについては一貫性のある結果は得られなかった。また、脳領域分割法については BAL の結果が総じて良かった。変数選択法と攪乱変数除去法については、一般線形モデルによる攪乱変数除去と二群検定法(有意水準 1%)の組合せが最適な結果となったが、この組合せを変えても、BAL による脳領域分割法を用いている限りは精度が大きく落ちるということにはなかった。したがって、現状の結果からは、BAL を用いた脳領域分割が特に推奨される。一方で、昨年度までの結果から入力変数の変数選択(または低次元化)と測定バイアス除去のための攪乱変数除去法の導入は必須であるものの、その方法については大きく精度の改善や悪化につながるようなものは同定できなかった。

検討3の結果から、完全な独立コホートに汎化するメランコリア特徴を有するうつ病の安静時機能結合の判別器(図6)を作成した。この判別器の12結合の線形加重和は、抑うつ症状と有意な正の相関があることから、状態を反映するバイオマーカーであると考えられた。

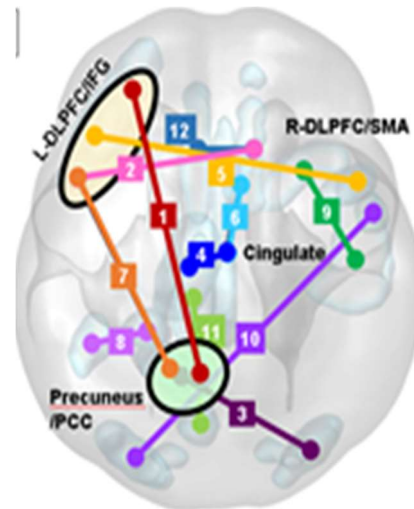


図6: 完全な独立コホートに汎化するメランコリア特徴を有するうつ病の安静時機能結合の判別器

検討4のサブタイプ分類法の開発においては、データ駆動的に治療反応性の違いに応じてうつ病患者群が3つのグループに分類できることを見出した。また、この3つのグループは右角回を中心とした12個の安静時機能結合性と幼児期トラウマ経験の大小によって特徴づけられていた。現在のところ、まだ少サンプルの解析結果であるため、その再現性は確認できていないが、これらの結果は、SSRI に対する治療反応性予測するためのモデルとして図7の決定木が有望であることを示唆している。

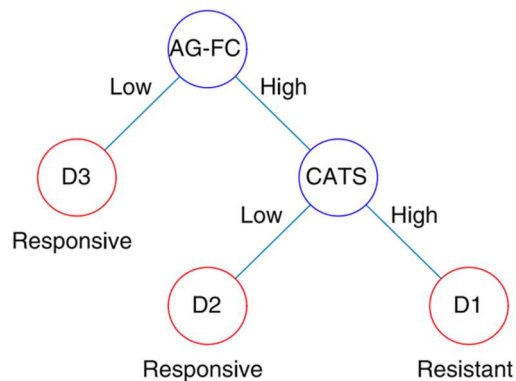


図7: 右角回を中心とした機能ネットワークの安静時機能結合性と幼児期トラウマ指標に基づく治療反応予測モデル(AG-FC は右角回を中心とした12個の安静時機能結合性のスコアであり、CATS は幼児期トラウマ指標である。)

E. 結論

本研究では、まず、検討1において、安静時脳機能画像に関しては、スパース推定 L1-Sparse Canonical Correlation Analysis(L1-SCCA) & Sparse Logistic Regression (SLR)による特徴選択、施設効果の線形回帰による regression-out、あるいは traveling subjects を用いた施設効果の測定バイアスの

分離と推定といった数理統計手法の利用がデータ解析に際しては必要と考えられた。また、測定バイアスに対する分類器の頑健性はアンサンブル学習の導入によって改善されることが期待された。一方、脳構造画像に関しては、これらの工夫や補正の必要性が小さいものと考えられた。

検討2では、解析パイプラインの最適化に向けたパラメータの網羅的探索を行った。その結果、脳領域分割法の選択が最も予測精度に良し悪しに影響し、現在利用できるものの中では BAL[®]が推奨されることが分かった。fMRI 撮像時の眼の開閉状態については、どちらを採用しても大きな違いはなく、変数選択法と攪乱変数除去法の導入は必須ながらも、その具体的な方法については、合理的なものであれば顕著な違いはなかった。

検討3では、安静時脳機能結合を用いて脳全体にわたる機能的結合のうち、ごく一部(わずか 0.2%)の結合から個人の『うつ度』(バイオマーカー)を測り、その大小でうつ病と健常対照者を判別することが可能となった。これらは独立した外部データにおいて汎化性能も確認できた。

検討4では、DSMに基づく医師の診断結果を明に用いない、教師なし学習法に基づくデータ駆動的なうつ病のサブタイプ分類をベイズ共クラスタリング手法を用いて試みた。その結果、治療反応性の違いに応じてうつ病患者群が3つのグループに分類できることを見出し、そのグループの違いは右角回を中心とした12個の安静時機能結合性と幼児期トラウマ経験の大小によって特徴づけられていることを見出した。

F. 健康危険情報

なし

G. 研究発表

1. 学会発表・招致講演等

Ichikawa N, Lisi G, Yahata N, Okada G, Takamura M, Hashimoto R, Yamada T, Yamada M, Suhara T, Moriguchi S, Mimura M, Yoshihara Y, Takahashi H, Kasai K, Kato N, Yamawaki S, Seymour B, Kawato M, Morimoto J, Okamoto Y. A classifier of melancholic depression with whole-brain resting-state connectivity. rtFIN2017, Nara, Japan, 2017.12.1, 市川奈穂, 岡本泰昌. 安静時fMRI活動を用いたうつ病の判別, 第15回日本うつ病学会, 東京, 2018.7.27 岡本泰昌, 市川奈穂. 脳機能画像研究からみたうつ

病の異種性, 第15回日本うつ病学会, 東京, 2018.7.28

岡本泰昌. 神経回路病態に基づくうつ病の診断・治療法の開発, 第26回脳の世紀シンポジウム『AIと脳』, 東京, 2018.9.12

Ichikawa N, Lisi G, Yahata N, Okada G, Takamura M, Hashimoto R, Yamada T, Yamada M, Suhara T, Moriguchi S, Mimura M, Yoshihara Y, Takahashi H, Kasai K, Kato N, Yamawaki S, Seymour B, Kawato M, Morimoto J, Okamoto Y. Melancholic depression biomarker of resting-state functional connectivity. AsCNP-ASEAN2019, Yogyakarta, Indonesia, 2019.3.2

H. 知的財産権の出願・登録状況

知的財産の内容(うつ症状の判別方法、うつ症状のレベルの判定方法、うつ病患者の層別化方法、うつ症状の治療効果の判定方法及び脳活動訓練装置)、種類・番号 PCT/JP2018/36952、出願年月日 2018.10.02、取得年月日、権利者国立大学法人広島大学、国際電気通信基礎技術研究所(ATR)

参考文献

- 1) World Health Organization. (2017). Depression and other common mental disorders: global health estimates. World Health Organization.
- 2) American Psychiatric Association. (2000). Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision (DSM-IV-TR).
- 3) Shimizu, Y., et al. (2015). Toward Probabilistic Diagnosis and Understanding of Depression Based on Functional MRI Data Analysis with Logistic Group LASSO. PLOS ONE, 10(5), e0123524.
- 4) Ichikawa N., et al. (2017). Identifying melancholic depression biomarker using whole-brain functional connectivity. arXiv:1704.01039.
- 5) Trivedi, M. H., et al. (2006). Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: Implications for clinical practice. American Journal of Psychiatry, 163(1), 28-40.
- 6) Yahata, N., et al. (2016). A small number of abnormal brain connections predicts adult autism spectrum disorder. Nature Communications, 7, 11254.
- 7) Neuropsychiatric Disorders Lab at Stanford University. Functional ROIs, https://findlab.stanford.edu/functional_ROIs.html.
- 8) Shirer W. R., et al. (2012). Decoding Subject-

Driven Cognitive States with Whole-Brain Connectivity Patterns. *Cerebral Cortex*, 22(1), 158–165.

- 9) Tokuda, T., et al. (2017). Multiple co-clustering based on nonparametric mixture models with heterogeneous marginal distributions. *PLOS ONE*, 12(10), e0186566.
- 10) Yoshida K, Shimizu Y, Yoshimoto J, Takamura M, Okada G, Okamoto Y, Yamawaki S, Doya K. Prediction of clinical depression scores and detection of changes in whole-brain using resting-state functional MRI data with partial least squares regression. *PLoS One*. 2017, 12: e0179638.