

脳情報に機械学習を用いることによる、 精神疾患(強迫症/強迫性障害)患者の判別

高木 優^{*123} 酒井 雄希^{*14} 阿部 能成^{*4} 西田 誠司^{*4} 中前 貴^{*14} 成本 迅^{*14}
 Yu Takagi Yuki Sakai Yoshinari Abe Seiji Nishida Takashi Nakamae Jin Narumoto

山田 恵^{*5} 福居 顯二^{*4} 田中 沙織^{*1}
 Kei Yamada Kenji Fukui Saori C. Tanaka

^{*1} 国際電気通信基礎技術研究所脳情報通信総合研究
 Computational Neuroscience Laboratories, Advanced Telecommunications Research Institute International

^{*2} 奈良先端科学技術大学院大学情報科学研究科
 Graduate School of Information Science, Nara Institute of Science and Technology

^{*3} 日本学術振興会特別研究員 DC1
 Research Fellow of Japan Society for the Promotion of Science

^{*4} 京都府立医科大学大学院医学研究科精神機能病態学
 Department of Psychiatry, Graduate School of Medical Science, Kyoto Prefectural University of Medicine

^{*5} 京都府立医科大学大学院医学研究科放射線診断治療学
 Department of Radiology, Graduate School of Medical Science, Kyoto Prefectural University of Medicine

We hypothesized that obsessive-compulsive disorder (OCD) is caused by impulsive choices in inter-temporal decision problems, e.g., maladaptive preference of immediate relief from anxiety by compulsions over a future normal life by temporarily restraining their compulsions. Combining machine learning techniques, we determined whether we could distinguish patients with OCD from healthy controls using functional network related to delay discounting that were obtained by resting-state functional magnetic resonance imaging. First, we detected the functional networks with which we predicted an individual's discount factor. Then using the detected network, we successfully made a high performance classifier to distinguish patients with OCD from healthy controls. Our findings suggest that OCD might exhibit altered neural organization related to delay discounting.

1. はじめに

精神医学分野では近年、疾患の症状を数理的なモデルで記述することにより、一見複雑に見える精神疾患の異常な行動の原理を解き明かそうとする研究が盛んである[1]。その対象とされる精神疾患の一つに、長時間の手洗いや鍵の確認などの繰り返しの行動を特徴とする強迫症/強迫性障害 (obsessive-compulsive disorder; OCD)がある。OCD の生涯有病率は2~3%とされており、患者の生活に与える影響は大きい[2]。その一方、OCD には確たる数理的モデルが存在しておらず、生物学的な理解もまだ進んでいないことから、依然として難治例も多く存在することが知られ、より有効な治療法が求められている。

我々は OCD を説明する可能性のあるモデルとして、時間割引モデルに着目している。時間割引モデルでは、パラメータの時間割引率が大きいと、将来の報酬を大きく割引いてしまい衝動的な行動を取ることで、小さいと抑制的な行動を取ることが説明できる。OCD の強迫行為は、不安を和らげるために衝動的に行われるが、そうした衝動性は時間割引率の異常と解釈できる。

更に OCD には、代表的な神経伝達物質であるセロトニンや、セロトニンの影響を強く受ける線条体という脳部位の異常が多く報告されており[3]、どちらも時間割引率との関連が深いことから[4][5]、分子生物学・脳神経科学的見地からも OCD と時間割引の関連が示唆される。

時間割引と脳画像研究について、先行研究では主に被験者に認知的課題を行わせた際の脳活動が機能的核磁気共鳴法 (functional magnetic resonance imaging; fMRI)を用いて調べられてきた[5]。実際、我々は過去に、健常者を人為的に低セロトニン状態へと操作することによって、時間割引課題において衝動的な選択が増加することに加え[4]、線条体の活動において衝動的な時間割引に関わる脳活動が優勢になっている [5]ことを発見した。更に我々は、OCD 患者が同様の課題を行った時の線条体の活動が、低セロトニン状態の健常者と同様に衝動的な脳活動が優勢になっていることを発見した[Sakai et al., submitted]。一方、近年の脳画像研究分野では、被験者が安静にしている時の脳活動(安静時脳活動)と、認知機能[6]や精神疾患[7]との関連を示す研究が盛んである。将来の臨床応用を

連絡先: 高木 優, 奈良先端科学技術大学院大学情報科学研究科,
 JSPS DC1, takagi.yu.tn0@is.naist.jp

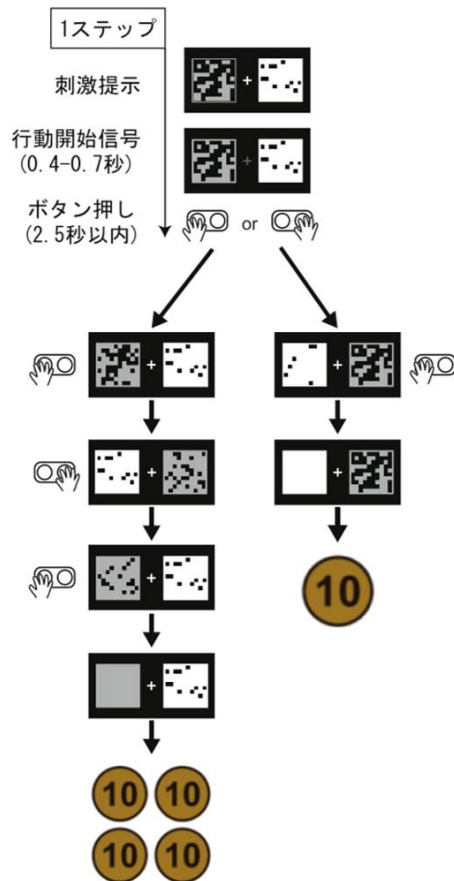


図1 時間割引課題の流れ各試行の始まりにおいて、スクリーンには中央の固視点を挟んで、左右に複数の黒色のモザイクで覆われた白色と灰色（実験では黄色を提示）の正方形が現れる。固視点が赤色になったと同時に被験者は右・左どちらかのボタンを押す。すると次の画面では、選択したボタンに対応する正方形のモザイクの数が減る。このステップを何回か繰り返す、どちらかの正方形のモザイクが完全になくなった時点で、白色の正方形なら10円、灰色の正方形なら40円の報酬を獲得する。被験者は一定時間中により多くの報酬を得ることを要求され、モザイクの初期状態、1ステップあたりモザイクの減る数と報酬の関係から、どちらの色を選ぶのが良いかを考えながらボタンを選択する。二つの選択肢の遅延時間は、灰色の正方形の平均遅延時間が白色の正方形の平均遅延時間よりも大きいという制約の下、各試行においてそれぞれ一様分布からランダムに選ばれた。

考える上で、撮像が簡便かつ短時間で終わる安静時脳活動の特性を明らかにする利点は大きい。

OCDと時間割引についても、それぞれ安静時脳活動の異常または関連を示す研究が存在する[6][8]。しかし、それら三者の関連を探った研究はこれまで存在しない。加えて、OCDと安静時脳活動の研究に関しては、特徴量作成の際の閾値設定方法に恣意性がある問題や、汎化性能が検証されていない問題がある[9]。また、時間割引と安静時脳活動の研究を行ったLiらは、予測の特徴量として用いる脳領域を時間割引課題中に賦活する脳領域に限定した[6]が、時間割引のような高次機能には幅広い脳領域の関与が考えられるため、他の脳領域が予測に寄与する可能性も否定できない。また、用いたアルゴリズムの性質

上、特徴量として用いた脳領域内における相対的な重要性も示されていない。

そこで、本研究では、OCDと安静時脳活動及び時間割引との関係を明らかにするために、多数の被験者から、fMRI計測による安静時脳活動データを取得した。更に、一部の被験者には時間割引課題も実施し、時間割引率を推定した。その後、異なる二つの手法でOCD患者と健常者を判別した。第一に、データ駆動的な手法として、時間割引の情報を用いずにOCD患者と健常者を判別する解析を行った。第二に、仮説駆動的な手法として、まず個人の時間割引率を安静時脳活動から予測し、その予測に有効であった脳情報を用いてOCD患者と健常者健常者と患者を正しく判別できるか調べた。いずれの解析においても、脳全体の情報を用いて解析を行ったが、この時、サンプル数に比して特徴量数が膨大になってしまう問題が生じる。このことによる過学習を避けるため、モデルの重みにスパース制約を課した機械学習アルゴリズムを用い、汎化性能の向上を狙った[10][11]。

2. 実験

2.1 被験者

京都府立医科大学精神医学教室において、63名の健常者（平均年齢:27.87歳±7.4SD, 男性:38名）及び同附属病院に通院中の48名のOCD患者（平均年齢:32.81歳±10.27SD, 男性:17名）に参加頂いた。実験は、京都府立医科大学全被験者の同意及び京都府立医科大学医学倫理委員会の承認に基づき、全被験者へ文書によって説明の上同意を得て行った。

2.2 fMRI 実験

全被験者は約6分40秒間の安静時脳活動実験に参加した。被験者はfMRIの中で、目を閉じ、深くものごとを考えず、寝ないように指示され、実験後の自己申告により寝たと判断された被験者は解析から除外した。全てのOCD患者は撮像中に強迫症状が生じてないことを実験後に医師が確認した。撮像装置はAchieva 3T (Philips社)を使用し、撮像方法はGradient-echo echo planner imaging法 (TR/TE/FA/Slice number/Matrix/FOV/Pixel size/Slice thickness/Slice order = 2000ms/30ms/80degree/39/64×64/192mm²/3×3×3/3.0mm (Gap:0mm)/Ascending)を使用した。また、実験初日に各実験参加者の構造画像を1×1mmの解像度で撮像した。

2.3 時間割引課題

fMRI実験に参加した被験者のうち48名に対し、先行研究と類似の時間割引課題を実施した[5]。図1に課題の流れを示す。

3. データ解析

3.1 時間割引係数の導出

個人の時間割引係数を導出するために、まず時間割引課題における各被験者の選択肢を即時報酬における遅延(D_s)と遅延報酬における遅延(D_L)の平面にプロットした。その後、以下のロジスティック関数により、遅延報酬を選ぶ確率値(P_L)を説明するモデルをあてはめた。

$$P_L = 1/(1 + \exp[-(\beta_L D_L + \beta_S D_S + \beta_0)])$$

ここで、P_L=0.5の時、

$$D_L = -\beta_S/\beta_L D_S - \beta_0/\beta_L \quad (1)$$

指数関数型の時間割引モデル及び線形の効用関数u(*)を仮定すると、遅延(Delay)に対する個人の主観的な割引価値は

$u(\text{報酬}) \times \gamma^{\text{Delay}}$ と表すことができ、このとき遅延報酬と即時報酬の値が等しい水準を示す式は次の形式となる。

$$D_L = D_S \cdot a / \ln \gamma, a = \ln[u(40)/u(10)].$$

ここで、 $u(40)=40$, $u(10)=10$ である。時間割引係数は切片に含まれる γ であり、式(1)における $-\beta_0/\beta_L$ と対応している。今回我々は個人の時間割引係数を $\gamma = \exp(-a\beta_L/\beta_0)$ として推定した。

3.2 安静時脳活動データからの特徴量抽出

全撮像データは、Statistical Parametric Mapping 8 (SPM8: Welcome Department of Cognitive Neurology, <http://www.fil.ion.ucl.ac.uk/spm/software/>) を用いて安静時脳活動解析における標準的な前処理を施した。具体的には、Realign, Slice-timing correction, Coregistering, Normalization 及び Smoothing (FWHM=8mm) を行った。更に、各被験者の全画像について、直前の画像から 0.5mm 以上動いたと判断される画像を解析から除外した[12]。

前処理を行った安静時脳活動データから、各被験者について、異なる関心領域 (Region of interest, ROI) 間の機能的結合 (Functional connectivity, FC) を計算した。FC は安静時脳活動解析において一般的に利用される特徴量であり、異なる ROI の時系列信号間のピアソンの相関係数により定義される。まず、各 ROI に含まれる全 Voxel の平均信号の時系列を抽出した。次に信号値からのノイズ除去のため、バンドパスフィルタ ($0.009 < f < 0.01$) をかけ、その後、9 つの説明変数 (全脳、白質、脳脊髄液の平均信号及び 6 つの体動補正パラメータ) による回帰を行った。回帰後の残差系列を機能的結合と関連する時系列信号値と考え、異なる ROI の時系列間でピアソンの相関係数を計算した。

ROI について本研究では、Brain Sulci Atlas (BAL) [13] に含まれる 137 の ROI に加え、Automated Anatomical Labeling Atlas [14] の小脳 (左右) 及び虫部の ROI を用いた。これら合計 140 の ROI 間の FC を特徴量として用いる場合をベースラインとして、本研究では、三つの異なる観点から ROI の組合せを用意した。第一に、BAL の線條体に加えて、先行研究を参考に、より細かく機能的に分割した ROI を二つ用意した。一つ目に、他の脳領域との FC のパターンを特徴量として k-means 法 ($k=6$ に設定) によりクラスタリングした線條体の ROI を、二つ目に、酒井らが時間割引モデルを用いて発見した、線條体の機能分化マップを用意した[15][Sakai et al., submitted]。第二に、脳部位の一つである小脳は信号値に欠損が出る場合が多いため、小脳の ROI を加える組合せと加えない組合せの二通りを用意した。最後に、OCD と時間割引のどちらも線條体との関連が強く示唆されることから、全 ROI 間で FC を計算する場合に加えて、線條体を ROI として含む FC のみで計算を行う場合も用意した。以上、合計 $3 \times 2 \times 2 = 12$ 通りの ROI の組合せを用意し、その間で計算された FC をその後の解析の特徴量として利用した。

3.3 データ駆動的に OCD 患者と健常者を判別

12 種類の特徴量の組合せそれぞれにつき、まず、データ駆動的な手法として、時間割引の情報を用いずに OCD 患者と健常者を判別する解析を、八幡・森本ら[Yahata and Morimoto, submitted]と同様の手法を用いて行った。本研究では、111 名のサンプル数に比して、特徴量の数は約 400 から 10000 あり、通常の多項線形判別モデルでは過学習により汎化性能が低下する恐れがある。そこで、モデルの重みにスパース制約を課すアルゴリズムである Sparse Logistic Regression (SLR) [10] を解析の内部で用いている。SLR は、多数の特徴量の中から判別に寄与しない特徴量を自動的に削除し、判別に有効な特徴量の

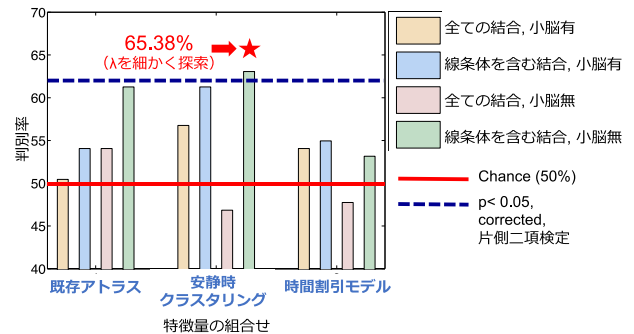


図3 データ駆動的な手法を用いた時の、特徴量の組合せ別の判別率(%)

線條体の分割方法	判別率	
	SLR	SVM
既存アトラス	46.8	57.7
結合パターン	47.7	55.9
時間割引モデル	54.1	61.3 ($p < 0.05$, uncorrected)

図4 時間割引の情報を入れた仮説駆動的な手法を用いた時の、特徴量の組合せ別の判別率(%)

みを抽出して、学習する推定手法である。SLR には、特徴量に比してサンプル数が少ない場合に未知のデータに対する汎化性能が向上する、学習時に調節すべきパラメータが無く、全自動で学習が出来るなどのメリットがある。判別精度の検証には Leave one out cross-validation (LOOCV) 法を用いた。また、八幡・森本の手法内で、スパース度合いをコントロールするためのパラメータ (λ) を決定する必要があるが、これはトレーニングセット内で 20 分割交差検定法により決定した。 λ は 0 から 1 の実数で、本研究では 0.005 から 0.250 の間を 0.005 刻みで探索した。

3.4 時間割引係数の予測に有効な特徴量から OCD 患者と健常者を判別

次に、まず時間割引係数を予測するモデルを作成し、その予測に有効な特徴量のみを用いて OCD 患者と健常者を判別するモデルを作った。時間割引係数の予測においても、判別と同様にサンプル数に比して特徴量の数が大きいと、重みにスパース制約を課すアルゴリズムである Sparse Linear Regression (SLR) を用いた。LOOCV 法を用いて各サンプルの予測値を計算し、予測値と実測値とのピアソンの相関係数により精度を検証した。

その後、相関係数が有意 ($p < 0.05$, 片側検定, uncorrected) に 0 より大きいモデルに限り、LOOCV を行った結果作られたサンプル数分のモデルにおいて、ゼロ以外の重みが有意な回数付与された FC を選別した[16] ($p < 0.001$, 片側二項検定, Bonferroni corrected)。その後、選別した FC のみを用いて SLR 及び Support Vector Machine (SVM) [17] により OCD 患者と健常者の判別モデルを作成し、LOOCV 法により判別精度を検証した。

4. 結果

4.1 OCD 患者と健常者のデータ駆動的な判別

時間割引の情報を使わずデータ駆動的に判別した結果を図3に示す。12 種類の FC の組合せのうち最高の成績を出したのは、FC のパターンによりクラスタリングされた線條体を使い、小

脳を含まず、線条体を含む FC のみを用いた組合せであった。この時 63.06%の精度で OCD 患者と健常者を判別可能であった ($p < 0.05$, 片側二項検定, Bonferroni corrected)。また、参考値として、最高の判別率が得られた組み合わせに限定して、より細かいスケール (0.005 から 0.250 の間を 0.001 刻み) で超パラメータ (λ) を探索した時は、65.38%とより高い判別成績を出すことができた ($p < 0.05$, 片側二項検定, Bonferroni corrected)。

4.2 時間割引モデルを用いた仮説駆動的な OCD 患者と健常者の判別

時間割引係数を FC から予測し、その予測に寄与した FC のみで患者と健常者を判別した結果を図 4 に示す。12 の組合せのうち、個人の時間割引係数を有意に回帰出来たモデルは三つあり、それらは、全ての FC を含み、小脳を除外した組合せであった ($p < 0.05$, 片側検定, uncorrected)。それらのモデルからスパースに抽出された特徴量を用いて、SLR と SVM を使って判別を試みた結果、時間割引モデルによる線条体の分割方法を用いた場合が最も判別成績が高く、特に、SVM を使った場合に高い判別成績を達成し、この時 61.3%の精度で OCD 患者と健常者を正しく判別出来た ($p < 0.05$, 片側二項検定, uncorrected)。また、時間割引係数の回帰に関してスパースに選ばれてきた FC に含まれる脳領域を見てみると、先行研究[6]では用いていない視覚や運動系に関わる脳領域も含まれていた。特に、両側外側前頭前野は偶然選ばれる確率に比べ有意に多い回数選ばれていた ($p < 0.05$, 両側二項検定, Bonferroni corrected)。

5. 考察・結論

本研究では、fMRI により取得された安静時脳活動データから、OCD 患者と健常者を完全にデータ駆動的に判別することに成功した。また、時間割引係数を回帰する特徴量を発見した上で、それにより OCD 患者と健常者を判別することにも成功した。いずれの解析においても、重みにスパース制約を課すアルゴリズムを用いることによって、脳全体の情報を使い特徴量選択の恣意性をなくすと同時に、交差検定による汎化性能が確認出来た。

時間割引の情報を使わずに行った判別において最も高い判別精度を出したのは、線条体が含まれる FC のみを用いた組合せであった。一方で、時間割引係数の回帰においては脳全体の FC を使うことが有効に働いた。この違いは、OCD 患者の症状が線条体を中心に局所的に表現されているのに対し、時間割引のような健常者にも存在するより一般的な個人の性質は脳全体に表現されているため、生じたものだと考えられる。

今後の方向性としては以下の二つが考えられる。第一に、本研究では日本国内の単施設のデータセットを用いたが、この結果が他施設・他国のデータセットに汎化するかどうかを検証する必要がある。第二に、今回特徴量の組合せとして 12 種類の FC の組合せを用意しその全てについて探索的に最高の判別率となる超パラメータを求めたが、この 12 種類の組合せの選別に関しても交差検定を行い、より適切な汎化性能を検証していく必要がある。

本研究は、OCD の生物学的な背景の理解に役立つだけでなく、OCD を説明する行動モデルとしての時間割引モデルの有用性も示すことが出来たと考えられる。これらの知見を活かすことにより、今後疾患の理解や治療に役立てていくことが期待される。

謝辞

本研究は、文部科学省脳科学研究戦略推進プログラム「BMI 技術を用いた自立支援、精神・神経疾患等の克服に向けた研究開発」、平成 22-25 年度「大阪大学社会経済研究所共同利用・共同研究費」及び文部科学省科学研究費補助金 新学術領域研究「予測と意思決定の脳内計算機構の解明による人間理解と応用」により実施された成果である。

参考文献

- [1] D. Lee, *Neuron*, **78**(2), 2013
- [2] Ruscio et al., *Mol Psychiatry*, **15**(1), 2010
- [3] Goddard et al., *Drug Discov Today*, **13**(7-8), 2008
- [4] Schweighofer et al., *J Neurosci*, **28**(17), 2008
- [5] Tanaka et al., *PLoS ONE*, **2**(12), 2007
- [6] Li et al., *J Neurosci*, **33**(11), 2013
- [7] Fox and Greicius, *Front Syst Neurosci*, **4**(19), 2010
- [8] Beucke et al., *JAMA Psychiatry*, **70**(6), 2013
- [9] Whelan and Garavan, *Biol Psychiatry*, **75**(9), 2014
- [10] Yamashita et al., *Neuroimage*, **42**(4), 2008
- [11] Toda et al., *Neuroimage*, **54**(2), 2011
- [12] Power et al., *Neuroimage*, **59**(3), 2013
- [13] Perrot et al., *Med Image Anal*, **15**(4), 2011
- [14] Tzourio-Mazoyer et al., *Neuroimage*, **15**(1), 2002
- [15] Kahnt et al., *J Neurosci*, **32**(18), 2012
- [16] Yamashita et al., *Sci Rep*, **5**, 2015
- [17] Cortes and Vapnik, *Machine learning*, **20**(3), 1995