

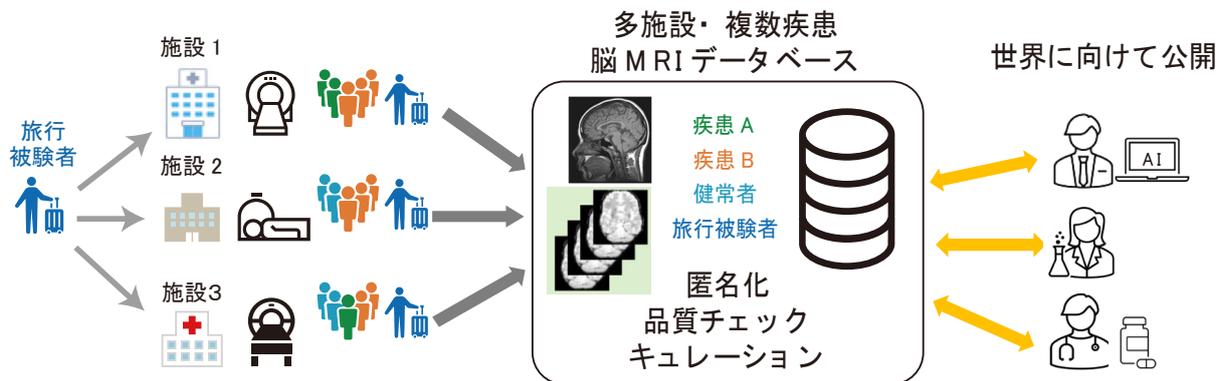
株式会社国際電気通信基礎技術研究所（ATR）
 国立大学法人東京大学
 国立研究開発法人日本医療研究開発機構（AMED）

多くの施設で集めた複数疾患の脳画像ビッグデータを一般公開 ～共通の脳回路マーカー開発促進で様々な精神疾患の診断・治療に貢献～

8月30日午前10時（英国時間）に Scientific Data 誌・オンライン版に掲載

本研究成果のポイント

- ✓ 近年、人工知能技術の発展と、ビッグデータ収集の環境整備により、様々な分野で人工知能が活躍しています。
- ✓ 医療分野でも、脳の fMRI 画像に人工知能技術を活用することで、疾患の診断に用いる試みが盛んになっています。
- ✓ この技術の実用化には、確かな再現性のため多施設において共通の撮像方法で複数疾患の患者から集めたビッグデータが必要となりますが、これまで整備されていませんでした。
- ✓ 本研究グループは、多施設で複数疾患の患者から集めた数千人規模の脳画像データをオンライン上で一般公開しました。さらに、測定方法の違いによる施設間差を除去するために同一の被験者が上記の多施設で撮像した「旅行被験者データ」も合わせて公開しました。
- ✓ また多くの研究者に広く使ってもらえるように、脳の領域間の活動の同期の強さを表す指標に変換したデータセット、前処理前の脳画像データセットと、必要性和用途に応じて使用できる複数のデータセットを用意しました。
- ✓ 精神疾患の脳回路マーカーの開発等に使用可能な公開データセットとして、精神疾患と発達障害の診断及び治療に貢献することが期待されます。



概要

田中沙織室長 (ATR 脳情報通信総合研究所 認知機構研究所 数理知能研究室)、笠井清登教授 (東京大学大学院医学系研究科 / 東京大学国際高等研究所 ニューロインテリジェンス国際研究機構 (IRCN)) らの研究グループは世界に先駆けて、多施設にて統一の撮像プロトコルで撮像した複数精神疾患の fMRI^[1] データ (データベース全体で 14 施設 2414 撮像、公開はそのうち最大 1627 撮像) および、旅行被験者データ (9 名の被験者が 12 施設で合計 143 撮像) を合わせてデータベース化し、オンラインプラットフォーム上 (<https://bicr.atr.jp/decnefpro/data/>) で公開しました。

背景

近年の科学全体の大きな動向のひとつに「オープンサイエンス」があります。オープンサイエンスは、科学的な知見をより広く社会へ伝えるという大きな概念ですが、より専門的にはオープンアクセスとオープンデータ (データシェアリング) を指します。その背景として、科学研究の再現性や透明性の担保があります。特に近年、様々な分野で活躍する人工知能技術の発展には、オープンデータが重要な役割を果たしています。人工知能技術を適用して再現性の高い結果を得るには、ビッグデータが不可欠となりますが、自施設のみで収集するには限界があり、オープンデータの必要性が認識されています。医療分野でも、脳の fMRI 画像に人工知能技術を活用することで、疾患の診断に役立つ試みが盛んになっていますが、疾患を対象とした大規模なオープンデータは限られており、再現性の担保は大きな課題となっていました。例えば単一施設で撮像された数十人程度の fMRI データに人工知能技術を適用して得られた結果は、他の施設で撮像したデータに適用した場合、再現できません。さらに、異なる施設で集めたデータには、撮像方法や機種などの違いが含まれており、単にデータを大量に集めて解析するだけでは、施設の影響のみを除くことができません。これらの問題を解決するには多施設において共通の撮像方法で複数疾患の患者から集めたビッグデータと、施設の影響を評価するために一人の被験者が多施設で撮像したデータ (旅行被験者データ) が必要となります。しかし現状、これらを満たす fMRI 公開データセットは整備されていませんでした。

研究内容

本研究では、14 施設にて統一の撮像プロトコルで撮像した複数疾患 (自閉スペクトラム症、大うつ、双極性障害、統合失調症、強迫症、慢性疼痛、脳卒中など) の MRI データを多施設・複数疾患データベースとして整備しました (図 1)。このデータベースは 2414 例 (疾患患者 993 例、健常者 1421 例) の安静時 fMRI データ^[2]、構造 MRI データ、デモグラフィックデータ (性別、年齢、利き手、臨床評価尺度) から構成されます。2414 例の被験者全員から、多施設・複数疾患のデータを研究プロジェクトに参画する研究機関内で共有し・合わせて解析する同意を得ております。このうち、公開の同意をいただいている被験者のデータについて公開しました。

さらに、測定の違いによる施設間差を除去するために同一の被験者が上記の多施設で撮像した「旅行被験者データ」(9 名の被験者が 12 施設で合計 143 撮像) も合わせてデータベース化し、非制限公開データセットとしてオンライン上で公開しました。

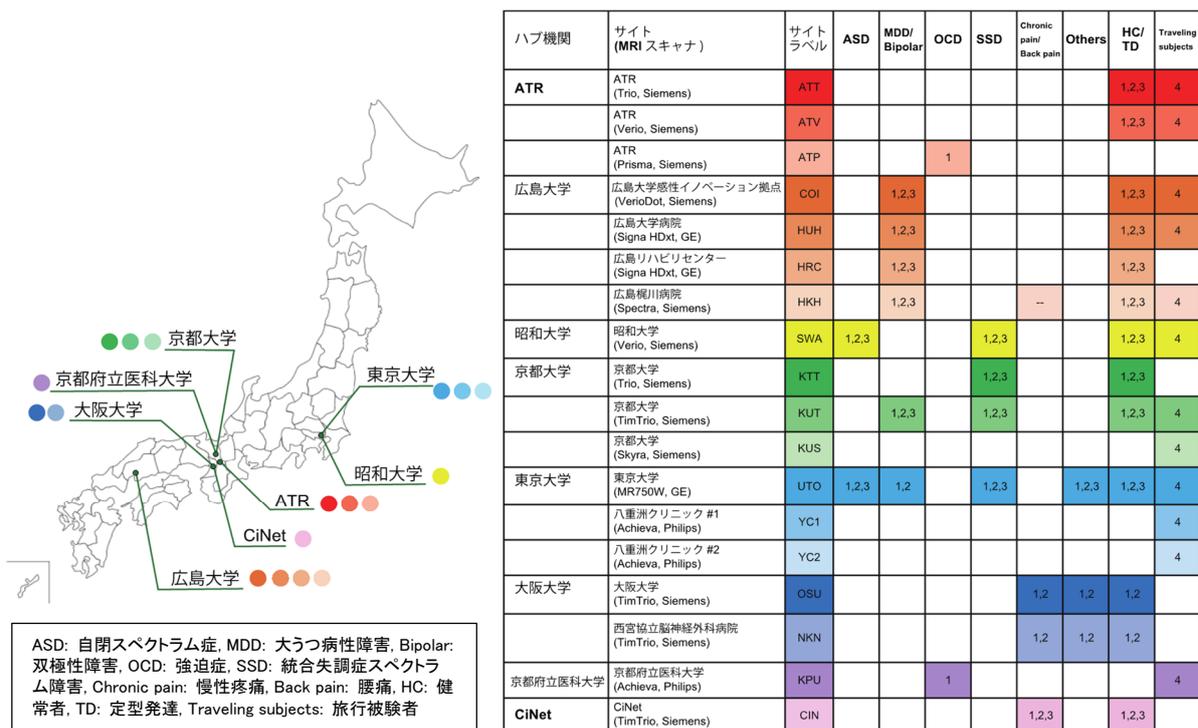


図1 データベースの概要

同意のレベルや利用者の目的に合わせて利用できる4つのデータセットを整備しました。

1. 安静時機能結合データ(制限公開)
2. 安静時 fMRI データおよび構造 MRI データ(制限公開)
3. 安静時 fMRI データおよび構造 MRI データ(非制限公開)
4. 旅行被験者の安静時 fMRI データおよび構造 MRI データ(非制限公開)

1. 安静時機能結合データ(制限公開)

安静時 fMRI データについて、前処理を行ったのち、脳を140の領域に分割し、各領域のMRI信号(BOLD信号)の時系列データに対して、その他の領域のBOLD信号との相関を計算します。これにより合計9730個の相関係数が算出されます。この相関係数(領域間機能結合^[3])を被験者ごとに計算した行列を機能結合行列と呼びます。施設ごとにこの機能結合行列を計算し、合計1625例分のデータを公開しています。また機能結合行列は計測時の体動などのノイズの影響を受けることが知られています。そこで、体動の影響を反映する指標(FD値)を個人ごとに算出し、体動の影響が大きい撮像は削除するとともに、デモグラフィックデータにFD値も合わせて公開しています。

2. 安静時 fMRI データおよび構造 MRI データ(制限公開)

MRIデータのノイズの除去手法や脳の分割方法などは現在も研究が盛んに行われており、安静時機能結合データも算出方法が進化しています。このような解析手法が発展する状況では、前処理を行っていない脳画像データを公開することがとても重要になります。そこで、安静時 fMRI データと構

造 MRI データを公開しました。MRI データの公開にあたり、すべてのデータから個人情報と連結する符号を削除し、脳画像からは顔部分を削除することで、参加者の個人同定が行われないような配慮を行い、また被験者の同意のレベルによってデータセットを分けています。1627 例分の制限公開データは、使用希望者の情報(所属、名前、連絡先)と使用方法について同意を得ました。独立して研究を行う研究者のみが書面にて使用申請を行い、ATR 脳情報通信総合研究所にて申請書の内容を査読し承認するプロセスを経て、使用が可能になります。

3. 安静時 fMRI データおよび構造 MRI データ(非制限公開)

2 のデータセットのうち、さらに非制限公開の同意を得られた 1410 名について、非制限公開データセットとして公開しています。非制限公開データは、使用方法について同意し、使用希望者の情報(所属、名前)を登録した上で、使用が可能になります。

4. 旅行被験者の安静時 fMRI データおよび構造 MRI データ(非制限公開)

同じ MRI データであっても、計測した施設によってデータの性質が全く異なってしまうという施設間差の問題があります。施設間差を除去して均質なデータベースとして統合する方法がこれまでも幾つか提案されていますが、その中でも最も効果的な方法として、「旅行被験者」と呼ばれる、同一の複数人の被験者が実際に多施設を訪れて撮像を行うことで「測定方法の違い」を同定する方法があります。このデータを使うことで、被験者の違いをなくし、純粋に測定方法の違いによる施設間差を知ることができます。9 人の被験者が国内 12 施設を訪れて、安静時 fMRI データおよび構造 MRI データを合計 143 撮像取得しました。ATR では、このデータセットを用いて、「測定方法の違いによる施設間差」のみを除去するハーモナイゼーション法を提唱し、施設間差を 3 割程度削減するとともに (Yamashita et al., 2019)、どこの施設でも約 70%の確率で健常者と大うつ病患者を判別する脳回路マーカーの開発に成功しました (Yamashita et al., 2020)。このデータセットは、使用方法について同意し、使用希望者の情報(所属、名前)を登録した上で、使用が可能になります。

今後の展望

統一プロトコルで撮像された多施設・複数疾患データと、旅行被験者データを合わせて公開することで、「測定方法の違いによる施設間差」のみを除去するハーモナイゼーションが可能となり、多施設から集めた脳画像データに対して人工知能技術である機械学習法を適用することが可能となります。このようなデータベースを公開したことで、世界中の全ての研究者がデータにアクセス出来るようになり、さらに精度の高い診断マーカーの開発や、より進化したハーモナイゼーション法の開発など、研究速度が画期的に速くなる可能性があります。また、今後は公開データを広く周知し世界中に利用してもらうことで、オープンサイエンスの推進に貢献していきます。

論文著者名とタイトル

Scientific Data 誌

(令和 3 年 8 月 30 日、10:00am (英国時間) オンライン版公開)

Tanaka SC, Yamashita A, Yahata N, Itahashi T, Lisi G, Yamada T, Ichikawa N, Takamura M, Yoshihara Y, Kunimatsu A, Okada N, Hashimoto R, Okada G, Sakai Y, Morimoto J, Narumoto J, Shimada Y, Mano H, Yoshida W, Seymour B, Shimizu T, Hosomi K, Saitoh Y, Kasai K, Kato N, Takahashi H, Okamoto Y, Yamashita O, Kawato M, Imamizu H.

A multi-site, multi-disorder resting-state magnetic resonance image database.

Scientific Data. (2021).DOI: 10.1038/s41597-021-01004-8

研究グループ* ()内は兼務先もしくは現所属を表記。

国際電気通信基礎技術研究所(ATR)

田中沙織, 山下歩(ボストン大学), 八幡憲明(東京大学・量子科学技術研究開発機構), ジュゼッペ・リジ, 山田貴志(ブラウン大学), 酒井雄希, 森本淳(京都大学), 島田育廣 (CiNet), 山下宙人(理化学研究所革新知能統合研究センター), 今水寛(東京大学大学院人文社会系研究科), 川人光男(理化学研究所革新知能統合研究センター)

東京大学

國松聡, 岡田直大, 笠井清登

昭和大学

板橋貴史, 橋本龍一郎, 加藤進昌

京都大学

吉原雄二郎, 高橋英彦(東京医科歯科大学)

広島大学

市川奈穂, 高村真広, 岡田剛, 岡本泰昌

大阪大学

清水豪士, 細見晃一, 齋藤洋一

京都府立医科大学

成本迅

情報通信研究機構 脳情報通信融合研究センター (CiNet)

眞野博彰, 吉田和子(オックスフォード大学), ベン・セイモア(オックスフォード大学)

研究支援

本研究は、AMED 脳とこころの研究推進プログラム・戦略的国際脳科学研究推進プログラム『人生ステージに沿った健常および精神・神経疾患の統合 MRI データベースの構築にもとづく国際脳科学連携』課題 JP20dm0307004(代表 笠井清登)および、AMED 脳科学研究戦略推進プログラム『DecNef を応用した精神疾患の診断・治療システムの開発と臨床応用拠点の構築』課題 JP17dm0107044 (代表 川人光男)の研究として行われたものです。また、AMED 脳とこころの研究推進プログラム・戦略的国際脳科学研究推進プログラム『脳科学と AI 技術に基づく精神神経疾患の診断と治療技術開発とその応用』課題 JP20dm0307008 (代表 川人光男)、『縦断的 MRI データに基づく成人期気分障害と関連疾患の神経回路の解明』課題 JP20dm0307002(代表 岡本泰昌)、『非線形動力学に基づく次世代 AI と基盤技術に関する研究開発』課題 JP20dm0307009 (代

表 合原一幸)の支援も受けています。また、笠井は東京大学国際高等研究所ニューロインテリジェンス国際研究機構からの支援も受けています。fMRI 計測については ATR 脳活動イメージングセンタの協力を得ました。

データ公開情報

- ATR 脳情報通信総合研究所のページから

<https://bicr.atr.jp/decnefpro/data/>

- Synapse から

<https://www.synapse.org/>

Synapse ID: syn22317076

* Synapse はアメリカの Sage Bionetworks が運営する科学データリポジトリです。

* Synapse からのアクセスは、Synapse へのアカウント登録が必要です。

お問い合わせ先

<広報担当連絡先>

株式会社国際電気通信基礎技術研究所(ATR) 経営統括部 企画・広報チーム

〒619-0288 京都府相楽郡精華町光台 2-2-2

Tel: 0774-95-1176, Fax: 0774-95-1178, E-mail: pr@atr.jp

<https://www.atr.jp/>

国立大学法人東京大学医学部附属病院 パブリック・リレーションセンター(担当:渡部、小岩井)

TEL:03-5800-9188(直通) E-mail:pr@adm.h.u-tokyo.ac.jp

国立大学法人東京大学国際高等研究所 ニューロインテリジェンス国際研究機構(広報担当)

E-mail:press@ircn.jp

<AMED の事業に関すること>

国立研究開発法人日本医療研究開発機構 疾患基礎研究事業部 疾患基礎研究課

戦略的国際脳科学研究推進プログラム

〒100-0004 東京都千代田区大手町 1-7-1 読売新聞ビル 22F

Tel: 03-6870-2286, Fax: 03-6870-2243, E-mail: brain-i@amed.go.jp

用語説明

[1] 機能的磁気共鳴画像 (functional Magnetic Resonance Imaging; fMRI)

酸化型と還元型ヘモグロビンの磁化率の違いを利用して、粗く言えば、脳全体の血流量の変化を画像化する技術です。酸化型と還元型ヘモグロビンの量の違いは脳活動の度合いを反映しているため、この画像を解析することで、各脳部位の活動度合いを推定することができます。

[2] 安静時 fMRI (resting-state fMRI)

ヒトの脳は巨大な情報ネットワークと見なすことができます。このネットワークは、遺伝で大まかな構造が決まり、様々な経験をすることで、そのヒト固有のネットワークが形作られます。最近では、わずか5～10分間安静にしているときの脳活動(安静時脳活動)を計測するだけで、そのヒトの脳内の領域同士がどのように繋がっているかを解読できるようになりました。これは、個人に特有な脳の配線図とも言えます。機械学習法による人工知能がこの脳の配線図を読み取ることで、年齢・個性・認知能力などを予測することも可能になってきました。

[3] 領域間機能結合 (functional connectivity)

空間的に隔たっている脳領域同士の活動パターンの同期関係(類似度)を表すもの。脳活動を反映するMRI信号(BOLD信号)の時間的変動の相関係数から評価を行いました。相関係数は、2領域間の脳活動の類似性が高い(=同時に活動が高くなったり低くなったりする)と1に近い値に、互いを抑制しあう関係では(一方の活動性が高いとき、他方の活動性が低いなど)-1に近い値に、互いに関連しないとき0に近い値を取ります。本研究では、140個の各脳部位から信号波形を取り出し、全ての脳部位ペア(9730個 $=140 \times 139 \div 2$)について相関係数を求めることで、個人の全脳にわたる機能的結合情報を含んだ脳の機能的回路図を得ました。