

**報道解禁日時（日本時間）**

テレビ、ラジオ、インターネット：2019年4月19日（金）午前3時

新聞：2019年4月19日（金）付け 朝刊

報道会見日時と場所：2019年4月12日（金）午後2時（場所：文部科学記者会）

※ 広島大学からは、大学院医系科学研究科精神神経医科学の岡本 泰昌教授が出席予定。

報道資料

2019年4月12日

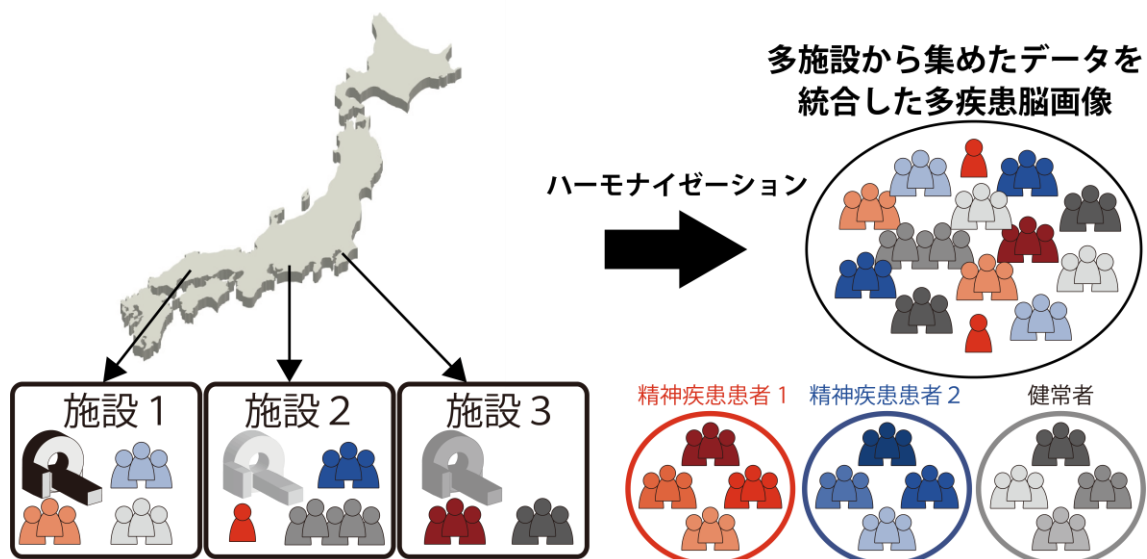
株式会社国際電気通信基礎技術研究所(ATR)  
 広島大学大学院医系科学研究科精神神経医科学  
 国立研究開発法人日本医療研究開発機構(AMED)

**脳画像の施設間差を減らすハーモナイゼーション法の開発に成功し、  
 多くの施設で集めた複数の疾患の脳画像をデータベース化**

4月18日 2pm（米国東部時間）・PLoS Biology 誌に掲載予定

**本研究成果のポイント**

- ✓ 異なった施設で取得した脳画像データを調和させ、均質なデータとする方法（＝ハーモナイゼーション法）を開発し、多施設から集めた fMRI データを1つのデータベースとして統合することに成功しました。
- ✓ 本研究では、旅行被験者データと呼ばれる、同一の複数人の被験者が実際に多施設に訪れて撮像を行うことで、脳画像における施設の違い（施設間差）のうち測定方法の違いを正確に調査することが出来るデータセットを取得しました。
- ✓ このデータセットを用いることで、複数の施設から集められた脳画像データから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開発し、施設間差を3割程度削減することに成功しました。
- ✓ この開発に用いられたデータを含む、多施設・多疾患の数千人規模の脳画像データをデータベースとして整備し、参加者から同意を得ているデータについてデータ使用登録者に対して公開しました(<https://bicyr-resource.atr.jp/decnefpro/>)。
- ✓ 本研究成果によって、施設によらずに使える精神疾患の脳回路マーカーなどを、世界に先駆けて開発して、精神疾患と発達障害の診断補助及び治療補助に貢献していきます。



## 概要

山下歩らATR脳情報通信総合研究所、広島大、東大、昭和大、京大、京都府立医科大学の研究グループは、機能的磁気共鳴画像(fMRI)データについて、異なった施設で取得した脳画像データを調和させ、均質なデータとする方法(=ハーモナイゼーション法)の開発に成功しました。また ATR・田中沙織室長、広島大・岡本泰昌教授らは、この開発に用いられたデータを含む、多施設で多疾患の数千人規模の脳画像データをデータベースとして整備し、その一部を登録者に対して公開しました。

近年、生物医学、心理学などの分野において、論文で発表された結果が再現出来ないことが指摘されています[1]。特に、1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である機械学習法を適用して得られた結果は、他の施設では再現できません[2]。この問題を解決するには多施設から集めた多数の脳画像データが必要ですが、計測した施設によってデータの性質が異なるという、極めて困難な問題がありました。米国 Human Connectome Project では一つの施設から多数の脳画像データを集めて、施設間差の問題を避けていましたが、世界中のどの施設でも再現可能な結果を得るためには、施設間差を解消する必要があります。

山下らは、fMRIに有効なハーモナイゼーション法を開発し、施設間差を3割程度減らしました。まず、9人の被験者が12施設を訪れて撮像し、全部で411サンプル取得しました。この旅行被験者データと、複数の施設から集められたデータ(9施設4疾患の被験者805人から1サンプルずつ805サンプル取得)の両者を組み合わせたデータセットに、施設の違い、疾患による変容などを推定する数学的なモデルを当てはめて、測定方法の違いによる脳画像の違いのみを除去するハーモナイゼーション法を開発しました。

本研究は、生命科学の分野で権威のあるPLOS Biology誌で、論文で発表された結果の再現性に関する問題などで注目を集めているmeta-researchというセクションの論文として掲載されます。

また、本研究で使用した多施設・多疾患で収集された大規模データ(総数2,409例)を世界的にも貴重なデータベースとして構築しました。このうち、研究参加者から同意を得ている1,828例のデータについて、参加者の個人同定が行われないうに十分に配慮した上で、登録した研究者に対して所定の審査を行ったのちに利用可能となる形で公開を始めました。

<https://bicr-resource.atr.jp/decnefpro/>

今後は、上記のハーモナイゼーション法を公開した脳画像データなどに適用して、施設によらずに使える精神疾患の脳回路マーカーなどを、世界に先駆けて開発して、精神疾患と発達障害の診断補助及び治療補助に貢献していきます。

## 背景

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、さまざま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来ました[3]。

これまでの多くの研究では、1施設で数十人くらいのfMRIデータから計算された脳の配線図と機械学習法を組み合わせることで、様々な認知機能の予測を行ってきました。しかしな

が、これらの研究結果を他施設で得られた fMRI データに当てはめてもほとんど再現出来ないことがわかってきました。この原因は、少数のデータサンプルで、多くの要因が関与するデータ(脳画像)の背後にある法則を学習しようとする、機械(コンピュータ)はそのデータサンプルだけにしか通用しない特殊なことを学習してしまう「過学習」という現象が起きるからと考えられています[2]。過学習を避けるには、多くの fMRI データが必要になりますが、ひとつの施設で集められる被験者には限りがあります。そこで、複数の施設が協力してデータを集めることが必要になります。しかし、同じ fMRI 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した。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きましたが、fMRI データの施設間差には 2 つの異なる施設間差が含まれていることを考慮していませんでした。それは、fMRI 装置や測定方法の違いによる生物学的に意味のない「測定方法の違い」から生じる施設間差と、生物学的に意味のある「被験者の違い」から生じる施設間差です。つまり、既存の手法では「測定方法の違い」のみならず、生物学的に意味のある「被験者の違い」までも同時に除去してしまう可能性が存在しました。そこで、本研究では旅行被験者データと呼ばれる、同一の複数人の被験者が実際に多施設に訪れて撮像を行うことで「測定方法の違い」を調査することが出来る特殊なデータセットを取得しました。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。複数の施設から集められたデータから、旅行被験者データで解った「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開発し、施設間差を 3 割程度削減することに成功しました。

## 研究内容

9 人の研究参加者が国内 12 施設を訪れて安静状態における脳活動を、各施設の fMRI 装置を用いて約 10 分間計測し、全部で 411 サンプル取得しました(旅行被験者データ)。この旅行被験者データに加えて、国内 9 施設 4 疾患(自閉症、うつ病、統合失調症、強迫性障害)の研究参加者 805 人の安静状態における脳活動を、各施設の fMRI 装置を用いて約 5 分~10 分間計測し、全部で 805 サンプルを取得しました(多施設多疾患データ)(図1)。脳を 268 個の小領域に分割し、ひとりひとり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ました。これを領域間機能的結合と呼びます。268 個の小領域の全てのペア(35,778 個)について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、全脳の回路図(= 35,778 個の数値からなるベクトル)が作成されます[4]。これを参加者全員分について求め、旅行被験者データと多施設多疾患データを組み合わせたデータセットを作りました。データセットを用いて、測定方法の違いによる施設間差、被験者群の違いによる施設間差が機能的結合に与える影響、さまざまな疾患による機能的結合の変容などを数理モデルに当てはめて推定しました。そして、測定方法の違いによる施設間差、被験者の違いによる施設間差が機能的結合に与える影響が、疾患による機能的結合の変容と比べて、同等もしくは大きいことを見出しました(図 2a)。さらに、我々は様々な測定方法の違いが機能的結合に与える影響の比較を行い、位相方向[5]、fMRI の製造会社、コイルのタイプ、装置の機種順で影響が大きいことを明らかにしました(図 2b)。また、各施設において同一の集団から被験者がサンプルされたと考えるのには無理があり(図 3a)、施設ごとに異なる部分集団からサンプルされることがわかりました(図 3b)。このことは、特定の施設のみで大規模なデータセットを集めたとしても、世界中のどの施設でも再現可能な結果を得ることが出来ないことを示しています。最後に、これらの施設間差の問題を克服するために、旅行被験者データを用いて、多

施設・多疾患データから測定方法の違いのみを除去するハーモナイゼーション法を開発し、施設間差を3割程度削減することに成功しました。

今回のハーモナイゼーション法の開発には、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・神経疾患の克服に向けた研究開発」の「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」課題において策定した「脳プロ統一プロトコル (<https://bicr.atr.jp/rs-fmri-protocol-2/>)」を用いて、同課題に参画した複数の施設(図4:ATR、広島大学、梶川病院、京都府立医科大学、京都大学、東京大学、昭和大学、八重洲クリニック、大阪大学、CiNet)にて撮像したデータを使用しました。これらの多施設・多疾患で収集された大規模データ(総数 2,409 例)を世界的にも貴重なデータベースとして整備しました(表1)。このデータベースは、疾患患者 988 名、健常者 1421 名の脳構造画像および脳、年齢・性別、患者については該当する疾患の診断に用いられる臨床評価尺度から構成されています。疾患の内訳は自閉症スペクトラムが 125 名、うつが 445 名、統合失調症が 159 名、強迫性障害が 110 名、慢性疼痛・腰痛が 107 名、その他の疾患が 42 名になります。2,409 例の参加者全員から多施設・多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。またこのうち疾患患者 706 名、健常者 1122 名に関しては公開の同意を得ております。このデータについて、利用を希望する研究者に対して所定の審査を行ったのちに利用可能となる形で公開を始めました(<https://bicr-resource.atr.jp/decnefpro/>)。使用希望者は、使用申請書をサイトからダウンロードし、必要事項を記載した上、サイト上に記載されているメールアドレス(decnef-db-admin@atr.jp)に送付することで、申請承認後に登録アカウントが発行されます。そのアカウントでログインすると、サイト上のデータが使用可能になります。現在は、8 つの機関(12スキャナ)について、安静状態の脳機能結合データと年齢・性別・臨床評価尺度等の情報が使用可能です。また、脳画像データの公開に同意を得られた参加者のデータについてのみ、2019 年 9 月以降に脳画像データの公開を予定しています。なお、全てのデータは個人を同定できない状態で公開します。具体的には、すべてのデータから個人情報と連結する符号を削除し、脳画像データ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なっています。

## 本研究の意義と今後の展望

本研究は、初めて、測定方法の違いによる施設間差、被験者の違いによる施設間差が機能的結合データに与える影響は、疾患による機能結合の変容と比べて、同等もしくは大きいことを見出しました。このことは、fMRI データを用いて疾患による変容を調査する際には施設間差に十分注意する必要があることを示しており、とても重要な知見となります。

また、各施設において同一の集団から被験者がサンプルされるのではなく、施設ごとに異なる部分集団からサンプルされることがわかりました。このことは、特定の施設のみで大規模なデータセットを集めても世界中のどの施設でも再現可能な結果を得ることが出来ないことを示しており、今後大規模データセットを集める際の重要な指針となります。

本研究では fMRI データにおける施設間差の問題を克服するために、旅行被験者データを用いて、多施設・多疾患データから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開発しました。このハーモナイゼーション法を用いることにより、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。さらに、多施設・多疾患で収集された大規模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、研究速度が画期的に速くなる可能

性があります。

今後、我々の研究グループにおいては開発したハーモナイゼーション法を公開した脳画像データなどに適用して、施設によらずに使える精神疾患の脳回路マーカなどを、世界に先駆けて開発して、精神疾患と発達障害の診断補助及び治療補助に貢献していく予定です。

### 【論文情報】

*PLOS Biology* 誌 (米国東部時間・2019年4月18日 2pm 公開)

Ayumu Yamashita, Noriaki Yahata, Takashi Itahashi, Giuseppe Lisi, Takashi Yamada, Naho Ichikawa, Masahiro Takamura, Yujiro Yoshihara, Akira Kunimatsu, Naohiro Okada, Hirotaka Yamagata, Koji Matsuo, Ryuichiro Hashimoto, Go Okada, Yuki Sakai, Jun Morimoto, Jin Narumoto, Yasuhiro Shimada, Kiyoto Kasai, Nobumasa Kato, Hidehiko Takahashi, Yasumasa Okamoto, Saori C Tanaka, Mitsuo Kawato, Okito Yamashita, and Hiroshi Imamizu

Harmonization of resting-state functional MRI data across multiple imaging sites via the separation of site differences into sampling bias and measurement bias. *PLOS Biology*. DOI: 10.1371/journal.pbio.3000042

<http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3000042>

### 【研究グループ】 \* ( )内は兼務先もしくは現所属を表記。

株式会社国際電気通信基礎技術研究所(ATR)

山下歩, 八幡憲明(東京大学・放医研), リシ・ジュゼッペ, 山田貴志(昭和大学), 酒井雄希(京都府立医科大学), 森本淳, 田中沙織, 川人光男(理化学研究所革新知能統合研究センター), 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)

東京大学大学院医学系研究科

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

広島大学

市川奈穂, 高村真広, 岡田剛, 岡本泰昌

ATR-Promotions 脳活動イメージングセンタ

島田育廣

山口大学

山形弘隆, 松尾幸治(現:埼玉医科大学)

京都府立医科大学

成木迅

## 【研究支援】

本研究は、国立研究開発法人日本医療研究開発機構 (AMED) ・「戦略的国際脳科学研究推進プログラム」の『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 JP18dm0307008 (代表 川人光男)、脳科学研究戦略推進プログラム BMI 課題の『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)、および「戦略的国際脳科学研究推進プログラム」の『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 JP18dm0307002 (代表 岡本泰昌)、『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 JP18dm0307004 (代表 笠井清登)、『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 JP18dm0307009 (代表 合原一幸)の研究として行われたものです。

一部は、内閣府 総合科学技術・イノベーション会議が主導する革新的研究開発推進プログラム (ImPACT) 「脳情報の可視化と制御による活力溢れる生活の実現」、日本学術振興会 科研費 26120002 と 15J06788 の助成を受けています。

## 【お問い合わせ先】

<研究内容に関すること>

(株)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/index\\_j.html](https://www.atr.jp/index_j.html)

<AMED の事業に関すること>

日本医療研究開発機構 戦略推進部 脳と心の研究課

〒100-0004 東京都千代田区大手町 1-7-1 読売新聞ビル 22F

Tel: 03-6870-2222, Fax: 03-6870-2244

E-mail: brain@amed.go.jp

## 補足説明

[1] 近年、生物医学、心理学、社会科学の多くの研究結果が再現出来ない問題が報告されています。

<https://www.nature.com/ndigest/journal/v10/n11/pdf/ndigest.2013.131128.pdf>

<http://science.sciencemag.org/content/349/6251/aac4716>

<https://www.nature.com/articles/s41562-018-0399-z>

[2] fMRI データから分類を行う場合には、基本的に機械学習が用いられます。機械学習では過学習の問題を避けるために、一人の被験者を除いて検証用として用いる交差検証法: leave-one-subject-out cross validation やデータを 10 分割し、10 分の 9 で学習し、残りの 10 分の 1 で検証を行う 10-fold cross validation を用いて分類器の評価を行うことが多い。しかし、単一施設から得られた少数のサンプルに対して機械学習を適用すると、予測のインフレーションを起こす危険が、精神医学分野でも近年認識されるようになってきました。少数のデータに対する機械学習では、学習用データにおける特定の施設の fMRI 装置や測定方法、実験者、参加者群などに存在する特定の傾向、あるいはノイズに対して過学習してしまう可能性が高い。例えば、脳の解剖画像から自閉スペクトラム症を判別する分類器は、開発に使われた英国の学習用データには感度も特異度も 9 割以上の高性能を示すが、日本人のデータでは 5 割になってしまうことが報告されています。学習用データとは全く異なる施設と被験者群からなる独立検証コホートで検証していない分類器は、科学的にも実用的にも意味が殆どないと言って過言ではありません。本研究では、多施設多疾患データには含まれていない山口大学などの完全に独立施設で撮像されたデータを用いて汎化検証を行いました。

[https://www.biologicalpsychiatryjournal.com/article/S0006-3223\(13\)00457-5/fulltext](https://www.biologicalpsychiatryjournal.com/article/S0006-3223(13)00457-5/fulltext)

Yoshihara Y, Sugihara G, Deoni D, et al. Discrimination of autistic adults from Controls using data on Whole-Brain MRI in a Japanese Sample. in The 17th Annual Meeting of the Organization on Human Brain Mapping 26-30 June, 2011.

<http://www.jneurosci.org/content/30/32/10612.long>

<https://www.sciencedirect.com/science/article/pii/S1053811909009045?via%3Dihub>

[3] 例えば、脳の配線図から個人の作業記憶の上限を予測することが出来ます。

<http://www.natureasia.com/ja-jp/srep/abstracts/61778>

[4] 空間的に隔たっている脳領域どうし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、268 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(35,778 個=268×267÷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。

[5] fMRI の画像は一般的に 2 次元の画像を取り重ねることで 3 次元の画像を構築します(金太郎あめのイメージです)。この 2 次元の画像を取り重ねる方向の違いが位相方向の違いであり、位相方向の違いによって fMRI 内の磁場の歪みが異なることが知られています。

精神疾患患者群 1

精神疾患患者群 2

健常者群



a 多施設多疾患データセット



b 旅行被験者データセット



\* 全ての施設で同一参加者



図 1 : 本研究で使用了データセットの概要と施設間差について

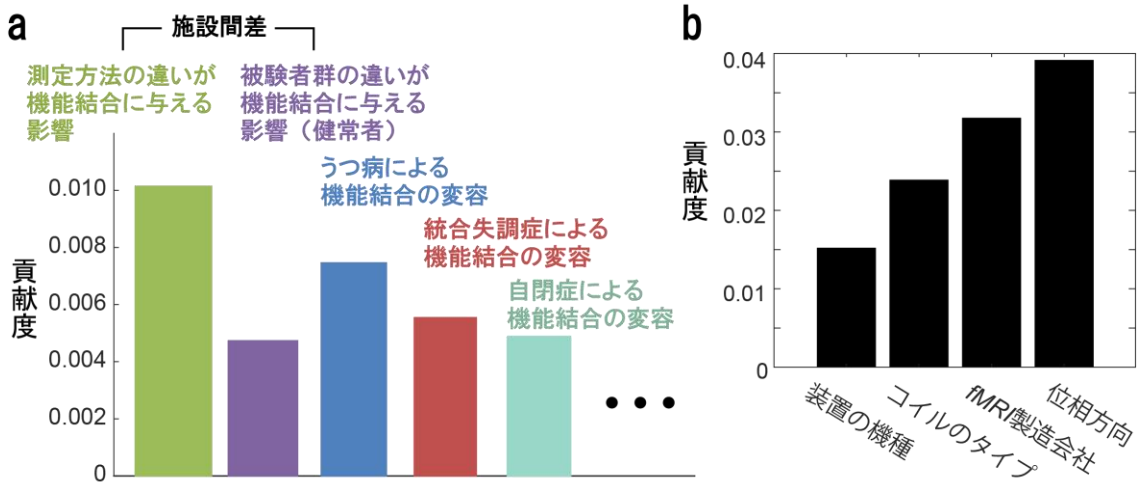


図 2 : (a)さまざまな要因が機能的結合に与える影響の比較 (b)さまざまな測定方法が機能的結合に与える影響の比較



**a 単一母集団モデル**    **b 施設毎に異なる集団モデル**  
 健常群の母集団                      健常群の母集団

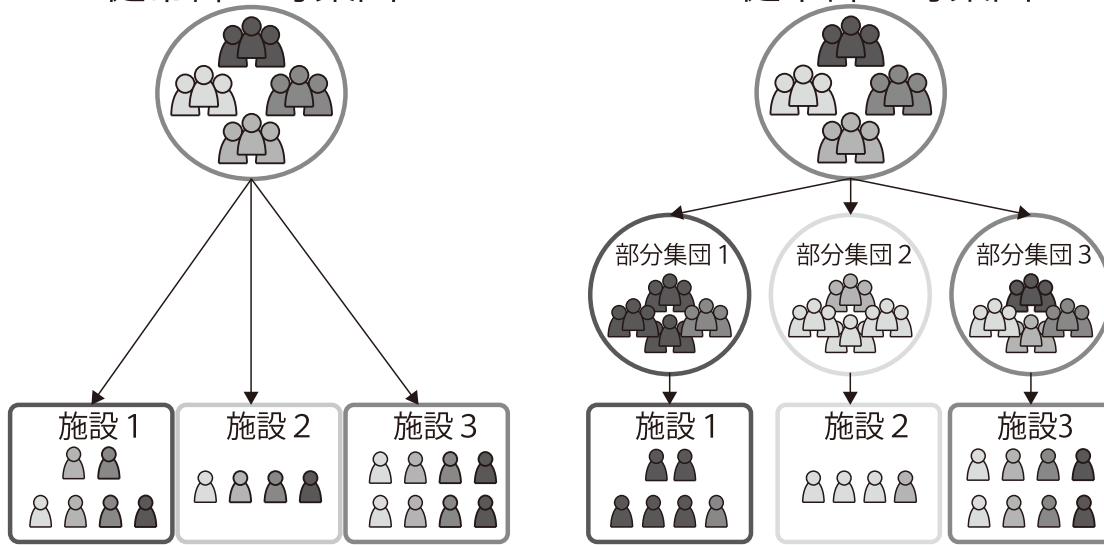


図 3: 参加者のサンプルされ方のモデル。(a)同じ母集団から各施設の参加者がサンプルされる。(b)施設毎に異なる部分集団から参加者がサンプルされる。



図 4: 撮像施設

サイト	スキャナ	自閉症 スペクトラム	うつ	統合失調症	強迫性障害	慢性疼痛 腰痛	その他	健常者	合計
東大	Philips, GE	10	103	36			32	170	351
阪大	Siemens					43	10	29	82
昭和	Siemens	115		19				110	244
京大	Siemens		20	104				244	368
広大	Siemens, GE		318			40		468	826
ATR	Siemens		4		5			271	280
京都府立医科	Philips				105			90	195
CiNet	Siemens					24		39	63
合計		125	445	159	110	107	42	1421	2409

表 1: 多施設・多疾患データベースの詳細。このデータベースのうち、本研究では東大からうつ病患者 34 名, 統合失調症患者 14 名, 健常者 142 名の計 190 名, 昭和から自閉症患者 49 名, 統合失調症患者 12 名, 健常者 40 名の計 101 名, 京大からうつ病患者 9 名, 統合失調症患者 22 名, 健常者 35 名の計 66 名, 広大からうつ病患者 118 名, 健常者 105 名の計 223 名, ATR から健常者 108 名, 京都府立医科から強迫性障害 65 名, 健常者 52 名の計 117 名の総勢 805 例のデータを使用しています。

	Sub.1	Sub.1	Sub.3	Sub.4	Sub.5	Sub.6	Sub.7	Sub.8	Sub.9
Age	25	27	26	26	32	24	25	28	30
Sex	M	M	M	M	M	M	M	M	M

	MRI1	MRI2	MRI3	MRI4	MRI5	MRI6	MRI7	MRI8	MRI9	MRI10	MRI11	MRI12
Site location	ATR	ATR	広大 COI	広大	梶川病院	京府医大	京大	京大	東大	昭和大	八重洲 クリニック	八重洲 クリニック
Scan make	SIE	SIE	SIE	GE	SIE	Phi	SIE	SIE	GE	SIE	Phi	Phi
Scan model	Trio	Verio	Verio	Signa HDxt	Spectra	Achieva	Skyra	Tim Trio	MR750W	Verio	Achieva	Achieva

	MRI1	MRI2	MRI3	MRI4	MRI5	MRI6	MRI7	MRI8	MRI9	MRI10	MRI11	MRI12	Sum
Sub.1	15	3	3	2	2	3	3	3	3	3	3	3	46
Sub.2	15	3	3	2	2	3	3	3	3	3	3	3	46
Sub.3	15	3	3	2	2	3	3	3	3	3	3	3	46
Sub.4	15	3	3	2	2	3	3	3	3	3	3	3	46
Sub.5	15	3	3	2	2	3	3	3	3	3	3	3	46
Sub.6	12	3	3	2	2	3	3	3	3	3	3	3	43
Sub.7	15	3	3	2	2	3	3	3	3	3	3	3	46
Sub.8	15	3	3	2	2	3	3	3	3	3	3	3	46
Sub.9	15	3	3	2	2	3	3	3	3	3	3	3	46
Sum	132	27	27	18	18	27	27	27	27	27	27	27	411

表 2: 旅行被験者の詳細。被験者情報と撮像した施設、スキャナおよび撮像数。本研究では 9 施設 12 個の fMRI 装置から得られた 411 サンプル全てを使用しています。