

アルツハイマー病の予兆候補の発見に役立つ機械学習モデル開発

～現実的な実験データの制約下で適用可能なモデル～

論文掲載

【本研究成果のポイント】

- アルツハイマー病患者の脳では、アミロイドβと呼ばれるタンパク質の蓄積が生じることが知られており、そのアミロイドβの脳内蓄積量を予測する特徴の発見のための機械学習モデルを開発しました。
- 機械学習による予測に通常は必要となるデータ（予測対象と入力のペアデータ）が十分でない場合にも使用が可能です。
- 公開マウス実験データで、行動特徴から海馬（記憶などを司る脳の部位）のアミロイドβ量を予測できることを実証しました。
- 今後、アルツハイマー病の早期発症予測法開発に貢献することが期待されます。

【概要】

アルツハイマー病は、脳の神経細胞が徐々に変性する進行性の疾患です。主に高齢者に見られ、記憶力の低下や認知機能の障害が特徴です。原因は完全には解明されていませんが、アルツハイマー病患者の脳では、神経細胞の変性に先立ってアミロイドβと呼ばれるタンパク質の蓄積が生じることが知られています。脳内のアミロイドβの蓄積を判定する方法は、現状では高額なコストや侵襲性などの問題を抱えています。そのため、脳内のアミロイドβの蓄積量を予測できる、簡便で非侵襲に計測可能なバイオマーカー（血液、尿、医療画像などから測定可能な、病気の状態を表す指標）があれば、アルツハイマー病の早期発症予測に有用なはずで

す。通常、機械学習を用いてアミロイドβの蓄積量をバイオマーカーから予測しようとすると、ペアデータ（バイオマーカーとアミロイドβ蓄積量を同じサンプルで観測したデータ）が必要になります。しかし、このようなペアデータの取得は高コストで労力もかかるため、バイオマーカー探索では避けられてきました。そこで、広島大学大学院統合生命科学研究科データ駆動生物学研究室の矢田祐一郎特任助教、本田直樹教授（兼任：京都大学生命科学研究科特命教授、自然科学研究機構生命創成探究センター客員教授）からなる研究グループは、ペアデータが限られている場合でも、アミロイドβ蓄積量の定量的予測を可能にする機械学習モデルを開発しました。今後、この技術を応用することで、アミロイドβ蓄積量の予測性に基づいた新たなアルツハイマー病バイオマーカーが開発されることが期待されます。本研究成果は、npj Systems Biology and Applications 誌に 2023 年 11 月 23 日に掲載されました。

【背景】

アルツハイマー病は、脳の神経細胞が徐々に変性する進行性の疾患です。主に高齢者に見られ、記憶力の低下や認知機能の障害が特徴で、認知症の中でも最も一般的な疾患です。高齢化が進む先進国では、患者数は増加の一步を辿っており、その対策は喫緊の課題です。アルツハイマー病患者の脳では、神経細胞が変性して細胞死が生じ、それが直接的な原因となり機能障害が起きます。アルツハイマー病患者の脳では、神経細胞の変性に10年から20年ほど先立って、アミロイド β と呼ばれるタンパク質の蓄積が生じることが知られています。脳脊髄液中のアミロイド β 量の測定や陽電子放射断層撮影(PET)によるアミロイド β イメージングは、アルツハイマー病の診断に有用です。しかし、それら脳内のアミロイド β 蓄積を間接的に判定する方法は、現状では高額なコストや侵襲性などのため、発症前の健康な方に対して適用する上での困難を抱えています。さらに、蓄積のごく初期段階での有効性は不明です。そのため、脳内のアミロイド β の蓄積量を予測できる、簡便で非侵襲に計測可能なバイオマーカーがあれば、アルツハイマー病の早期発症予測などに有用なはずで

す。バイオマーカー候補から脳内のアミロイド β 量を予測する問題を考える上では、通常はバイオマーカー候補とアミロイド β 量を同じ個体から計測した「ペアデータ」が必要になります。アルツハイマー病の新規バイオマーカーを探索する際はモデル動物を用いる事が多いですが、金銭的なコスト及び労力がかかるため、アミロイド β 量を同じサンプルで定量する例は多くありません。そのため、アミロイド β の予測性に基づくバイオマーカー探索をするためには、ペアデータがあまりない場合でも予測可能な方法が必要でした。

【研究成果の内容】

(1) アミロイド β 量予測のための階層ベイズモデルと学習アルゴリズムの開発

本研究では、アミロイド β とバイオマーカー候補のペアデータが限られている場合に対しても、バイオマーカーからアミロイド β を予測することができる機械学習モデルと学習アルゴリズムを提案しました。この機械学習モデルは、マウスなどのモデル動物を使用して、いくつかのバイオマーカー候補が計測されている状況を想定しています。

まず、提案モデルでは、アミロイド β 量が時間に対してシグモイド曲線(図1の左上図のようなS字型の曲線)に従って蓄積すると仮定します(図1)。曲線の最大値や時間的位置、急峻さを決めるパラメータは、個体ごとに異なります。次に、アミロイド β の蓄積状態を、アルツハイマー病の進行度を表す指標と見立てて、バイオマーカー候補はその指標に基づいた値が観測されると仮定します。ただし、観測値がどの程度指標の影響を受けるかは、バイオマーカー候補によって異なります。

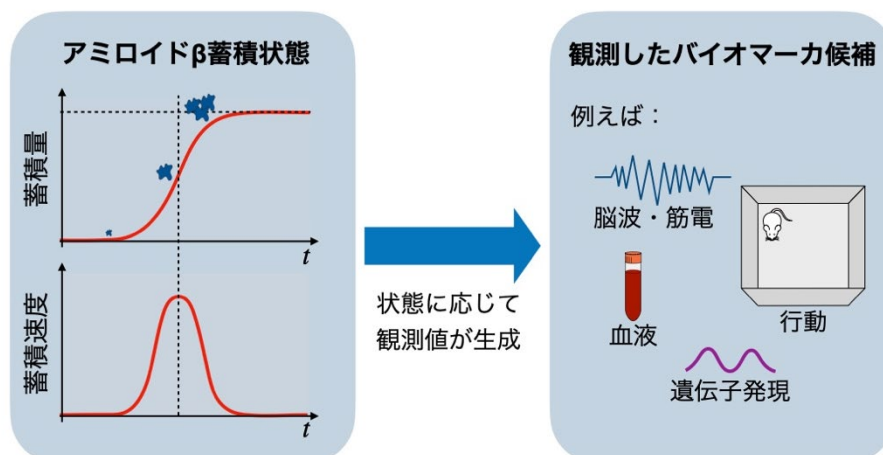


図1：提案モデルのコンセプト

提案モデルの学習では、「シグモイド曲線のパラメータ」と「各バイオマーカー候補がアミロイドβ状態からどの程度影響を受けるか」をデータから推定します。ここでは、ペアデータが多くない場合でも、ベイズ学習（事前に手に入るデータから得た知識を元にして、新たなデータを学習して知識を更新する機械学習法）と半教師あり学習正解のあるデータと無いデータの両方を使って学習する機械学習法」というテクニックを使用することで、学習を可能にしました。学習したモデルを用いると、バイオマーカー候補の観測値から、アミロイドβ量の予測値を計算することができます。

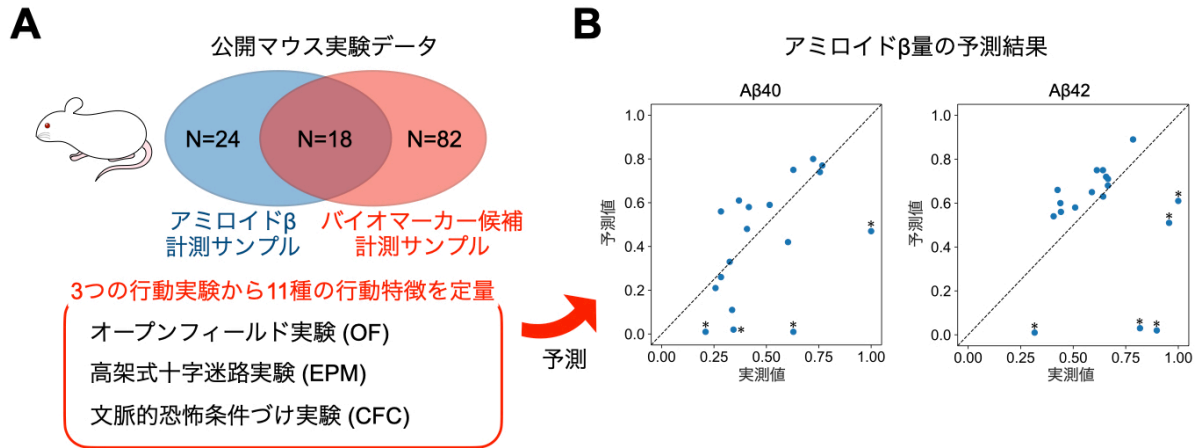


図 2：公開マウス行動実験データによるアミロイドβ量の予測

(2) モデルマウス行動データでの検証

提案したモデルと学習アルゴリズムを、公開されているマウスデータに適用し、予測性能を評価しました。このデータでは健康な野生型のマウスと、アルツハイマー病のモデルマウスである 5xFAD マウスに対して、何点かの月齢でいくつかの行動実験を実施し、複数の行動特徴を計測しており、また、アミロイドβ量を電気化学発光免疫測定法（標的タンパク質と抗体の反応を電気化学的に検出した結果からタンパク質量を高感度で定量する手法）により定量しています。ただし、大半のマウスは行動特徴とアミロイドβ量のどちらかしか計測されておらず、一部の個体のみ両者が取得されたペアデータがあります（図 2A）。

このデータを使用して提案した機械学習モデルを学習し、学習に使用しなかったマウスの全行動特徴からアミロイドβ量が予測できるかを交差検証すると、（図 2B）のように、大半のマウスでは海馬のアミロイドβ量が予測できることがわかりました（Aβ40、Aβ42 はアミロイドβの主要なサブタイプ）。ペアデータの数をさらに少なくした場合でもあまり予測性能は変わらず、少ないペアデータに対して十分に適用可能な機械学習モデルになっていることがわかりました。

(3) 予測に重要なバイオマーカー候補の同定

次に、アミロイドβ量の予測にどの行動特徴が重要だったかを評価しました。初めに、全行動特徴から特定の一つの特徴を除外した場合の予測性能を、除外する特徴を入れ替えながら順次評価しました（図 3A）。さらに、除外した場合に予測誤差が大きくなった順に行動特徴を並び替え、上から順番に使用する特徴数を増やしていった場合の予測誤差を評価しました（図 3B）。すると、全特徴を使用した場合と同程度の予測性能が出るのは 11 特徴中 10 個を使用した場合で、ある程度良い性能が出るのも 5 つの特徴を使用した場合になることがわかりました。この 5 つの特徴には、3 つの異なる実験から取得された特徴が含まれていました。これは、多様な方法で取得された複数の特徴がアミロイドβ量を予測するバイオマーカーとして有用である可能性を示唆しています。

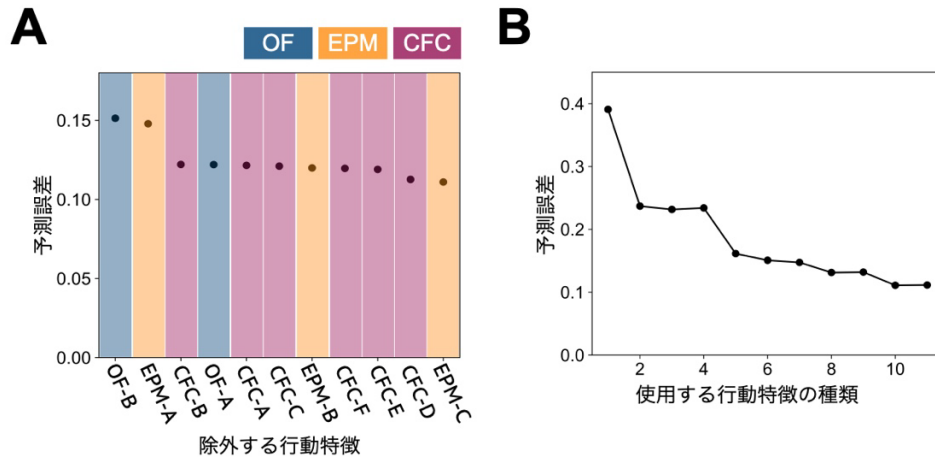


図3：アミロイドβ量の予測に重要な行動特徴の評価

【今後の展開】

本研究で開発した機械学習モデルと学習アルゴリズムを利用することで、アミロイドβの予測に基づいたアルツハイマー病バイオマーカーの開発に貢献することが期待されます。アルツハイマー病発症前から蓄積が開始するアミロイドβの蓄積量を定量的に予測できるバイオマーカーは、アルツハイマー病の早期発症予測に重要と考えられます。提案モデルを利用することで、これからアミロイドβ予測バイオマーカーの探索実験をする際に、全ての個体でペアデータを取得する必要がないため、効率のよい実験が可能です。また、これまでに取得されたモデル動物でのバイオマーカー探索実験データを、アミロイドβ予測バイオマーカー探索に有効活用することもできます。

本研究はモデル動物での実験データを対象としていますが、今後ヒトでのデータも扱える機械学習モデルへと発展させていくことが期待されます。また、パーキンソン病など他の神経変性疾患でも、異常なタンパク質の蓄積に続いて神経細胞の変性が起き、機能障害が引き起こされる、という流れは共通しています。そのため、提案モデルのアプローチは他の神経変性疾患にも応用可能と考えられます。今後、それらの疾患に対しても、予測に基づいたバイオマーカーの開発に提案モデルが貢献することが期待されます。

【掲載論文】

論文タイトル

Few-shot prediction of amyloid β accumulation from mainly unpaired data on biomarker candidates

著者

Yuichiro Yada^{1,*}, Honda Naoki^{1,2,3,*}

1：広島大学大学院統合生命科学研究科

2：京大大学生命科学研究科

3：自然科学研究機構生命創成探究センター

*：責任著者

掲載誌

npj Systems Biology and Applications

DOI 番号

<https://doi.org/10.1038/s41540-023-00321-5>

【プロジェクトについて】

本研究は、JST【ムーンショット型研究開発事業目標 2050年までに、超早期に疾患の予測・予防をすることができる社会を実現】(JPMJMS2024-9)の支援を受けたものです。