

厚生労働科学研究費補助金（障害者政策総合研究事業）

分担研究報告書

大規模疫学研究データによる入院外統合失調症等有病率の推定に関する研究

研究分担者 He Yupeng 藤田医科大学医学部公衆衛生学講座 助教

研究要旨。

統合失調症を有する成人 223 人と精神障害を有さない成人 1776 人の健康に関連する情報と身体的・精神的・社会的併存症状のデータから、機械学習を用いて、統合失調症の症例を正確に分類できるモデルを構築した。機械学習モデルの感度、特異度、陽性的中率、陰性的中率は、それぞれ 0.56、0.97、0.69、0.95 であった。機械学習モデルの陽性的中率はロジスティック回帰モデルの陽性的中率を上回った。個人レベルで統合失調症の有無を判別することに使用するには限界はあるが、集団レベルでの統合失調症の有病率を推測するためには使える可能性がある。

A. 研究目的

松永研究分担者の研究結果（参照：分担研究報告書）から、個人特性や身体的・精神的・社会的併存症状をもとに統合失調症を有する者の判別が可能であると考えた。本研究では機械学習を用いて、統合失調症の症例を正確に判別できるモデルを構築した。

B. 研究方法

2022 年 2 月に行ったインターネット調査にて、統合失調症を有する成人 223 人と精神障害を有さない成人 1776 人から個人特性と身体的・精神的・社会的併存症状を尋ねた。同調査の詳細は松永研究分担者の分担研究報告書を参照されたい。

インターネット調査の質問項目への回答は、1 つの回答変数（統合失調症と診断された）と特徴変数（個人特性、身体的・精神的・社会的併存症状）にフォーマットされた。機

械学習モデルとして、人工ニューラルネットワーク (artificial neural network: ANN) を適用して、統合失調症の症例を分類するためのモードを構築した。本研究では、5 つの隠れ層（各層のニューロン：128-64-32-16-8）、HeNormal 重み初期化器、隠れ層の ReLU 活性化関数、出力層の Sigmoid 活性化で ANN を構成した。多重ロジスティック回帰モデルを比較対照として適用した。

機械学習モデルと多重ロジスティック回帰モデルのそれぞれで、感度・特異度・陽性的中率・陰性的中率を算出し、比較した。受信者動作特性曲線下面積と精度も算出した。（倫理面への配慮）

本研究はヘルシンキ宣言および人を対象とする生命科学・医学系研究に関する倫理指針（文部科学省、厚生労働省、経済産業省）に則って実施した。藤田医科大学医学研究倫理審査委員会の審査を受け、藤田医科大

表.統合失調症例の予測値と観察値：機械学習とロジスティック回帰式の比較

		機械学習		ロジスティック回帰式	
観察値 (統合失調症)	あり	19	24	16	27
	なし	346	11	337	20
		なし	あり	なし	あり
		予測値 (統合失調症)		予測値 (統合失調症)	
		感度	0.56	感度	0.63
		特異度	0.97	特異度	0.94
		陽性的中率	0.69	陽性的中率	0.57
		陰性的中率	0.95	陰性的中率	0.95

学長の承認を得て実施した。

C. 研究結果

機械学習モデルの感度、特異度、陽性的中率、陰性的中率は、それぞれ 0.56、0.97、0.69、0.95 であった (表)。受信者動作特性曲線下面積(0.86 vs. 0.78)、精度 (0.93 vs. 0.91)、特異度(0.97 vs. 0.94)、陽性的中率(0.69 vs. 0.57)においてロジスティック回帰モデルよりも優れていた。ロジスティック回帰モデルは感度 (0.63 vs. 0.56) において機械学習モデルより優れていた。睡眠薬の使用、年齢、世帯収入、雇用形態が、変数の重要度の上位 4 位を占めた。

D. 考察

機械学習モデルの陽性的中率 0.69 はロジスティック回帰モデルの陽性的中率 0.57 を上回った。統合失調症を有する人を抽出する場合、機械学習のほうがロジスティック回帰モデルよりも優れている可能性がある。

その一方、機械学習モデルの陽性的中率

は 0.69 に留まった。個人レベルで統合失調症の有無を判別することに使用するには限界はある。しかし、この機械学習モデルを既存の一般住民を対象とした大規模疫学研究データに当てはめ、集団レベルでの統合失調症の有病率を推測するためには使える可能性がある。

本研究の限界として、対象者の統合失調症・精神障害の有無が自己申告に基づく点が挙げられる。また、本研究は統合失調症を有する人と精神障害を有さない人からなるデータを用いて行われた。統合失調症ではない精神障害を有する人が機械学習モデルで統合失調症と判定されてしまう可能性も考えられる。これらの懸念を解決するためには本研究で開発した機械学習モデルを精神科医が診断を確定した症例に当てはめ、妥当性をさらに検討する必要がある。そのために、この太田班では、精神科医が診断を確定した人 (精神科受診患者) を対象にして研究を進める予定である。詳しくは李分担研究者の分担研究報告書を参照されたい。

E. 結論

さらに妥当性を検証する必要があるが、統合失調症を判別できる機械学習モデルを構築できる可能性が示された。

F. 研究発表

1. 論文発表

投稿中: He Y, Matsunaga M, Li Y, Kishi T, Tanihara S, Iwata N, Tabuchi T, Ota A. Artificial neural network for classifying the schizophrenia case: using Japanese online survey data.

2. 学会発表

なし

G. 知的財産権の出願・登録状況（予定を含む。）

1. 特許取得

なし

2. 実用新案登録

なし

3. その他

なし