

ライフログデータを活用したうつ病の 再燃検知法に関する基礎検討

平尾 歩¹ 田近 亜蘭² 古川 壽亮³ 池田 和司¹ 吉本 潤一郎^{1,a)}

概要: うつ病は悲しみや興味の喪失が持続する症状を特徴とする精神疾患であり、再燃率が高いことが知られている。再発時の病状の進行を防ぎ、その予後を改善するためには、早期発見・早期介入が不可欠である。本研究では、医師やカウンセラーが直接介入することなく、早期にうつ病の再燃兆候を検知するシステムの実現を目指し、ライフログデータからうつ病の再燃を予測する方法論について検討を行った。具体的には、87名の寛解期うつ病患者からスマートフォンとウェアラブルウォッチを用いて取得した1年間に渡るライフログデータから、定期的にK6やPHQ-9により評価したうつ病の再燃期にあるか否かを判別する予測モデルを構築した。そして、Leave-one-subject-out 交差検証法によってその予測モデルの汎化性能を評価し、その問題点と今後の課題を明らかにした。

キーワード: うつ病, 再発・再燃予測, ライフログ, ロジスティック回帰モデル, イベントトリガー平均

1. はじめに

持続的な抑うつ気分と意欲の低下を主症状とするうつ病は、一度寛解しても再燃率が高い精神疾患である。例えば、うつ病の初回の治療後の再燃率は約60%であり[1]、うつ病を2回以上経験した際に診断される反復性うつ病では、70-90%が再燃するとされている[2][3]。また、うつ病エピソードの回数や長さは、寛解後における記憶機能低下のリスク要因となり得る[4]。その一方で、2回目うつ病エピソード時でも早期介入できれば、その期間が大幅に短縮できることが報告されている[5]。したがって、精神医学の臨床現場においては、うつ病の再燃を早期に検知あるいは予知し、介入に繋げることが、反復性うつ病の予後を改善するための重要な課題とされている[6]。

現在、うつ病の標準的な診断基準としてはDSM5[2]が用いられており、その評価方法にはHamilton Rating Scale for Depression (HAM-D) [7]やBeck Depression Inventory-II (BDI-II) [8], Patient Health Questionnaire-9 (PHQ9) [9], kessler 6 scale (K6) [10]などが存在している。また既に脳活動などから予測する研究としては内側前頭前野の気分連動反応[11]などがある。これらの予測にはMRIなどの機器、また医師の協力が必要であり、そのために2つの課題

が存在する。1) MRIを用いた予測は計測費用が高額であること、2) 患者の精神状態によっては来院が難しく。また頻繁に計測するものではないことから計測していない期間の患者の症状の状態がわからないことである。

一方、近年のスマートフォンやウェアラブルデバイスの普及やセンシング技術の向上によって、ユーザの日々の活動をライフログとして継続的かつ長期に渡って記録することが容易になってきている。ライフログとして記録することができる身体活動や睡眠状態は、うつ症状と相関しており[12]、ライフログの活用は、うつ病再燃の早期の予知・検知するための有望な情報を含んでいると考えられる。実際に、寛解期のうつ病患者を対象とした[13]では、約1年間に渡って収集したライフログデータに対してパネルベクトル自己回帰分析を適用した結果、長時間睡眠が3週間後のうつ病の再燃を予測するための主要因子となり得ることが示唆されている。そこで、本研究では、スマートデバイスとライフログの活用したうつ病再燃予知システムの開発を長期的な目標に見据え、機械学習法を用いた再燃予測モデルの構築とその予測精度の評価および改善方法について検討した。具体的な検討項目は以下の通りである。

- I. [13]の報告に基づき、睡眠時間を説明変数（入力データ）として早期のうつ病再燃の予測が実用的な精度で可能であるかどうか。
- II. 睡眠以外のライフログ情報を用いることによって、項目I.の予測精度を改善することが可能であるかどうか。

¹ 奈良先端科学技術大学院大学

² 京都大学医学部附属病院精神科神経科

³ 京都大学大学院医学研究科

a) juniti-y@is.naist.jp

III. 項目 II. で有意な予測精度の改善が認められるならば、どの活動が予測因子として重要であるのか。

これらの研究課題に対する答えを導くために、[13]にて用いられたデータを二次解析し、ロジスティック回帰を弱学習器とするアンサンブル学習法によりうつ病再燃予測モデルを構築した。予測モデルの性能は、Leave-One-Subject-Outを用い、PPV, NPV, ROC 曲線で評価した。

2. 手法

2.1 倫理的配慮

本研究は、奈良先端科学技術大学院大学 (2017-M-5, 2017-M-6), 京都大学大学院医学研究科 (R0591-6), 名古屋市立大学病院 (50-16-0001), 高知大学医学部 (28-67), 広島大学病院 (C-105-3), 東邦大学医学部 (A17062) の倫理審査委員会による承認を得て実施した。

2.2 研究参加者

本研究は、寛解期うつ病患者を対象にした観察研究 [13] で収集されたデータの二次解析である。データ収集方法や研究参加者の特徴の詳細は [13] でも述べられているが、以下ではそれを簡潔に説明する。研究参加者は、2016年3月から2017年4月にかけて、4つの大学 (高知大学, 広島大学, 名古屋市立大学, および, 東邦大学) の付属病院と、それらの関連病院やクリニックに外来通院しているうつ病患者から募り、京都大学が参加登録の中央事務局となった。興味を示した患者に対しては、本研究の意義や趣旨を説明し、そのうち選択基準を満たし、かつ、同意書を取得できた患者を研究参加者として登録した。

なお、患者は6つの選択基準により選択された。1) 年齢が22歳以上69歳以下である。2) DSM-5の反復性大うつ病性障害の基準を満たす。3) Beck Depression Inventory-II (BDI-II) のスコアが9点以下である [8]。4) 不安障害や気分変調症の合併の有無は問わない。5) 携帯電話 (iPhone) の操作に十分に慣れている。6) 本研究の意義とリスクを十分に理解し、自ら参加を希望している。

また、以下の除外基準のうち1つでも該当する患者は、本研究の参加者からは除外した。1) 精神病性障害、双極性障害、物質使用障害、パーソナリティ障害のいずれかを合併している。2) 日本語が理解できない。3) 痛みや麻痺、ゴムバンドのアレルギー等の理由でシルミー W20 を装着できない。4) 他の治験に参加中である。5) その他、この研究への参加が不適切と判断される。

研究途中で、この研究への参加続行が不可能と主治医が判断した場合また躁転のように診断に変更があった場合の中止基準のどちらかに合致した場合は、その研究参加を中止した。

以上によって、100名が研究参加者として登録されたが、そのうち10名は研究途中で参加同意を撤回した。また、1

表 1 くらしアプリで記録される活動の種類

Table 1 Activity category recorded by Kurashi app

No.	活動内容	No.	活動内容
1	睡眠	9	買い物
2	ゴロゴロ, ボーッと	10	病院
3	食事・おやつ	11	交流・交際
4	風呂	12	スポーツ・運動
5	仕事・勉強	13	趣味・娯楽・習い事
6	通勤・通学	14	読書・新聞・雑誌
7	家事	15	TV・DVD・音楽
8	育児・介護	16	その他

くらしアプリでは、iPhone 内蔵のセンサ類によって計測される位置情報・歩数情報・動作情報をもとに、30分ごとの活動内容が上記16種類の1つに分類される。なお、この分類方法は総務省統計局社会生活基本調査 (<https://www.stat.go.jp/data/shakai/2016/index.html>) による分類に基づき選定された。

名は事後に選択基準を満たしていなかったことが判明したため研究対象から除外した。さらに、2名は研究期間全般に渡って重度のうつ症状を呈していた可能性があったため (後で説明する K6 のスコアが研究期間全般に渡って極度に高かったため)、それも研究対象から除外した。したがって、最終的には、計 87 名分のデータが本研究の研究対象となった。

2.3 解析データ

研究参加者のライフログデータは、スマートフォン Apple iPhone 用アプリである「くらしアプリ」*1 とリストバンド型生体センサ Silmee W20 (TDK 社) 用いて収集した。くらしアプリは、研究参加者本人が普段利用している iPhone にインストールするか、もしくは、研究参加者が iPhone を所有していない場合はくらしアプリがインストールされた iPhone を貸与した状態で研究を開始した。くらしアプリでは、iPhone から位置情報・歩数情報・動作情報が自動的に採取され、各時刻での歩数や移動距離が計算される。また、これらの情報に基づいて、30分ごとの活動内容が事前に定義された以下の16種類の1つに自動分類される (表 1)。研究参加者には、毎日活動内容が正しく分類されているかどうかを確認し、もし、誤りがあった場合には各自でそれらを修正するよう教示を行った。また、研究参加者自身にとって楽しいと感じた活動があれば、丸、二重丸、花丸によってそれを「元気度の自覚」としてアノテーションするよう教示を行った。Silmee W20 は、くらしアプリの補助センサーとして利用し、各時刻での歩数、エネルギー消費量、睡眠状態、脈拍数、紫外線レベル、会話時間が記録された。

ライフログデータに基づく再燃予測が実態とどれだけ一致するかを評価するために、期間中、研究参加者の抑うつ

*1 <https://apps.apple.com/jp/app/id1147944616>

症状の程度も、簡易調査票 K6 日本語版 [14] [10] を用いて調査した。K6 は a) 神経過敏, b) 絶望感, c) 焦燥感, d) 抑うつ気分, e) 気力・意欲の低下, f) 無価値感の 6 項目を過去 7 日に渡って振り返り, その頻度に応じて 5 段階 (0: 全くない, 1: 少しだけ, 2: ときどき, 3: たいてい, 4: いつも) で評価する手法である。6 項目に対する回答は合計値 (最小: 0 点, 最大: 24 点) によって集約され, 高得点ほど精神障害傾向が高いとされている。研究参加者には, 毎週任意のタイミングでくらしアプリを用いて K6 の質問に回答するように教示した。

2.4 特徴量抽出

くらしアプリと Silmee W20 から収集された情報は, 日ごとに以下のように集計された。一般に, 生活活動には個人内でも大きな多様性 [15] と週単位での周期性 [16] が存在する。この点を考慮して, くらしアプリおよび Silmee W20 から収集される基礎データ (歩数, 移動距離, 1 日あたりの各活動の時間など) は, その日を含む直前 7 日間の移動平均と移動標準偏差の 2 つの統計量により, 生活活動の経時変化を評価した。また, 16 種類の活動内容が同定できなかった未記入時間の多さや当日の抑うつ気分に対する自己評価と強い相関が推測される元気度の自覚は, 各日の合計点によって評価した。また, 先行研究 [13] では, 睡眠障害や食事摂取の状況が有望なうつ再燃の予測因子となり得ることが示唆されたことから, これら 2 項目についてはより細かい粒度で集計した。具体的には, 睡眠障害の程度を前日 12 時から当日 12 時までの合計睡眠時間, 起床時刻, 入眠時刻, 最長連続睡眠時間の 4 項目について, それぞれを直前 7 日間の移動平均, 移動標準偏差, 最大値, 最小値の 4 つの統計量で評価した。また, 起床後に社会活動に取り組むことができたかを, 朝 (5 時~10 時) の紫外線暴露レベルの直前 7 日間の移動平均と移動標準偏差によって評価した。日中の睡眠状況は, 当日 10 時から 18 時までの睡眠時間とゴロゴロ, ボツとの時間の 2 項目について, それぞれを直前 7 日間の移動平均, 移動標準偏差, 最大値, 最小値の 4 つの統計量で評価した。食事の状況については, その摂取時間に応じて朝食 (4 時台~10 時台), 昼食 (11 時台~15 時台), 夕食 (18 時台~22 時台) に分類し, それぞれを, 直前 7 日間の合計摂取回数によって評価した。以上の結果, 研究参加者の各日の生活活動状況は計 79 変数の値として集計された。

ただし, 研究参加者のうち 1 人でも全計測日中 30% 以上の欠損値が存在した変数は, 本研究の解析対象からは除外した。その結果, Silmee による計測が由来の変数はすべて除外され, くらしアプリによる計測が由来の計 64 変数が本研究の解析対象となった。

解析対象となった変数の中に欠損値が存在する場合は, 直前直後の非欠損値を用いて線形補間した。さらに, 生活

様式の個人差の影響を削減するために, 各変数ごとに Z スコアに変換したもの ((元の値 - 期間全体に渡る平均値) / (期間全体に渡る標準偏差)) を予測のための特徴量とした。

2.5 特徴量選択

計 64 個の特徴量からうつ病再燃予測に有用やものを絞り込むために, イベントトリガー平均法を利用した以下の変数選択法を適用した。まず, うつ病をスクリーニングするための標準的なカットオフ値は, K6 では 10 点以上 [17], PHQ9 では 5 点以上 [18] とされている。そこで, $K6 \geq 10$ が記録された日を「うつ病再燃期」, $K6 < 9$ が記録された日を「寛解期」と定義した。上記の各時点から 1 日ごとに 10 日前まで遡り, n 日前の各特徴量の値を対応する群 (「うつ病再燃期」群あるいは「寛解期」群のいずれか) に割り当てた。そして, 各特徴量ごとに各 n 日前 ($n = 0, \dots, 10$) のうつ病再燃期群と寛解期群との統計的な差を 2 標本 t 検定 (有意水準 0.05, 多重比較補正なし) によって評価し, 合計 4 日以上に有意差が認められたものを K6 に基づく有望な特徴量と定義した。

一方, PHQ9 については, 以下の手続きによって特徴量の順位付けを行った。まず, $PHQ \geq 5$ が記録された日を「うつ病再燃期」, $PHQ < 5$ が記録された日を「寛解期」と定義した。上記の各時点から 1 日ごとに 10 日前まで遡り, n 日前の各特徴量の値を対応する群に割り当てた。そして, 各特徴量ごとに各 n 日前 ($n = 0, \dots, 10$) のうつ病再燃期群と寛解期群との統計的な差をマンホイットニーの U 検定 (有意水準 0.05, 多重比較補正なし) によって評価し, 有意差が認められた日数が多い順に特徴量を順位付けした。この際, 同じ活動の種類 (表 1) から算出された特徴量については, 最上位のみを採用し, それ以外のものはこの順位付けから除外した。この順位付けによって, 上位 5 位以内に入ったものを PHQ9 に基づく有望な特徴量と定義した。

最後に, K6 と PHQ9 いずれでも有望な特徴量とされたものを以下の予測モデルの入力として選択した。

2.6 うつ病再燃予測モデルの学習と評価方法

うつ病再燃の予測可能性とともに特徴量選択の効果も検証するために, 以下の 3 種類の予測モデルを構築した。

モデル 1 [13] において睡眠に関する特徴量がうつ病再燃を予測するための主要因子と示唆されたことから, 予測対象日から n 日前 ($n = 0, \dots, N = 20$) の前日 12 時から当日 12 時までの合計睡眠時間 (直前 7 日間の移動平均) を入力として, 予測対象日がうつ病再燃期か寛解期かを分類するモデル。

モデル 2 2.4 節で抽出された全ての特徴量に関して, 予測対象日から n 日前 ($n = 0, \dots, N = 20$) のものを入力として, 予測対象日がうつ病再燃期か寛解期かを分類するモデル。

モデル 3 2.5 節で選択された一部の特微量に関して、予測対象日から n 日前 ($n = 0, \dots, N = 20$) のものを入力として、予測対象日がうつ病再燃期か寛解期かを分類するモデル。

各 3 つの予測モデルは、ロジスティック回帰モデルを弱学習器とするアンサンブル学習法 (Logistic Regression Ensembles; LORENS) [19] を単純化したものを用いて構築した。弱学習器となる各ロジスティック回帰モデルは、K6 計測日の n 日前 ($n = 0, \dots, N = 20$) の第 i 番目の特微量 ($i = 1, \dots, I$) を $x_{n,i}$ 入力して、その K6 計測日がうつ病再燃期である確率 $p_{n,i}$ を予測するものであり、以下の式で定義される。

$$p_{n,i} = 1 / (1 + \exp(-w_{n,i} \cdot x_{n,i}))$$

ここで、 $w_{n,i}$ はモデルパラメータであり、訓練データから最尤推定法により決定されるものである。また、 I はモデルに入力される特微量の数であり、モデル 1 では $I = 1$ 、モデル 2 では $I = 64$ 、モデル 3 では 2.5 節で説明した手順により選択された特微量の数である。各弱学習器からの出力は、

$$p = \frac{1}{I(N+1)} \sum_{i=1}^I \sum_{n=0}^N p_{n,i}$$

によって算術平均化され、予測モデル全体としての出力は以下で与えられる。

$$y = \begin{cases} 1 & (p \geq 0.5 \text{ の場合}) \\ 0 & (p < 0.5 \text{ の場合}) \end{cases}$$

ここで、 $y = 1$ はその日がうつ病再燃期 (正例) であることを、 $y = 0$ は寛解期 (負例) であることを表すものとする。

この予測モデルの汎化性能は、1 人の参加者から得られたデータをテストデータとし、残りの参加者から得られたデータをモデルパラメータ決定のための訓練データとするランを、すべての組み合わせに関して実行する Leave-One-Subject-Out (LOSO) 交差検定法により評価した。性能評価指標には、それぞれ以下の式で定義される陽性的中率 (Positive Predictive Value; PPV) と陰性的中率 (Negative Predictive Value; NPV) を用いた。

$$PPV = (\#TP) / (\#TP + \#FP)$$

$$NPV = (\#TN) / (\#TN + \#FN)$$

ここで、 $\#TP$ 、 $\#FP$ 、 $\#TN$ 、および、 $\#FN$ は、それぞれ、真陽性、偽陽性、真陰性、および、偽陰性となったサンプル数である。また、予測結果がチャンスレベルよりどの程度勝っているかを評価するために Receiver Operating Characteristic (ROC) 曲線の Area Under Curve (AUC) を求めた。

3. 結果

3.1 選択された特微量

K6 に基づく有望なものとして、以下の 13 個の特微量が選択された。

- (1) 「スポーツ・運動」時間の直前 7 日間移動平均
- (2) 「スポーツ・運動」時間の直前 7 日間移動標準偏差
- (3) 「ゴロゴロ・ボーッと」時間の直前 7 日間移動平均
- (4) 「ゴロゴロ・ボーッと」時間の直前 7 日間移動標準偏差
- (5) 「通勤・通学」時間の直前 7 日間移動平均
- (6) 日中 (10~18 時) の睡眠時間の直前 7 日間移動平均
- (7) 「仕事・勉強」の直前 7 日間移動平均
- (8) 最長連続「睡眠時間」の直前 7 日間移動標準偏差
- (9) 日中 (10~18 時) の「ゴロゴロ・ボーッと」時間の直前 7 日間最大値
- (10) 日中 (10~18 時) の「ゴロゴロ・ボーッと」時間の直前 7 日間移動平均
- (11) 日中 (10~18 時) の「ゴロゴロ・ボーッと」時間の直前 7 日間移動標準偏差
- (12) 前日 12 時~当日 12 時の入眠時刻の直前 7 日間移動標準偏差
- (13) 歩数の直前 7 日間移動平均

また、K6 に基づく有望なものとして、以下の 5 個の特微量が選択された。

- (1) 「ゴロゴロ・ボーッと」時間の直前 7 日間移動平均
- (2) 「通勤・通学」時間の直前 7 日間移動平均
- (3) 日中 (10~18 時) の睡眠時間の直前 7 日間移動平均
- (4) 「朝食」回数の直前 7 日間合計値
- (5) 前日 12 時~当日 12 時の起床時刻の直前 7 日間移動平均

したがって、2.6 節で定義した予測モデル 3 への入力には、以下の 3 つの特微量のみが選択された。

- 「ゴロゴロ・ボーッと」時間の直前 7 日間移動平均
- 「通勤・通学」時間の直前 7 日間移動平均
- 日中 (10~18 時) の睡眠時間の直前 7 日間移動平均

3.2 予測モデルの汎化性能

2.6 節で定義した 3 つの予測モデルの汎化性能を LOSO 交差検証法で評価した結果、各モデルの性能指標は表 3.2 の通りとなった。事前に変数選択を行ったモデル 3 が、陽性適中率 (PPV)、陰性適中率 (NPV)、および、AUC のいずれの指標においても最良となった。一方で、最良だったモデル 3 においても、AUC は 0.57 であり、その ROC 曲線はチャンスレベル (AUC=0.50) に比べてわずかに良い程度に留まった (図 1)。

表 2 各予測モデルの汎化性能評価結果

特徴量の種類	モデル 1	モデル 2	モデル 3
	$I = 1$	$I = 64$	$I = 3$
含まれる活動	睡眠	全て	ゴロゴロ・ボーッと 通勤・通学 睡眠
PPV	12.5%	13.9%	15.9%
NPV	87.3%	88.1%	89.4%
AUC	0.48	0.52	0.57

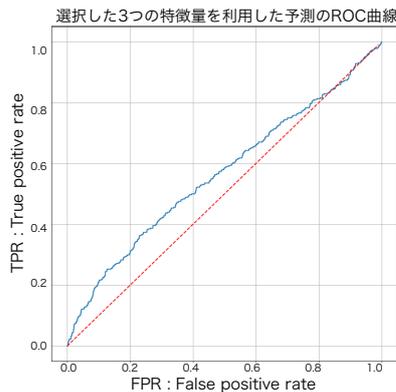


図 1 モデル 3 の ROC 曲線

4. 考察

4.1 結果のまとめと問題点

本研究では、[13] で収集されたデータを二次解析することにより、スマートデバイスによるライフログ計測を活用したうつ病再燃予測システムの実現可能性を検討した。この目的のために、先行研究の知見を基にライフログからの特徴量抽出をマニュアルで行い、その特徴量を入力とし、うつ病再燃期か寛解期かを予測する二値分類器を、ロジスティック回帰モデルを弱学習器とするアンサンブル学習法により構築した。また、二値分類器への入力が高次元となるため、イベントトリガー平均法による特徴量選択の効果を検討した。LOSO 交差検証法で構築した予測器の汎化性能を評価したところ、イベントトリガー平均法による特徴量選択を導入した効果は、陽性適中率、陰性適中率、および、AUC の 3 つの指標に関して認められた。一方で、今回構築したモデルのうち、最良のものであっても AUC は 0.57 に留まり、その予測精度はチャンスレベル相当のものであった。

本報告の内容は予備検討の段階であり、まだいくつか検討すべき事項がいくつかある。第一に特徴量選択方法についてである。今回の検討では、イベントトリガー平均法を適用する際に全ての解析対象データを利用している。これは、後の予測器構築およびその性能評価のステップから見ると、本来完全に隠蔽されるべきテストデータの情報が漏洩していることとなる。実用的な予測器を構築する際に

は、nested 交差検定法を用いるなど特徴量選択の際にはテストデータの情報が漏洩しないような工夫が必要である。

特徴量選択時にテストデータの情報が漏洩し、予測精度を過大評価する状況にあるにも関わらず、最良のモデルでも、AUC は 0.57 とチャンスレベル相当に留まった点は、特徴量抽出と予測モデル構築についても問題が残されている。今回は、実装の容易さと解釈可能性の高さから、予測モデルにはロジスティック回帰を弱学習器として利用したが、その表現能力は線形な判別関数に制約されるため、今回の二値分類問題を解くには不十分であった可能性がある。今後、ランダムフォレストやカーネル法などの非線形モデルも弱学習器として採用することも検討する必要がある。

特徴量抽出も予測精度向上のためには重要な問題である。今回見逃してしまった重要な特徴量がないかを検討するために、以下では、その簡単な解析を行った。

4.2 新たな特徴量の探索

[13] では睡眠がうつ病再燃の主要因子として報告されているが、一般的にも睡眠障害はうつ病の中核症状の一つだと考えられている [12]。一方で、本研究では睡眠時間については考慮したが、そのタイミングについては全く考慮されていなかった。そこで、入眠時刻と起床時刻の 2 点に関してうつ病再燃の予測因子となり得るかについて簡単な予備解析を行った。

まず、解析対象となった 87 名の研究参加者を、再燃群と非再燃群の 2 群に分割した。ここで、再燃群には研究期間中に 1 度でも $K6 \geq 10$ が計測された研究参加者を割り当て、それ以外の患者を非再燃群に割り当てた。ただし、1 年間の研究期間中に週に 1 度の頻度で合計 53 回の K6 を計測する機会があったにも関わらず、実際に K6 を計測した回数とその半分に対応する 27 回未満だった研究参加者については、再燃のタイミングを見逃している可能性が高いため解析対象から除外した。その結果、再燃群には 35 名が、非再燃群には 31 名が割り当てられた。

そして、各研究参加者ごとに研究期間を通した入眠時刻の最頻値と標準偏差、起床時刻の最頻値と標準偏差の 4 つの統計量を算出し、再燃群と非再燃群との群間比較をマンホイットニーの U 検定 (有意水準 0.05, 多重比較補正なし) を用いて評価した。

その結果、入眠時刻については最頻値と標準偏差のいずれも再燃群と非再燃群の間に有意な差は認められなかった。また、起床時刻の最頻値についても再燃群と非再燃群の間に有意な差は認められなかった。一方で、起床時刻の標準偏差については再燃群と非再燃群の間に有意な差は認められた (図 2)。このことから、睡眠の時間よりも、起床時刻のばらつきがうつ病再燃の兆候を示す特徴量となり得る可能性が示唆された。ただし、これはグループ解析の結果であるため、兆候というよりはむしろうつ病再燃のリス

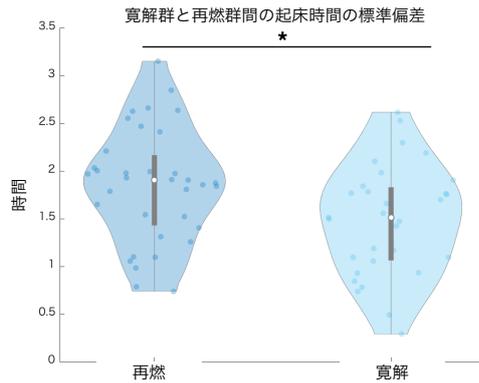


図2 再燃群と非再発群の起床時刻の標準偏差

ク要因となっている可能性も否定できないことに注意が必要である。

4.3 今後の課題

以上を整理すると、今後解決すべき課題として以下が挙げられる。

- 特徴量抽出: 睡眠に代表されるように、各活動の継続時間や回数だけではなく、その開始や終了のタイミングを考慮した特徴量の抽出。
- 特徴量選択: 特徴量選択時における、汎化性能検証用テストデータの情報漏洩防止。
- 予測モデルの構築: ランダムフォレストや決定木、カーネル法のような非線形モデルの導入。

これらの問題を解決しながら、ライフログデータからのうつ病再燃予測が実用的な精度で可能かどうかについて検討を進めるとともに、生理指標（心拍変動や皮膚電位活動など）も活用した予測方法についても今後検討予定である。

謝辞

本研究の一部は、厚生労働科学研究費補助金（課題ID: 17948487）および JST 未来社会創造事業（課題ID: 20349577）の支援を受けて実施された。

参考文献

[1] Hirschfeld, R. M.: Clinical importance of long-term antidepressant treatment, *The British Journal of Psychiatry*, Vol. 179, No. S42, pp. s4-s8 (2001).

[2] American Psychiatric Association: *Diagnostic and statistical manual of mental disorders (DSM-5®)*, American Psychiatric Association (2013).

[3] Burcusa, S. L. and Iacono, W. G.: Risk for recurrence in depression, *Clinical psychology review*, Vol. 27, No. 8, pp. 959-985 (2007).

[4] Gorwood, P., Corruble, E., Falissard, B. and Goodwin D Phil, F Med Sci, G. M.: Toxic effects of depression on brain function: impairment of delayed recall and the cumulative length of depressive disorder in a large sample of depressed outpatients, *American Journal of Psychiatry*, Vol. 165, No. 6, pp. 731-739 (2008).

[5] Kupfer, D. J., Frank, E. and Perel, J. M.: The advantage

of early treatment intervention in recurrent depression, *Archives of General Psychiatry*, Vol. 46, No. 9, pp. 771-775 (1989).

[6] Kraus, C., Kadriu, B., Lanzenberger, R., Zarate Jr, C. A. and Kasper, S.: Prognosis and improved outcomes in major depression: a review, *Translational psychiatry*, Vol. 9, No. 1, pp. 1-17 (2019).

[7] Hamilton, M.: A rating scale for depression, *Journal of neurology, neurosurgery, and psychiatry*, Vol. 23, No. 1, p. 56 (1960).

[8] Hiroe, T., Kojima, M., Yamamoto, I., Nojima, S., Kinoshita, Y., Hashimoto, N., Watanabe, N., Maeda, T. and Furukawa, T. A.: Gradations of clinical severity and sensitivity to change assessed with the Beck Depression Inventory-II in Japanese patients with depression, *Psychiatry research*, Vol. 135, No. 3, pp. 229-235 (2005).

[9] Kroenke, K., Spitzer, R. L. and Williams, J. B.: The PHQ-9: validity of a brief depression severity measure, *Journal of general internal medicine*, Vol. 16, No. 9, pp. 606-613 (2001).

[10] Kessler, R. C., Andrews, G., Colpe, L. J., Hiripi, E., Mroczek, D. K., Normand, S.-L., Walters, E. E. and Zaslavsky, A. M.: Short screening scales to monitor population prevalences and trends in non-specific psychological distress, *Psychological medicine*, Vol. 32, No. 6, pp. 959-976 (2002).

[11] Farb, N. A., Anderson, A. K., Bloch, R. T. and Segal, Z. V.: Mood-linked responses in medial prefrontal cortex predict relapse in patients with recurrent unipolar depression, *Biological psychiatry*, Vol. 70, No. 4, pp. 366-372 (2011).

[12] Nutt, D., Wilson, S. and Paterson, L.: Sleep disorders as core symptoms of depression, *Dialogues in clinical neuroscience*, Vol. 10, No. 3, p. 329 (2008).

[13] Kumagai, N., Tajika, A., Hasegawa, A., Kawanishi, N., Horikoshi, M., Shimodera, S., Ken'ichi Kurata, Chino, B., Furukawa, T. A.: Predicting recurrence of depression using lifelog data: an explanatory feasibility study with a panel VAR approach, *BMC psychiatry*, Vol. 19, No. 1, p. 391 (2019).

[14] Furukawa, T. A., Kawakami, N. et al.: The performance of the Japanese version of the K6 and K10 in the World Mental Health Survey Japan, *International journal of methods in psychiatric research*, Vol. 17, No. 3, pp. 152-158 (2008).

[15] Schönfelder, S. and Axhausen, K. W.: On the Variability of Human Activity Spaces, *The Real and Virtual Worlds of Spatial Planning*, Springer Berlin Heidelberg, Berlin, Heidelberg, chapter 17, pp. 237-262 (2004).

[16] Hu, F. and Smeaton, A. F.: Periodicity intensity for indicating behaviour shifts from lifelog data, *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, IEEE, pp. 970-977 (2016).

[17] 川上 憲人: 一般住民におけるトラウマ被害の精神影響の調査手法マニュアル (2015年2月版), <http://plaza.umin.ac.jp/heart/pdf/151026.pdf> (2015).

[18] 米永裕紀, 青山真帆, 森谷優香, 五十嵐尚子, 升川研人, 森田達也, 木澤義之, 恒藤 暁, 志真泰夫, 宮下光令: 緩和ケア病棟のケアの質および遺族の悲嘆・うつの地域差: 全国遺族調査の結果から, *Palliative Care Research*, Vol. 13, No. 3, pp. 235-243 (2018).

[19] Lim, N., Ahn, H., Moon, H. and Chen, J. J.: Classification of high-dimensional data with ensemble of logistic regression models, *Journal of Biopharmaceutical Statistics*, Vol. 20, No. 1, pp. 160-171 (2009).