

デジタルバイオマーカー活用による メンタルヘルス不調予測に向けた試み ——スマートウォッチを用いた縦断研究——

A Study for Prediction of Mental Health Problems Through the Use of Digital Biomarkers :
A Smartwatch Longitudinal Study

高橋 史也^{✉1)}, 橋本 里奈¹⁾, 滝沢 龍^{✉1,2)}

Fumiya Takahashi[✉], Rina Hashimoto, Ryu Takizawa[✉]

現状のメンタルヘルスの評価は自己報告に基づいて行われることが中心だが、リコールバイアスなどのバイアスの影響を受けやすく精度を欠くことが指摘されてきた。また、メンタルヘルス不調の早期発見のために、毎日質問紙に回答することは負担が大きい。そのような背景から、近年注目されているのがデジタルバイオマーカーである。デジタルバイオマーカーは、デジタルデバイスを通して取得される日々のさまざまなデータ（身体活動指標、睡眠指標、生理指標など）のことで、データを簡便かつ継続的に取得できることなどからメンタルヘルス不調の早期発見に適した形態である。本研究では、Fitbit社の新型スマートウォッチ Fitbit Sense が感知情報に基づいて算出する身体活動指標、睡眠指標、生理指標を用いて、自記式の質問紙 DASS21 に基づく抑うつ、不安、ストレスを説明することを目的とした。対象者は18～29歳の健常者と閾下症状をもつ者88名で、4週間スマートウォッチを常時着用し、着用中1週間ごとにweb上の質問紙に回答した。Fitbit Sense から得られたデジタルバイオマーカーを説明変数とし、DASS21の抑うつ、不安、ストレスを目的変数とするステップワイズ法による重回帰分析を行った結果、直前2週間におけるデジタルバイオマーカーの変化（Fitbit Sense から推算されたレム睡眠の時間の増加と深い睡眠の時間の減少）がその直後のメンタルヘルス不調を予測している可能性が示された。また、それは主観的評価では予測できず、デジタルバイオマーカーによってのみ予測ができる可能性を示唆していた。なお、本結果はベースライン値およびデモグラフィック指標の値を統制して得られた統計学的により頑健な示唆である。加えて、Fitbit Sense から推算されたレム睡眠の時間の増加に関しては、抑うつ得点の上昇のみならず、重症度のカテゴリー移行（悪化側）も予測できる可能性が明らかになった。本結果は、メンタルヘルス領域におけるデジタルバイオマーカーの応用可能性を裏付けた。

著者所属：1) 東京大学大学院教育学研究科臨床心理学コース, Department of Clinical Psychology, Graduate School of Education, The University of Tokyo
2) ロンドン大学, キングズ・カレッジ・ロンドン, 精神医学・心理学・神経科学研究所, Institute of Psychiatry, Psychology and Neuroscience, King's College London

✉ E mail : beatenaciousman@gmail.com ; takizawar-ky@umin.ac.jp

受付日：2024年10月23日

受理日：2025年1月27日

doi : 10.57369/pnj.25-074

はじめに

メンタルヘルス不調は発見が遅れると、不登校や休職など日常生活に支障をきたしたり、最悪の場合は自殺につながったりすることもある。そのため、予防や早期発見が非常に重要である。現状の標準的なメンタルヘルスのアセスメントは、本人の語りや質問紙などの主観的評価に基づいて行われるが、主観的評価はリコールバイアス³³⁾や社会的望ましさのバイアス²⁸⁾の影響を受けやすく精度を欠くことが指摘されてきた¹⁸⁾。加えて、自分の状態を自分で正確に把握することには難しさがあることが指摘されており³⁰⁾、気づいたら重症化していたというケースも少なくない。

近年、そのようなメンタルヘルスのアセスメント手法における課題の解決に向けてデジタル技術の活用が注目を集めている。デジタル技術を活用した客観的評価指標が「デジタルバイオマーカー(digital biomarkers)」である。デジタルバイオマーカーは、「ポータブル、ウェアラブル、埋め込み型、または消化可能な機器に埋め込まれたセンサーによって収集される客観的、定量化可能、生理学的、および行動的なデータ」として定義される²⁹⁾。デジタルバイオマーカーに非常に類似した概念として、デジタル表現型(digital phenotyping/digital phenotype)がある。デジタル表現型は、「パーソナルデジタルデバイスから得られるデータを使って個人レベルのヒト表現型を瞬間瞬間で定量化すること」と定義される³⁴⁾。デジタルバイオマーカーやデジタル表現型という概念は誕生して間もなく、両者の明確な違いは報告されていない。本研究では、以下デジタルバイオマーカーという用語を用いる。

デジタルバイオマーカーは、上で述べた主観的評価の精度に関する課題解決に寄与するだけではない。例えば、メンタルヘルス不調の予防や早期発見のためには日々のセルフモニタリングが必要不可欠であるが、毎日質問紙に回答することは時間や手間の観点で大きな負担であり現実的ではない。デジタルバイオマーカーはデジタルデバイスを身につけておくだけで自動的・受動的にデータを取得することができるため、セルフモニタリングを容易にすることができる。また、日常生活で得られるリアルタイムデータを

継続的かつ長期的に測定することができる点も予防や早期発見に適していると考えられる。加えて、本研究はCOVID-19の世界的流行下で行われたが、メンタルヘルス領域においても対面の援助機会が失われているなかで、対人接触を伴わないパーソナルデバイスを用いたデータ取得の意義はますます高まっていくであろうし、医療資源やコストの削減にも貢献する可能性を秘めている。

これまで、メンタルヘルス分野においてもデジタルバイオマーカーを活用する研究が活発に行われている。諸外国において2004年頃から研究数は確実に増えており²¹⁾、近年、国内でも少なからず研究がなされてきているが、課題克服の困難さから臨床上確実に実用価値のあるデジタルバイオマーカーの特定にまでは至っていない³¹⁾。本研究では、課題として挙げられているデジタルデバイスの技術的課題やサンプルサイズの不十分さに焦点をあてながら、デジタルバイオマーカー活用によるメンタルヘルス不調予測に向けた試みを行った。

まず、デジタルデバイスの技術的課題に関して、多くの研究がデジタルバイオマーカーを取得するために侵襲性の低さなどを考慮してウェアラブルデバイスを使用しているが、バッテリー容量の限界から毎日充電する必要があることや防水機能をもたないなどの技術的な課題が存在していた。それらが継続的な測定に影響を与えているという報告もあった³⁷⁾。本研究では、ウェアラブルデバイスとしてスマートウォッチのFitbit Senseを使用した。Fitbit SenseはFitbit社が日本では2020年に発売した新型の市販デバイスである。Fitbit Senseは1週間に一度の充電で継続的使用が可能であり防水性も備えている。また、われわれが調査した限りにおいて、Fitbitを使用した研究の多くは旧型モデルのFitbitを使用していた。精度向上の観点でも新型デバイスを使用した研究は重要であると考えられる。

次に、サンプルサイズの不十分さに関して、本研究領域における先行研究の多くは探索的な研究のために数十名程度の小さなサンプルサイズにとどまっていることが報告されている²⁹⁾。デバイス1台あたりのコストが高いことに加え、長期の縦断調査となればよりコスト的負担は大きくなることもその一因かもしれない。しかし、膨大で多様なデータであるデジタルバイオマーカーを扱う研究において頑健な結果を得るためには十分なサンプルサイズを確保す

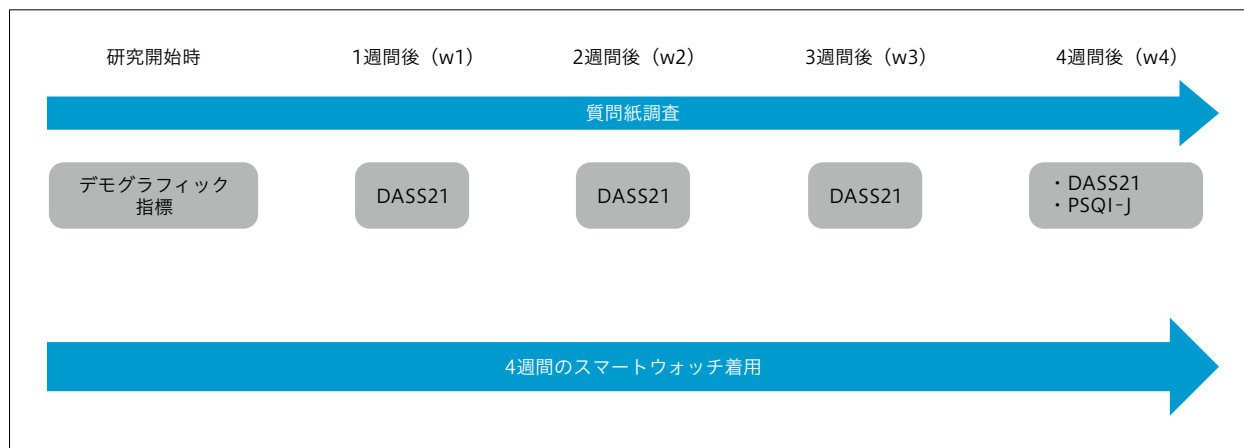


図 研究の流れ

ることが重要である。本研究では、従来の研究、特に日本国内の従来の研究と比較してより大きなサンプルサイズでデータを取得した。

これまでの議論をふまえ本研究の目的は以下3点である。

〔目的 1〕

Fitbit社の新型スマートウォッチFitbit Senseが感知情報に基づいて算出する身体活動指標、睡眠指標、生理指標を用いて、自記式の質問紙 Depression, Anxiety, Stress Scale 21 (DASS21) に基づく抑うつ、不安、ストレスを説明する。スマートウォッチはこれまでの技術的課題を克服した新型デバイスを使用し、比較的大きなサンプルからデータを収集する。

〔目的 2〕

目的 1 で説明されたデジタルバイオマーカーに対して、それと関連する主観的指標との比較を行い、両者の同質性や異質性を検討する。つまり、関連する主観的指標でもメンタルヘルス不調の予測の可能性があるのか、それとも、客観的指標であるデジタルバイオマーカーのみで予測の可能性があるのか、あるいは、どちらでも予測が可能だが組み合わせることでより予測の可能性が高まるのかということをも明らかにする。

〔目的 3〕

より臨床的意義の大きい結果を得るために、目的 1 と目的 2 を通してメンタルヘルス不調の予測を説明するデジタルバイオマーカーおよび主観的指標が特定された場合、それらを用いて、軽度から中程度、中程度から重度などのメンタルヘルス不調の重症度のカテゴリー移行（悪化側）を予測できる可能性があるのかを検証する。

1. 方 法

1. 研究協力者

本研究は2021年8月から2022年5月にかけて実施された。「18～29歳である」「現在精神疾患の治療中ではない」「未婚かつ子どもがいない」「妊娠中ではない」のすべてを満たす研究協力者を募集した。募集にあたっては、機縁法のほか、研究室ホームページや、研究協力者を募るソーシャルネットワークのTwitter（現在のX）アカウントでも告知した。募集期間内に上記条件を満たす105名からの応募があったが、研究協力者自身の都合やスマートウォッチの不具合により17名が中途ドロップアウトになった。そのため、研究を完遂した研究協力者は88名であった（男性43名、女性45名、平均年齢 23.02 ± 2.66 歳）。

2. 手続き

スマートウォッチを研究協力者の自宅に郵送し、4週間の常時着用（入浴時は除く）を求めた。4週間の着用終了後に、われわれが研究協力者のアカウントにログインし、4週間分のデータ出力を行った。研究協力者は、スマートウォッチの画面上およびスマートフォンのアプリケーション上で日々のデータを確認することができた。

質問紙調査はweb上（Google フォーム）で行った。調査のタイミングは、研究開始時、開始から1週間後（以下、w1）、2週間後（以下、w2）、3週間後（以下、w3）、4週間後（以下、w4）の合計5時点であった。回答に要する時間はいずれも10～15分程度であった。以上の研究の流れを図に示した。

3. デジタルデバイスとデジタルバイオマーカー

本研究ではウェアラブルデバイスとしてスマートウォッチの Fitbit Sense を使用した。Fitbit Sense は Fitbit 社が日本では 2020 年に発売した新型の市販デバイスで、健康管理に重点を置いている。決済、通話、メッセージ、音楽コントロールなどとしての用途ももつ多機能型デバイスではあるが、本研究では研究協力者ごとに使用条件がバラつくことを防ぐためにそれらの機能は使用しないよう求めた。また、防水機能も備えているが、耐水性は時間の経過とともに低下すること、石鹸やシャンプーなどによって悪影響を受けること、研究協力者の皮膚への影響などを考慮して水中や入浴中の使用は制限した。ただし、入浴後の再着忘れによるデータ欠損の問題も指摘されているため³⁵⁾、研究開始時に再着忘れを忘れないようにとの注意喚起を行った。さらに、Fitbit Sense は 1 週間に 1 度の充電で継続的使用が可能であり負担が少ないデバイスといえる。

Fitbit Sense には以下のセンサーおよびモーターが搭載されている¹²⁾。

- ・運動パターンを追跡する 3 軸加速度計
- ・ジャイロ스코プ
- ・高度の変化を追跡する高度計
- ・GPS 受信機と GLONASS 内蔵（ワークアウト時に位置情報を追跡）
- ・マルチパス光学式心拍数トラッカー
- ・ECG アプリと EDA スキャンアプリに対応した多目的電気センサー
- ・手首の皮膚温センサー
- ・周辺光センサー
- ・マイク
- ・スピーカー
- ・振動モーター

上記センサーやモーターを用いて取得しデータとして出力できるデジタルバイオマーカー一覧を身体活動指標、睡眠指標、生理指標に分類し、それぞれの小分類と詳細を表 1 に示した。表 1 内の詳細は Fitbit 公式ホームページを参考に記載した¹³⁾。

4. Fitbit の測定精度

Fitbit の測定精度に関して、歩数は一貫して正確であるという報告や¹⁴⁾、歩数と心拍は平均して正確であるという報告がある⁹⁾。Fitbit の睡眠データに関しては、従来の脳波を用いた睡眠の推定と比較して一定の信頼性があるという

報告もある¹⁶⁾。一方で、消費カロリーに関しては不正確であるという指摘や^{9,14)}、心拍と睡眠データの正確性に関して問題点を指摘する研究もある⁴⁾。

このように、Fitbit の精度に関してはいまだ明確な結論が出ていないのが現状であると考えられる。その要因としては、研究の質や Fitbit のモデルの違いなどが挙げられているが⁹⁾、われわれが調査した限りにおいても、Fitbit の精度を検証した研究の多くは旧型モデルの Fitbit を使用していた。

本研究で使用する新型モデルの Fitbit Sense では精度がさらに向上している可能性が高い。また、「臨床現場でのウェアラブルデバイスの信頼性については議論の余地があるが、人間の生理学研究では、その利用が近年受け入れられつつある」と述べている研究があるように³⁷⁾、精度の向上は今後も重要と考えられる一方、市場規模が年々拡大しているスマートウォッチに対する期待値が大きいのも事実である。そのため、本研究において Fitbit Sense を使用することは一定の妥当性があると考えた。

5. 評価指標（質問紙）

1) デモグラフィック指標

年齢、性別、現在の立場・身分（学生または社会人）、居住形態（同居または独居）、飲酒の有無、喫煙の有無、運動習慣の有無、BMI、親の教育年数について尋ねた。親の教育年数は両親の最終学歴から教育を受けていた年数を割り出し、両親のうち長いほうを選んだ。親の教育年数は個人または家族の社会状況を表す Socioeconomic Status (SES) の指標として用いた。これらすべてのデモグラフィック指標は研究開始時に尋ねた。

2) メンタルヘルス指標

(1) Depression, Anxiety, Stress Scale (DASS21)

DASS21²²⁾は抑うつを測定する質問が 7 問、不安を測定する質問が 7 問、ストレスを測定する質問が 7 問の合計 21 問で構成される。日本語版の信頼性と妥当性は確認済みである²³⁾。

「この 1 週間のあなたに起こった感覚にもっともあてはまる番号に○をつけてください」との質問に対し、「まったくあてはまらない (0)」「時々、あてはまる (1)」「かなりあてはまる (2)」「ほとんどあてはまる (3)」の 4 件法で各項目への回答を求めた。抑うつ、不安、ストレスそれぞれ 7 問に関して、合計を 2 倍した得点によって重症度を評価する (0 点～42 点)。抑うつに関しては、0 点～9 点が

表 1 Fitbit Sense から取得されたデジタルバイオマーカー一覧

大分類	小分類	詳 細
身体活動指標	歩数	3 軸加速度計により計測
	移動距離	歩数×歩幅により計測
	登った階数	高度計センサーにより計測
	総消費カロリー	基礎代謝率, アクティビティデータ, 心拍数により計測
	運動による消費カロリー	アクティビティデータ, 心拍数により計測
	静止状態の時間	1.5 メッツ*以下の活動（心拍数により計測）
	ややアクティブな時間	1.5～3.0 メッツの活動（心拍数により計測）
	かなりアクティブな時間	3.0～6.0 メッツの活動（心拍数により計測）
	とてもアクティブな時間	6.0 メッツ以上の活動（心拍数により計測）
睡眠指標	就寝時間	体が完全に安静で 1 時間ほど動いていないとデバイスは睡眠中であることを記録
	起床時間	
	睡眠時間	合計時間－（覚醒時間または落ち着きのない眠りの時間）により計測
	ベッドにいた時間	動きのパターン, 心拍変動により計測
	目覚めた回数	
	目覚めていた時間	
	浅い睡眠の時間	
	深い睡眠の時間	
生理指標	レム睡眠の時間	
	安静時心拍数	光学式心拍センサーによって腕上部の毛細血管の容量変化を検出し計測
	心拍変動	心拍間隔の変動
	呼吸数	夜間の 1 分間の平均呼吸数, 光学式心拍センサーにより計測
	皮膚温の変動	皮膚の表面温度で, 深部体温とは異なる
	血中酸素飽和度	血液に含まれる酸素の割合で, デバイスの背面にある赤色光と赤外線センサーにより計測
	皮膚電気活動 (EDA**)	皮膚の汗の量の変化で, デバイスの電極により計測

*メッツ (METs) とは, 運動強度の単位で, 安静時を 1 としたときと比較して何倍のエネルギーを消費するかで活動の強度を示したものである¹⁹⁾.

**EDA: electrodermal activity

正常範囲内, 10 点～13 点が軽度, 14 点～20 点が中程度, 21 点以上が重度と評価される。不安に関しては, 0 点～7 点が正常範囲内, 8 点～9 点が軽度, 10 点～14 点が中程度, 15 点以上が重度と評価される。ストレスに関しては, 0 点～14 点が正常範囲内, 15 点～18 点が軽度, 19 点～25 点が中程度, 26 点以上が重度と評価される。DASS21 は w1, w2, w3, w4 の合計 4 時点で尋ねた。

3) 睡眠指標

(2) ピッツバーグ睡眠質問票 (PSQI)

ピッツバーグ睡眠質問票 (Pittsburgh Sleep Quality Index: PSQI)⁶⁾は, 18 項目で構成され, 睡眠の質, 入眠時

間, 睡眠困難, 睡眠時間, 睡眠効率, 睡眠薬の使用, 日中覚醒困難の 7 つの指標から, 過去 1 ヶ月における主観的な睡眠の質を総合的に評価する。日本語版 PSQI-J の信頼性と妥当性は確認済みである¹¹⁾

就寝時刻, 睡眠潜時, 起床時間, 睡眠時間に関する質問項目については該当する数字を記入し, それ以外の質問項目についてはリッカート尺度 4 件法 (0～3) で評価する。点数が高いほど睡眠の質が悪いことを示す。PSQI-J は w4 の 1 時点で尋ねた。

6. 倫理的配慮

本研究は、東京大学ライフサイエンス委員会倫理申請専門委員会による審査および承認を受けたのち、東京大学大学院教育学研究科長の承認を得て実施された（承認番号：21-193）。本研究の概要、参加の任意性と撤回の自由、個人情報保護の保護、研究成果の発表、研究参加者にもたらされる利益と不利益、資料・情報の取り扱い方針、費用負担などを記載したインフォームド・コンセント資料を研究開始時のweb上質問紙1ページ目に記載し、研究協力者から同意を得た。

7. 解析

1) 目的1における解析の概観

まず、短期的なメンタルヘルス不調の予測に寄与するスマートウォッチから取得されるデジタルバイオマーカーを特定するための解析を行った（解析1, N=80）。解析1（w4を目的変数）では、w2とw3の2週間における各デジタルバイオマーカーの1日平均を説明変数とし、w4の抑うつ、不安、ストレスのそれぞれを目的変数とする重回帰分析（ステップワイズ法）を行った。それぞれの重回帰式において、w1の目的変数（ベースライン）および、目的変数と有意な関連があるデモグラフィック変数を統制変数として投入した。

精神疾患の世界的な診断基準として用いられているDSM-5を参照すると、例えば、うつ病の診断基準は症状が2週間以上続いていることと記載されており、2週間での変動が不調を示す最低基準になると考え、本研究においても説明変数をw2とw3の2週間における1日平均として算出した。

すべての解析において、説明変数と目的変数に時間的前後関係があるため、因果関係の強い推論が可能となる。さらには、w1の目的変数（ベースライン）を統制変数として加えることで、変化（不調）を予測するデジタルバイオマーカーを特定することができる。統計学的には、予測に関してベースライン値の影響を取り除いたより頑健な結果を得ることができる。

欠損値に関しては、欠損値を含むデータを除外して解析する完全ケース分析を実施した。行った解析によって欠損値の数が若干異なる。

分析はオープンソースの統計ソフトウェア環境であるR4.2.2上で行った。検定はすべて有意水準を5%に設定して両側検定を行った。

表2 重回帰分析で投入した説明変数

身体活動指標	総消費カロリー
	運動消費カロリー
	歩数
	登った階数
	静止状態の時間
	ややアクティブな時間
	かなりアクティブな時間
睡眠指標	とてもアクティブな時間
	レム睡眠の時間
	深い睡眠の時間
	睡眠効率
生理指標	睡眠断片化
	安静時心拍数

2) 重回帰分析における説明変数の選択手法

表1にFitbit Senseから取得したデジタルバイオマーカーの一覧を記載したが、サンプルサイズに対して説明変数の数が多いことから、類似変数をまとめたり、先行研究を参考にメンタルヘルスと関連の低いと考えられる変数や、継続測定や計算アルゴリズムの問題で欠損の多い変数は削除したりすることで説明変数を以下の13個に絞った（表2）。具体的には、身体活動指標の「移動距離」は、歩数×歩幅により算出されているため「歩数」のみにまとめ、睡眠指標の「就寝時間、起床時間、睡眠時間、ベッドにいた時間、目覚めた回数、目覚めていた時間」は、先行研究で扱われることが多かった「睡眠効率、睡眠断片化」という指標に代表させた（先行研究に沿って睡眠効率は睡眠時間/ベッドにいた時間より算出し⁷⁾、睡眠断片化は目覚めた回数/睡眠時間より算出した²⁴⁾）。また、先行研究では、「浅い睡眠の時間」より「深い睡眠の時間」とメンタルヘルスとの関連を報告していることが多かったことから^{32,36)}、「深い睡眠の時間」のみを選択した。加えて、メンタルヘルスとの関連が指摘されている生理指標は安静時心拍数、心拍変動、皮膚電気活動が中心であったが、心拍変動はデータの欠損が多かったため採用せず、皮膚電気活動はその瞬間の値を測定するスポットチェックのみの機能であったため、本研究が焦点をあてている継続的データの取得には適わないと考え採用しなかった。

続いて、Akaike's Information Criterion (AIC) を最小にするという基準で説明変数を絞り込むステップワイズ法

(変数増減法)により、重回帰分析に投入する説明変数を統計的に上記13個から一定数に絞り込んだ。絞り込まれた変数すべてを説明変数として重回帰分析を行い、その結果有意でなかった説明変数は除去することで、最終的な重回帰式を得た。

3) 目的2における解析の概観

上記の解析1でメンタルヘルス不調の予測に寄与する可能性が明らかとなったデジタルバイオマーカーに対して、それと関連する主観的指標を抽出し解析を行った(解析2)。解析2では、説明変数としてデジタルバイオマーカーに主観的指標を加えた重回帰分析を行い、デジタルバイオマーカーと主観的指標それぞれが目的変数に有意な影響を与えているかを検証する。

4) 目的3における解析の概観

上記の解析1および2でメンタルヘルス不調の予測に寄与することが明らかとなったデジタルバイオマーカーおよび主観的指標を用いて、DASS21のカットオフ値をもとにした重症度のカテゴリー移行(悪化側)を予測できる可能性があるのかを検証するための解析を行った(解析3)。解析3では、w1(ベースライン)におけるそれぞれの目的変数(抑うつ、不安、ストレス)のカテゴリーが悪化側に移行した群とそれ以外の群(移行しなかったまたは改善側に移行した)をそれぞれダミー変数1と0として目的変数に、解析1および2で特定されたデジタルバイオマーカーおよび主観的指標を説明変数としてロジスティック重回帰分析を行った。

II. 結 果

1. 研究協力者の属性

研究協力者の属性として、年齢、性別、現在の立場・身分(学生または社会人)、居住形態(同居または独居)、飲酒の有無、喫煙の有無、運動習慣の有無、BMI、親の教育年数を表3に示した。

2. 統制変数の検討

重回帰分析の目的変数となるw4の抑うつ、不安、ストレスにおいて、デモグラフィック指標(性別、現在の立場・身分、居住形態、飲酒の有無、喫煙の有無、運動習慣の有無、BMI、親の教育年数)による有意な差や関連があるかを検討した(N=80)。その結果を表4および表5に示した。この分析により有意な差や関連がみられた変数は重

表 3 研究協力者の属性

属性 (N=88)	平均 (標準偏差) または n (%)
年齢 [平均 (標準偏差)]	23.02 (2.66)
性別 (n (%))	
男性	43 (48.9)
女性	45 (51.1)
立場・身分 (n (%))	
学生	59 (67.0)
社会人	29 (33.0)
居住形態 (n (%))	
同居	44 (50.0)
独居	44 (50.0)
飲酒の有無 (n (%))	
する	46 (52.3)
しない	42 (47.7)
喫煙の有無 (n (%))	
する	4 (4.5)
しない	84 (95.5)
運動習慣の有無 (n (%))	
あり	58 (65.9)
なし	30 (34.1)
BMI [平均 (標準偏差)]	21.05 (3.11)
親の教育年数 [平均 (標準偏差)]	15.44 (1.72)

重回帰分析における統制変数として用いた。

シャピロ・ウィルク検定により、すべての目的変数で正規性が棄却された〔抑うつ (w4) $W=0.87$, $P<0.001$; 不安 (w4) $W=0.77$, $P<0.001$; ストレス (w4) $W=0.83$, $P<0.001$ 〕。そのため、性別、現在の立場・身分、居住形態、飲酒の有無、喫煙の有無、運動習慣の有無に関しては、ノンパラメトリック検定のウィルコクソンの順位和検定を用いて有意差を検討した。BMIと親の教育年数に関しては、ノンパラメトリック検定の相関分析(スピアマンの順位相関係数)を用いて有意な関連があるかどうかを検討した。その結果、w4の抑うつにおいて居住形態による有意差があり($W=573$, $P<0.05$)、独居のw4抑うつ得点(6.65点)のほうが同居のw4抑うつ得点(4.51点)より有意に高かった。その他のデモグラフィック指標における統計的に有意な差異は確認されなかった。

3. 解析1の結果

目的変数の経時的な平均値の推移を表6に示し、解析1における結果全体を表7に示した。まず、w2とw3の2週間における各デジタルバイオマーカーの1日平均を説明変数、w4の抑うつを目的変数、w1の抑うつと居住形態を統

表4 統制変数ごとの目的変数の平均値（標準偏差）

	男性 (N=37)	女性 (N=43)	P 値
抑うつ	5.68 (6.17)	5.35 (4.78)	0.71
不安	2.54 (3.35)	2.60 (3.48)	0.97
ストレス	4.98 (5.90)	5.67 (6.25)	0.53
	学生 (N=54)	社会人 (N=26)	P 値
抑うつ	5.81 (5.46)	4.85 (5.43)	0.36
不安	2.89 (3.86)	1.92 (2.08)	0.66
ストレス	5.70 (6.40)	4.62 (5.33)	0.55
	同居 (N=37)	独居 (N=43)	P 値
抑うつ	4.51 (5.38)	6.65 (5.33)	0.03*
不安	2.32 (3.38)	2.86 (3.45)	0.33
ストレス	4.51 (5.64)	6.32 (6.46)	0.15
	飲酒する (N=40)	しない (N=40)	P 値
抑うつ	5.55 (5.77)	5.45 (5.14)	0.79
不安	2.15 (2.81)	3.00 (3.90)	0.40
ストレス	5.20 (6.07)	5.50 (6.12)	0.87
	喫煙する (N=4)	しない (N=76)	P 値
抑うつ	5.50 (8.54)	5.50 (5.31)	0.65
不安	0.50 (1.00)	2.68 (3.45)	0.17
ストレス	3.00 (2.00)	5.47 (6.18)	0.67
	運動習慣あり (N=50)	なし (N=30)	P 値
抑うつ	5.64 (5.57)	5.27 (5.29)	0.86
不安	2.52 (3.38)	2.67 (3.50)	0.92
ストレス	5.48 (6.04)	5.13 (6.19)	0.69

* $P<0.05$

制変数として、ステップワイズ法による重回帰分析を行った。その結果、統制変数である w1 の抑うつ ($\beta=0.66$, $P<0.001$)、睡眠指標であるレム睡眠の時間 ($\beta=0.29$, $P<0.01$)、睡眠指標である深い睡眠の時間 ($\beta=-0.27$, $P<0.01$) が有意となり、モデルも有意となった (adj $R^2=0.51$, $P<0.001$)。多重共線性の問題はみられなかった ($VIF<2$)。また、身体活動指標および生理指標はいずれも有意にならなかった。

次に、w2 と w3 の 2 週間における各デジタルバイオマーカーの 1 日平均を説明変数、w4 の不安を目的変数、w1 の不安を統制変数として、ステップワイズ法による重回帰分析を行った。その結果、統制変数である w1 の不安 ($\beta=0.49$, $P<0.001$)、睡眠指標であるレム睡眠の時間 ($\beta=0.31$, $P<0.01$)、睡眠指標である深い睡眠の時間 ($\beta=-0.26$, $P<0.05$) が有意となり、モデルも有意となった (adj $R^2=0.31$, $P<0.001$)。多重共線性の問題はみられなかった ($VIF<2$)。また、身体活動指標および生理指標はいずれも有意にならなかった。

表5 統制変数ごとの目的変数との相関係数

BMI	抑うつ	0.04
	不安	-0.06
	ストレス	0.00
親の教育年数	抑うつ	0.05
	不安	-0.07
	ストレス	-0.06

表6 目的変数の経時的な平均値の推移

	w1	w2	w3	w4
抑うつ	5.80 (6.58)	6.30 (6.59)	5.68 (6.02)	5.50 (5.43)
不安	3.25 (3.51)	3.13 (3.84)	3.23 (5.41)	2.58 (3.40)
ストレス	6.58 (6.02)	6.83 (7.24)	5.53 (6.34)	5.35 (6.06)

() 内は標準偏差。

次に、w2 と w3 の 2 週間における各デジタルバイオマーカーの 1 日平均を説明変数、w4 のストレスを目的変数、w1 のストレスを統制変数として、ステップワイズ法による重回帰分析を行った。その結果、統制変数である w1 のストレス ($\beta=0.71$, $P<0.001$) が有意、睡眠指標であるレム睡眠の時間 ($\beta=0.14$, $P<0.10$) が有意傾向となり、モデルも有意となった (adj $R^2=0.51$, $P<0.001$)。多重共線性の問題はみられなかった ($VIF<2$)。また、身体活動指標および生理指標はいずれも有意にならなかった。

4. 解析2の結果

解析1の結果、w4 の抑うつ、w4 の不安それぞれを目的変数とした重回帰分析において w2 と w3 の 2 週間における「レム睡眠の時間」と「深い睡眠の時間」が統計学的に有意な影響があり、w4 のストレスを目的変数とした重回帰分析においては w2 と w3 の 2 週間における「レム睡眠の時間」が有意傾向ながら影響があることが明らかとなった。このことから、主観的な睡眠の質を計測する「PSQI の総合得点」を説明変数として新たに投入し、再度上記3つの重回帰分析を行った。

解析2における結果全体を表8に示した。まず、w4 の抑うつを目的変数、w2 と w3 の 2 週間における「レム睡眠の時間」と「深い睡眠の時間」の 1 日平均および w1~w4 の 4 週間における平均的な「主観的睡眠の質 (PSQI 総合得点)」を説明変数、w1 の抑うつを統制変数として重回帰

表 7 ステップワイズ法による重回帰分析の結果（解析 1）

目的変数（w4 の抑うつ）			目的変数（w4 の不安）			目的変数（w4 のストレス）		
説明変数	標準偏回帰係数 (β)	VIF	説明変数	標準偏回帰係数 (β)	VIF	説明変数	標準偏回帰係数 (β)	VIF
w1 の抑うつ 睡眠指標	0.66***	1.01	w1 の不安 睡眠指標	0.49***	1.02	w1 のストレス 睡眠指標	0.71***	1.00
レム睡眠の時間	0.29**	1.41	レム睡眠の時間	0.31**	1.42	レム睡眠の時間	0.14 [†]	1.00
深い睡眠の時間	-0.27**	1.40	深い睡眠の時間	-0.26*	1.40			
決定係数 (adjR ²)	0.51***		決定係数 (adjR ²)	0.31***		決定係数 (adjR ²)	0.51***	

* $P<0.05$, ** $P<0.01$, *** $P<0.001$, [†] $P<0.10$

表 8 重回帰分析の結果（解析 2）

目的変数（w4 の抑うつ）			目的変数（w4 の不安）			目的変数（w4 のストレス）		
説明変数	標準偏回帰係数 (β)	VIF	説明変数	標準偏回帰係数 (β)	VIF	説明変数	標準偏回帰係数 (β)	VIF
w1 の抑うつ 睡眠指標	0.62***	1.51	w1 の不安 睡眠指標	0.41***	1.38	w1 のストレス 睡眠指標	0.66***	1.00
レム睡眠の時間	0.29**	1.41	レム睡眠の時間	0.31**	1.42	レム睡眠の時間	0.15 [†]	1.00
深い睡眠の時間	-0.26**	1.42	深い睡眠の時間	-0.23*	1.43			
主観的睡眠の質 (PSQI 総合得点)	0.07	1.53	主観的睡眠の質 (PSQI 総合得点)	0.16	1.37	主観的睡眠の質 (PSQI 総合得点)	0.09	
決定係数 (adjR ²)	0.51***		決定係数 (adjR ²)	0.32***		決定係数 (adjR ²)	0.51***	

* $P<0.05$, ** $P<0.01$, *** $P<0.001$, [†] $P<0.10$

分析を行った。その結果、統制変数である w1 の抑うつ ($\beta=0.62$, $P<0.001$), レム睡眠の時間 ($\beta=0.29$, $P<0.01$), 深い睡眠の時間 ($\beta=-0.26$, $P<0.01$) が有意となり、モデルも有意となった (adj R²=0.51, $P<0.001$)。多重共線性の問題はみられなかった (VIF<2)。また、主観的睡眠の質は有意にならなかった。

次に、w4 の不安を目的変数、w2 と w3 の 2 週間における「レム睡眠の時間」と「深い睡眠の時間」の 1 日平均および w1~w4 の 4 週間における平均的な「主観的睡眠の質 (PSQI 総合得点)」を説明変数、w1 の不安を統制変数として重回帰分析を行った。その結果、統制変数である w1 の不安 ($\beta=0.41$, $P<0.001$), レム睡眠の時間 ($\beta=0.31$, $P<0.01$), 深い睡眠の時間 ($\beta=-0.23$, $P<0.05$) が有意となり、モデルも有意となった (adj R²=0.32, $P<0.001$)。多重共線性の問題はみられなかった (VIF<2)。また、主観的睡眠の質は有意にならなかった。

次に、w4 のストレスを目的変数、w2 と w3 の 2 週間における「レム睡眠の時間」の 1 日平均および w1~w4 の 4 週間における平均的な「主観的睡眠の質 (PSQI 総合得点)」を説明変数、w1 のストレスを統制変数として重回帰

分析を行った。その結果、統制変数である w1 のストレス ($\beta=0.66$, $P<0.001$) が有意、レム睡眠の時間 ($\beta=0.15$, $P<0.10$) が有意傾向となり、モデルも有意となった (adj R²=0.51, $P<0.001$)。多重共線性の問題はみられなかった (VIF<2)。また、主観的睡眠の質は有意にならなかった。

5. 解析 3 の結果

解析 1 の結果、w4 の抑うつ、w4 の不安それぞれを目的変数とした重回帰分析において「レム睡眠の時間」と「深い睡眠の時間」が統計学的に有意な影響があり、w4 のストレスを目的変数とした重回帰分析においては「レム睡眠」が有意傾向ながら影響があることが明らかとなった。続けて、解析 2 を行ったところ、いずれの目的変数に対しても主観的睡眠指標は有意な影響がみられなかった。以上の結果をふまえて、解析 3 では、w1 におけるそれぞれの目的変数のカテゴリーが w4 において悪化側に移行した群とそれ以外の群（移行しなかったまたは改善側に移行した）をそれぞれダミー変数 1 と 0 として目的変数に、解析 1 で特定されたデジタルバイオマーカー（「レム睡眠の時

表9 ロジスティック回帰分析の結果（解析3）

目的変数（w4の抑うつ）			目的変数（w4の不安）			目的変数（w4のストレス）		
説明変数	オッズ比	95%信頼区間	説明変数	オッズ比	95%信頼区間	説明変数	オッズ比	95%信頼区間
レム睡眠の時間	1.04	1.00～1.09	レム睡眠の時間	1.04	0.99～1.09	レム睡眠の時間	1.04	0.99～1.10
深い睡眠の時間	1.05	0.99～1.11	深い睡眠の時間	1.02	0.94～1.12			

間」と「深い睡眠の時間」を説明変数としてロジスティック回帰分析を行った（表9）。w4において抑うつのカテゴリーが悪化側に移行した群は11名、不安のカテゴリーが悪化側に移行した群は6名、ストレスのカテゴリーが悪化側に移行した群は4名であった。ロジスティック回帰分析の結果、w4の抑うつのカテゴリー移行（悪化側）に関して、「レム睡眠の時間」の増加が有意に影響していた（オッズ比1.04、95%信頼区間1.00～1.09）。

III. 考 察

本研究全体をまとめると、若年層の健常者および閾下症状をもつ者を対象として、Fitbit Sense から得られたデジタルバイオマーカーを説明変数とし、DASS21の抑うつ、不安、ストレスを目的変数とするステップワイズ法による重回帰分析を行った結果、直前2週間におけるデジタルバイオマーカーの変化（Fitbit Sense から推算されたレム睡眠の時間の増加と深い睡眠の時間の減少）がその直後のメンタルヘルス不調の予測に資する可能性が示された。また、それは主観的評価では予測できず、デジタルバイオマーカーによってのみ予測ができる可能性を示唆する結果であった。なお、本結果は目的変数のベースラインおよび目的変数と有意な関連があったデモグラフィック指標の値を統制して得られた統計学的により頑健な示唆である。加えて、Fitbit Sense から推算されたレム睡眠の時間の増加に関しては、抑うつ得点の上昇のみならず、悪化側のカテゴリー移行の予測にも寄与する可能性が明らかになった。解析1ではFitbit Sense から推算されたレム睡眠の時間の増加が抑うつ得点の上昇を予測していることを明らかにしたが、わずかな点数の上昇が悪化を示しているとは厳密には言い難い。そのため、カットオフ値をもとに重症度のカテゴリーが悪化側に移行することを予測する可能性を示すことにより臨床的意義があると考えられる。つまり、解析3の結果は、抑うつ悪化の早期発見に向けたスマートウォッチの応用可能性を見出したといえるだろう。また、

不安、ストレスに関してデジタルバイオマーカーがカテゴリー移行に有意な関連を見いだせなかった要因としては、不安、ストレスは悪化側にカテゴリー移行した人数が本研究では少なかったことが考えられる。

レム睡眠の時間の増加とメンタルヘルス不調との関連は多くの研究で報告されている。例えば、うつ病に罹患するとレム睡眠の時間が増加することが明らかにされたり^{3,27)} 全般性不安障害の子どもが健常の子どもに比べてレム睡眠の時間が長いことが示されている²⁾。また、日常的な軽度のストレスがレム睡眠の時間の増加をもたらすことも報告されている²⁶⁾。レム睡眠中は、身体は休んでいる一方脳は活発に活動し交感神経優位になっていることが明らかになっており²⁵⁾、Fitbit Sense から推算された直近の「レム睡眠の時間の増加」がメンタルヘルス不調を予測するという本研究の結果は理論的にも矛盾しないと考えられる。

深い睡眠の時間の減少とメンタルヘルス不調との関連についても多くの研究で報告されている。例えば、うつ病に罹患している自殺未遂経験のある青年の睡眠ステージを分析し、睡眠時間の最後の3分の1で深い睡眠の割合が減少していることを報告している研究など⁵⁾、高レベルの不安により深い睡眠が減少することを示す研究が存在している¹⁷⁾。また、絶え間ない過度のストレスにより発生すると考えられている燃え尽き症候群の人は深い睡眠が減少しているという報告もある⁸⁾。深い睡眠中は、身体も脳も休息し副交感神経優位になっていることが明らかになっており¹⁾、Fitbit Sense から推算された直近の「深い睡眠の時間の減少」がメンタルヘルス不調の予測に資する可能性を示したという本研究の結果に関しても理論的に矛盾しないと考えられる。

続いて、身体活動指標に関して、解析1の結果から、いずれの変数もメンタルヘルス不調の経時的変化に有意な影響を及ぼしていないことが示された。しかし、うつ病や抑うつ状態の症状の1つとして、興味や喜びの喪失、易疲労性、意欲の低下などに起因する身体活動量の低下がみられることは広く知られていることである。うつ病と家にもこも

りがちなライフスタイルの有意な相関関係を示した報告がある¹⁵⁾。また、実際に、スマートフォンのセンサーを用いて身体活動量を計測し抑うつ症状との関連を調べた研究も複数存在する。例えば、大学生 83 名を対象に 9 週間の追跡調査を行い、抑うつ症状が重度な人ほど静止時間が長く、訪問する場所の数が少ないことを報告している研究がある³⁵⁾。本研究においても、身体活動指標のいずれかの変数がメンタルヘルス不調の経時的変化に有意な影響を与えるだろうとの予想をしていたが異なる結果となった。その要因の 1 つとして、まず、デジタルバイオマーカーの変動の期間設定が影響した可能性が考えられる。本研究では、2 週間におけるデジタルバイオマーカーの変動を予測の期間として設定したが、3 週間、4 週間、あるいはそれ以上の期間であれば身体活動量の低下がメンタルヘルス不調を予測する可能性は十分考えられる。

次に、本研究の対象が健常者や閾下症状をもつ者であり、身体活動を要する生活が特に問題なく成り立っていたことの影響もあるかもしれない。臨床群や精神疾患既往歴のある人を対象とした場合であれば、2 週間の身体活動でも不調を捉えることができる可能性はある。さらに、本研究が外出自粛や行動制限のあったコロナ禍で行われたことにも注意をする必要がある。日本国内で COVID-19 の感染拡大前後における 1 週間あたりの身体活動時間は約 60 分（約 3 割）減少していたことが報告されている³⁸⁾。本研究のデータが十分なばらつきを反映できていたかは留意しなければならない。

生理指標の安静時心拍数に関しても、解析 1 の結果からメンタルヘルス不調の予測に有意な影響を及ぼしていなかった。安静時心拍数は自律神経によって支配され、交感神経の影響を受けて上昇し、副交感神経の影響を受けて低下する。そのため、交感神経を優位にする不安やストレスとの関連が推測されたが異なる結果となった。その要因として時間的な前後関係の影響があるのかもしれない。つまり、不安やストレスが安静時心拍数に影響を与えるという方向性での関連はあるが、安静時心拍数が不安やストレスに影響を与えるという方向性での関連はないのかもしれない。つまり、安静時心拍数を不安やストレスの悪化に対して予測に資する因子として用いることには大きな意義がない可能性が考えられる。他の要因としては、本研究で用いたスマートウォッチの測定精度の影響が推測される。序論でも述べたように、Fitbit による心拍の測定精度に関しては研究者の間で賛否が分かれており、測定精度の影響がな

かったとは言い切れない。今後の技術的な向上が望まれる。また、心拍間隔の周期的な変動である「心拍変動」に関しては今後検討の余地がある。本研究では、データの欠損が多く指標として採用しなかったが、メンタルヘルス研究では、心拍数より心拍変動に注目しているケースが多く、有効な予測因子になる可能性は考えられる。

1. 本研究の限界と今後の課題

1) 結果の一般化可能性

本研究では機縁法およびソーシャルネットワークサービスの 1 つの Twitter（現在の X）を使用して研究協力者を募集した。機縁法は同質の集団をサンプリングしてしまう可能性があり、Twitter に関しても研究募集にたどり着くという点で何かしらのバイアスが生じている可能性は否定できない。より厳密な結果を得るために、今後はランダムサンプリングによる再検証が求められる。

2) 対象者の属性

本研究では 18~29 歳の若年層を対象として睡眠の質に関する有益な結果を得た。しかし、本研究の結果はあくまで若年層における結果であることに留意しなければならない。例えば、深い睡眠の時間は加齢とともに減少することが指摘されており²⁰⁾、睡眠は年齢の影響を大きく受けることが知られている。デジタルバイオマーカーを効果的に社会実装していくために、年齢の影響を詳細に考慮したさらなる研究が求められる。

また、本研究は健常者および閾下症状をもつ者を対象としているが、「現在精神疾患の治療中ではない」という質問によるオンライン上の自己申告をもとに選別している。厳密には Structured Interview for DSM-V (SCID) やミニメンタルステート検査 (Mini-Mental state : MINI) などの精神医学的評価をしなければならない。さらに、解析 1 のステップワイズ法による重回帰分析の結果のみでは回帰係数含め臨床的意義をもちにくいことをふまえ、本研究では解析 3 で重症度のカテゴリーが悪化側に移行することの予測まで行った。しかし、健常者および閾下症状をもつ者が発症という状態なるまで、そしてその予後まで追ったわけではないため、臨床的意義の小さい変動を予測したに過ぎない可能性は否定できない。今後の研究において、健常者および閾下症状をもつ者を対象とする場合より長期での縦断調査が望まれる。

3) 研究デザイン

第一に、本研究領域は萌芽期にあり、欧米の先行研究の

多くは数十名程度の小さなサンプルサイズにとどまっていた。本研究の解析1および解析2では、先行研究に比べると比較的大きなサンプルサイズによる検証を行うことができたが十分とはいえず、解析3のサンプルサイズは非常に小さい。加えて、サンプルサイズが不十分なことにより、本研究ではメンタルヘルス不調と有意な関連がみられなかった指標も存在する可能性がある。そのため、より大規模なサンプルサイズでの再検証によって本研究と同様の結果が得られるかどうか確かめる必要があるだろう。本研究は萌芽期研究として予測に資する変数を探索したことにとどまったが、今後大規模研究が蓄積し、予測する変数およびアウトカムのカテゴリーがより明確になることで、感度、特異度、予測精度などの評価も可能となるだろう。大規模研究を行ううえで研究者自身がデバイスを大量に準備するには多額の費用がかかるため一定の限界はあるが、本研究のように市販のデバイスを用いるのであれば、すでにデバイスを所有している人を募集にかけることで十分なサンプルサイズを実現できるかもしれない。その際はサンプリングバイアスに留意する必要がある。

第二に、解析2に関して、w1~w4の4週間を測定期間としたPSQI総合得点を説明変数として投入しているが、解析に用いたデジタルバイオマーカーの期間はw2とw3の2週間であるため、完全に重なり合っていないことは限界点である。

第三に、本研究はメンタルヘルスの主観的評価の課題を乗り越えるために、デジタルバイオマーカーという客観的指標を導入しその有効性を検証したものであるが、現状確立した客観的評価手法が存在しないため、主観的指標を使用せざるを得なかった。この点は本研究の限界であるが、ほとんどの先行研究も同様の手法を用いており、萌芽期の研究手法としては典型的である。ただし、コストやリソースの問題はあるが、精神科医などによる面接評価を行えば現状においてもより精度を高めることはできるだろう。

第四に、スマートウォッチの装着による結果への影響について、実験中に皮膚接触アレルギーが生じた被験者は除外することでその影響を低減したが、それ以外の装着によるストレスや健康不安の喚起などが結果に与える影響に関しては十分考慮ができていない。コストやリソースの問題はあるが、研究開始1ヵ月のデータは用いないなど、研究デザインを工夫することで装着の影響をより厳密に低減させることができるだろう。

4) デバイス技術

本研究で用いたスマートウォッチ (Fitbit Sence) は、これまで指摘されてきた防水性や充電の課題を克服し、ユーザーにとってより使用しやすいデバイスであったといえる。一方、上でも述べたようにFitbitの測定精度に関しては議論の余地が残るため、本研究結果は推測の域を出ないことは大きな限界点である。測定精度のさらなる改善は今後期待されるところである。

加えて、今後の展望として早期発見から介入までを1つのデジタルデバイス上で行うことができると理想的だろう。つまり、デジタルバイオマーカーの変動によってメンタルヘルス不調を検知し、リアルタイムに本人へフィードバックすることでマインドフルネス瞑想や種々のリラクゼーション法を促すような仕組みである。心拍の変動からストレスを検知し、緩やかなテンポのバイブレーションを送ることで不安を低減させる介入まで行っている研究がすでに存在する¹⁰⁾。さらには、そのようなフィードバック介入で実際にリラクゼーション効果が得られたかをデジタルバイオマーカーで検知し、さらにフィードバックする仕組みまで実装できるとより有益である可能性もある。

5) 人工知能・機械学習の活用と異分野専門家の協働

本研究で目的としたような早期発見を可能にするためにはデジタルバイオマーカーによりわずかな変化を検出することが求められる。一方で、臨床的に意義のあるわずかな変化を感度特異度高く検出することの難しさは、他の医学領域における検査の現状をみても自明である。そのような場面で活用が期待されるのが人工知能や機械学習であろう。放射線科における画像診断の領域をはじめとして、人工知能や機械学習の医学への導入は著しく進歩しており、本研究領域でもサポートベクターマシン (Support-Vector Machine : SVM)、ランダムフォレスト (Random Forest : RF)、ナীবベイズ (Naive Bayes : NB)、k近傍法などの機械学習を活用した研究が行われている。客観的評価が特に難しい精神医学の領域において、デジタルバイオマーカーをより実用的なものにしていくために、人工知能や機械学習を取り入れた研究の進歩は欠かせないかもしれない。ただし、本研究で用いたような線形回帰などに比べると特徴量の重要度の判定が困難であることには留意しなければならない。また、こうした高度な解析を行うためには、メンタルヘルスの専門家と統計学者やエンジニア、データサイエンティストなどの異分野の専門家同士の協働が重要だろう。

おわりに

本研究では、若年層の健常者および閾下症状をもつ者を対象として、スマートウォッチから得られたデジタルバイオマーカーを説明変数とし、DASS21の抑うつ、不安、ストレスを目的変数とするステップワイズ法による重回帰分析を行った結果、直前2週間におけるデジタルバイオマーカーの変化（Fitbit Sense から推算されたレム睡眠の時間の増加と深い睡眠の時間の減少）がその直後のメンタルヘルス不調の予測に資する可能性が示された。また、それは主観的評価では予測できず、デジタルバイオマーカーによってのみ予測ができる可能性を示唆していた。なお、本結果は目的変数のベースラインおよび目的変数と有意な関連があったデモグラフィック指標の値を統制して得られた統計学的により頑健な示唆である。加えて、Fitbit Sense から推算されたレム睡眠の時間の増加に関しては、抑うつ得点の上昇のみならず、重症度のカテゴリー移行（悪化側）も予測できる可能性が明らかになった。

なお、本論文に関連して開示すべき利益相反はない。

謝辞

本研究はJSPS 科研費（16H05653, 19K03278, 22H01091 & 22K18582）と、Royal Society and The British Academy（Newton International Fellowship Alumni follow on funding）と、東京大学社会連携講座「“はたらく”に遊びを」と、東京大学卓越研究員制度と、東京大学大学院教育学研究科附属海洋教育センター・海洋教育基盤研究プロジェクトの助成を受けて行われた。

文献

- 1) Ako, M., Kawara, T., Uchida, S., et al. : Correlation between electroencephalography and heart rate variability during sleep. *Psychiatry Clin Neurosci*, 57 (1) ; 59-65, 2003
- 2) Alfano, C. A., Reynolds, K., Scott, N., et al. : Polysomnographic sleep patterns of non-depressed, non-medicated children with generalized anxiety disorder. *J Affect Disord*, 147 (1-3) ; 379-384, 2013
- 3) Benca, R. M. : Sleep in psychiatric disorders. *Neurol Clin*, 14 (4) ; 739-764, 1996
- 4) Benedetto, S., Caldato, C., Bazzan, E., et al. : Assessment of the Fitbit Charge 2 for monitoring heart rate. *PLoS One*, 13 (2) ; e0192691, 2018
- 5) Boaf, A., Armitage, R., Greenham, S., et al. : Sleep architecture in adolescents hospitalized during a suicidal crisis. *Sleep Med*, 56 ; 41-46, 2019
- 6) Buysse, D. J., Reynolds, C. F. 3rd., Monk, T. H., et al. : The Pittsburgh Sleep Quality Index : a new instrument for psychiatric practice and research. *Psychiatry Res*, 28 (2) ; 193-213, 1989
- 7) Cacioppo, J. T., Hawkley, L. C., Berntson, G. G., et al. : Do lonely days invade the nights? Potential social modulation of sleep efficiency. *Psychol Sci*, 13 (4) ; 384-387, 2002
- 8) Cheeta, S., Ruigt, G., van Proosdij, J., et al. : Changes in sleep architecture following chronic mild stress. *Biol Psychiatry*, 41 (4) ; 419-427, 1997
- 9) Chevanec, G., Golaszewski, N. M., Tipton, E., et al. : Accuracy and precision of energy expenditure, heart rate, and steps measured by combined-sensing Fitbits against reference measures : systematic review and meta-analysis. *JMIR Mhealth Uhealth*, 10 (4) ; e35626, 2022
- 10) Costa, J., Adams, A. T., Jung, M. F., et al. : EmotionCheck : leveraging bodily signals and false feedback to regulate our emotions. *UbiComp '16 : Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. Association for Computing Machinery, New York, p.758-769, 2016
- 11) 土井由利子, 簗輪真澄, 大川匡子ほか : ピッツバーグ睡眠質問票日本語版の作成. *精神科治療学*, 13 (6) ; 755-769, 1998
- 12) Fitbit Inc. : Fitbit Sense マニュアル. 2022
- 13) Fitbit Inc. : ホームページ. 2022 (<https://myhelp.fitbit.com/s/home?language=ja>) (参照 2022-12-15)
- 14) Germini, F., Noronha, N., Borg Debono, V., et al. : Accuracy and acceptability of wrist-wearable activity-tracking devices : systematic review of the literature. *J Med Internet Res*, 24 (1) ; e30791, 2022
- 15) Giuntella, O., Hyde, K., Saccardo, S., et al. : Lifestyle and mental health disruptions during COVID-19. *Proc Natl Acad Sci U S A*, 118 (9) ; e2016632118, 2021
- 16) Haghighyeh, S., Khoshnevis, S., Smolensky, M. H., et al. : Accuracy of wristband Fitbit models in assessing sleep : systematic review and meta-analysis. *J Med Internet Res*, 21 (11) ; e16273, 2019
- 17) Horváth, A., Montana, X., Lanquart, J. P., et al. : Effects of state and trait anxiety on sleep structure : a polysomnographic study in 1083 subjects. *Psychiatry Res*, 244 ; 279-283, 2016
- 18) Huckvale, K., Venkatesh, S., Christensen, H. : Toward clinical digital phenotyping : a timely opportunity to consider purpose, quality, and safety. *NPJ Digit Med*, 2 (1) ; 88, 2019
- 19) 厚生労働省 : メッツ/METs. 健康用語辞典. e-ヘルスネット. 2021 (<https://kenet.mhlw.go.jp/information/information/dictionary/exercise/ys-004>) (参照 2022-12-15)
- 20) 厚生労働省 : ノンレム睡眠. 健康用語辞典. e-ヘルスネット. 2021 (<https://kenet.mhlw.go.jp/information/information/dictionary/heart/yk-048>) (参照 2022-12-15)
- 21) Liang, Y., Zheng, X., Zeng, D. D. : A survey on big data-driven digital phenotyping of mental health. *Inf Fusion*, 52 ; 290-307, 2019
- 22) Lovibond, S. H., Lovibond, P. F. : Manual for the Depression Anxiety & Stress Scales, 2nd ed. Psychology Foundation, Sydney, 1995
- 23) 三谷智子, 村上由希, 今村行雄 : DASS (depression, anxiety, stress scale) 21 日本語版の作成と妥当性の検証. *日本公衆衛生*

- 24) Morrell, M. J., Finn, L., Kim, H., et al. : Sleep fragmentation, awake blood pressure, and sleep-disordered breathing in a population-based study. *Am J Respir Crit Care Med*, 162 (6) ; 2091-2096, 2000
- 25) Narkiewicz, K., Somers, V. K. : The sympathetic nervous system and obstructive sleep apnea : implications for hypertension. *J Hypertens*, 15 (12 Pt 2) ; 1613-1619, 1997
- 26) Nollet, M., Hicks, H., McCarthy, A. P., et al. : REM sleep's unique associations with corticosterone regulation, apoptotic pathways, and behavior in chronic stress in mice. *Proc Natl Acad Sci USA*, 116 (7) ; 2733-2742, 2019
- 27) Nutt, D., Wilson, S., Paterson, L. : Sleep disorders as core symptoms of depression. *Dialogues Clin Neurosci*, 10 (3) ; 329-336, 2008
- 28) Paulhus, D. L. : Socially desirable responding on self-reports. *Encyclopedia of Personality and Individual Differences* (ed by Zeigler-Hill, V., Shackelford, T.K.). Springer, Cham, 2017
- 29) Piau, A., Wild, K., Mattek, N., et al. : Current state of digital biomarker technologies for real-life, home-based monitoring of cognitive function for mild cognitive impairment to mild Alzheimer disease and implications for clinical care : systematic review. *J Med Internet Res*, 21 (8) ; e12785, 2019
- 30) Sariyska, R., Rathner, E. M., Baumeister, H., et al. : Feasibility of linking molecular genetic markers to real-world social network size tracked on smartphones. *Front Neurosci*, 12 ; 945, 2018
- 31) Sobolev, M., Gullapalli, B. T., Rahman, T. : Advancing the science of digital biomarkers. *Proceedings of the 2021 Workshop on Future of Digital Biomarkers (DigiBiom '21)*. Association for Computing Machinery, New York, p.1-2, 2021
- 32) Steiger, A., Pawlowski, M. : Depression and sleep. *Int J Mol Sci*, 20 (3) ; 607, 2019
- 33) Stone, A. A., Shiffman, S. : Capturing momentary, self-report data : a proposal for reporting guidelines. *Ann Behav Med*, 24 (3) ; 236-243, 2002
- 34) Torous, J., Kiang, M. V., Lorme, J., et al. : New tools for new research in psychiatry : a scalable and customizable platform to empower data driven smartphone research. *JMIR Ment Health*, 3 (2) ; e16, 2016
- 35) Wang, R., Wang, W., DaSilva, A., et al. : Tracking depression dynamics in college students using mobile phone and wearable sensing. *Proc ACM Interact Mob Wearable Ubiquitous Technol*, 2 (1) ; 43, 2018
- 36) Wichniak, A., Wierzbicka, A., Jernajczyk, W. : Sleep as a biomarker for depression. *Int Rev Psychiatry*, 25 (5) ; 632-645, 2013
- 37) Wright, S. P., Hall Brown, T. S. H., Collier, S. R., et al. : How consumer physical activity monitors could transform human physiology research. *Am J Physiol Regul Integr Comp Physiol*, 312 (3) ; R358-367, 2017
- 38) Yamada, M., Kimura, Y., Ishiyama, D., et al. : Effect of the COVID-19 epidemic on physical activity in community-dwelling older adults in Japan : a cross-sectional online survey. *J Nutr Health Aging*, 24 (9) ; 948-950, 2020

A Study for Prediction of Mental Health Problems Through the Use of Digital Biomarkers : A Smartwatch Longitudinal Study

Fumiya TAKAHASHI¹⁾, Rina HASHIMOTO¹⁾, Ryu TAKIZAWA^{1,2)}

1) Department of Clinical Psychology, Graduate School of Education, The University of Tokyo

2) Institute of Psychiatry, Psychology and Neuroscience, King's College London

Current mental health assessments are mainly based on self-reports using questionnaires, but it has been pointed out that they are susceptible to biases such as recall bias and lack of accuracy. In addition, while daily monitoring is essential for the early detection of mental health problems, answering daily questionnaires is a major burden in terms of time and effort. Against these backgrounds, digital biomarkers have been attracting attention in recent years. Digital biomarkers are suitable for early detection because they can automatically and passively acquire data simply by wearing a digital device, and they also enable continuous and long-term measurement of real-time data obtained in daily life.

In this study, the objective of this study was to explain depression, anxiety, and stress based on the self-reported questionnaire, Depression, Anxiety, Stress scale 21 (DASS21) using physical activity indices, sleep indices, and physiological indices calculated by Fitbit Sense. The subjects were 88 healthy individuals aged 18 to 29 years and those with subthreshold symptoms, who wore the smartwatch constantly for 4 weeks and answered a web-based questionnaire (DASS21) every week while wearing the watch.

Stepwise multiple regression analysis indicated that the changes in digital biomarkers (increased time in REM sleep and decreased time in deep sleep calculated by Fitbit Sense) in the preceding 2 weeks might predict mental health problems in the following week. The results also suggested that this could not be predicted by the subjective assessment, but only by digital biomarkers. Note that these are statistically more robust findings obtained by controlling for the baseline values and the demographic indicators. In addition, the results of this longitudinal study revealed that an increase in the duration of REM sleep calculated by Fitbit Sense might predict not only an increase in depression scores, but also a shift in severity category (toward the worse side). These findings suggest the applicability of digital biomarkers in the field of mental health.

Authors' abstract

Keywords digital mental health, digital biomarkers, smartwatch, early detection, longitudinal study