

令和2年度厚生労働科学研究補助金
(倫理的法的社会的課題研究事業)

分担研究報告書

精神科領域への人工知能技術導入と、その課題

研究分担者 菅原 典夫 (獨協医科大学・精神神経医学講座・准教授)

研究要旨

精神科領域への AI 導入について、その現状と、そこから考えられる課題について、検討を行った。他の医学領域と比較した場合、精神科領域では病因や臨床上の訴えに心理社会的背景が反映されやすく、診断の基準に客観的な要因が組み込まれることが少ないため、AI の導入に対する壁になると考えられる。しかし、これまでの、この領域における AI 関連論文数 2,866 件のうち、実に 2,120 件が 2016 から 2020 年の間に公表されたものであり、まさに現在、飛躍的に増加している。研究の内容としては、これまでも比較的導入が容易とみられていた画像所見による診断学的研究が多いものの、こうした所見を治療反応性の予測にまで拡張した研究や、さらには AI を搭載したロボットやスマートデバイスによる介入についての研究、自然言語処理による診断学的研究など、機械学習の活用は確実にその領域を広げており、我が国からの成果発信も期待できる状況である。ただ、診断学的研究においては臨床的に判別が難しい症例を峻別できるものではないうえ、診断根拠の説明を求められた際の回答など、特に司法精神医学の領域において課題となりえる。また、現在のところ、治療デバイスとしては、患者の人体に対して生物学的介入を行うものは開発されていないため、大きな問題は生じていないものの、将来においては、そうした介入の安全性も課題となるであろう。

A. 研究目的

人工知能 (AI; Artificial Intelligence) の利活用はあらゆる分野において進行しており、精神科領域も、その例外ではない。現在においても、精神疾患そのものの機序は解明されておらず、さらには心理社会的背景により症状が変わり得るといったこの領域に特有の事情は、AI の導入に対する壁となっている。しかし、比較的導入が容易とみられていた画像診断以外の領域でも、その利活用に関する研究が進行しており、これまでの手法では扱いきれなかった膨大な情報を解釈することで、診断・治療から予防に至るまでの大きな変革が期待されている。しかし、ヒトの能力を超えるともされる AI を、医療の、それも、ある意味では心を取り扱う臨床の現場に導入することに対する不安も大きく、それに対して、現在のところ明確な回答は存在しない。本研究では、**精神科領域への AI 導入について、その現状と、そこから考えられる課題について、論点を整理することを目的に考察を行う。**

B. 研究方法

1. 精神科領域の特徴および AI について研究の現状と、それについての考察

精神科臨床の視点から、主に文献 (書籍、学術論文) の検討により、精神科領域と他の医学領域との比較検討および、医療現場への AI 導入の現状と、それらについての課題を考察する。

2. 精神科領域における AI 関連論文数の検討

臨床医学文献データベースとして、PubMed を用いた。その検索式として、(artificial intelligence [TW] OR machine learning [TW] OR natural language processing [TW] OR text mining [TW] OR deep learning [TW]) AND psychiatry [ALL] として、令和 3 年 5 月 5 日に検索を実施した。

(倫理面への配慮)

本研究では、基本的に個人情報を取り扱うことはない。

C. 結果および考察

1. 精神科領域と他の診療科領域との比較

ひとくちに精神疾患といっても、その病因として心理社会的背景の寄与が比較的に大きいものと、小さいものがある。¹⁾ 前者の場合、患者の訴えを疾患によるものと捉えるより、主として心理社会的背景によるものと考えた方がよいだろうし、後者の場合は、その訴えを疾患による認知や体験の歪みと理解すべきである。とはいえ、患者の訴えが 100% 疾患によるものと断ずることは実際のところ困難であり、疾患による影響が大と考えられる場合でも、その心理社会的背景への配慮は必要であり、そのため、精神科治療の現場においては薬物療法や心理教育 (疾病教育) だけでなく、単に病者を気遣うという以上の精神療法的接近が求められている。

ところで、患者の訴えにおける心理社会的背景と疾患の割合は、同じ病名であっても各患者によって異なるであろうし、同じ患者であっても場面によって違うことが想定され、臨床医を悩ませている。伝統的な精神医学では、患者が語る心の動きを診察者が共感的に追うことが出来る (了解可能) か否かを重視しており、正常な心の動きからの逸脱を病的なものと解釈していた。¹⁾ しかし、この判断は診察者の経験や想像力によっても変動する余地を残しており、臨床医が自ら下す了解可能の判断に悩み、客観性あるいは普遍性という点からみた弱点ともなっていた。そうした流れをうけ、米国精神医学会の精神障害の診断・統計マニュアル 第 5 版 (DSM-5) などに代表される操作的診断を用いることが、世界的な主流になっているが、ここまで述べてきた問題が解決されたわけではない。

一例として、我が国で大きな公衆衛生上の問題となっているうつ病について、DSM-5 による診断基準を表 1 に示す。ここに機能障害などの要件が加わることで、うつ病と診断することが出来る。しかし、どの項目において、いったいどの程度の訴えであれば、基準を満た

していると言えるのか、目の前の患者がその項目を満たしていると言ってよいのか、臨床医が日々悩むことに違いはない。そのため、こうした操作的診断基準に対する批判は根強く、症候学的あるいは生物学的観点から診断を見直すべきという意見も多い。また、他の診療科領域における画像診断の進歩はめざましいものであるが、精神科領域においては一部の疾患を除いて診断に必須のものとはされていない。しかし、ある意味で客観的把握が難しい臨床症状よりも、機械学習との相性がよいと考えられてもおり、ここに研究上の未開地を求めて報告が相次いでいるのが現状である。こうした背景を踏まえ、AIの導入という視座から精神科領域と、他の診療科領域との間の違いについて道喜がまとめた要点を表2に示す。²⁾

表1 うつ病 (DSM-5)の診断基準

以下の9項目のうち、5つが同じ2週間の中に存在して、少なくとも1つは1)か2)を含んでいること

- 1) ほとんど1日中、毎日の抑うつ気分
- 2) ほとんど1日中、毎日の興味または喜びの喪失
- 3) 食事療法をしていないのに体重減少、または体重増加 (5%以上の変化)
- 4) ほとんど毎日の不眠または睡眠過多
- 5) ほとんど毎日の精神運動制止、あるいは焦燥
- 6) ほとんど毎日の疲れやすい、気力の減退
- 7) ほとんど毎日の無価値感、過剰であるか不適切な罪責感
- 8) ほとんど毎日の、思考力や集中力の減退、または決断困難
- 9) 死についての反復思考、自殺念慮、自殺企図、ハッキリした計画

表2 AI導入の際の課題の精神医療とその他の医療との比較

	精神医療	その他の医療
病因	心理社会的背景の考慮が必須で、AIのモデルに組み込みにくい。	心理社会的背景の関与は少なく、精神医療ほど考慮しなくてもモデルは組めるかもしれない。
画像診断	一部の疾患で必須だが、それだけで診断が確定することは少ない。	疾患によっては、画像検査が診断に大きく関与。特に病理診断科や放射線診断科はほとんどが画像。
診断基準	客観的な指標が少ない。	客観的な指標がメイン。
診察	家族や知人との人間関係も観察や洞察が必要。	患者個人だけみても成立する場合が多い。
治療	薬物療法、疾病教育以外に、精神療法が含まれる。	薬物療法や手術の選択などのガイドラインが多く存在する。

2. 精神科領域における AI 関連論文数の推移

前節で述べたように、精神科領域特有の課題はあるものの、検索を行ったところ、2,866 件の文献が報告されていることを確認した。このうち、2001 から 2005 年までの文献は 21 件、2006 から 2010 年までは 88 件、2011 から 2015 年では 314 件であったが、2016 から 2020 年の間では実に 2,120 件と、それまでにない飛躍的な増加 (図 1) を認め、この領域における研究が進展していることがうかがえる。

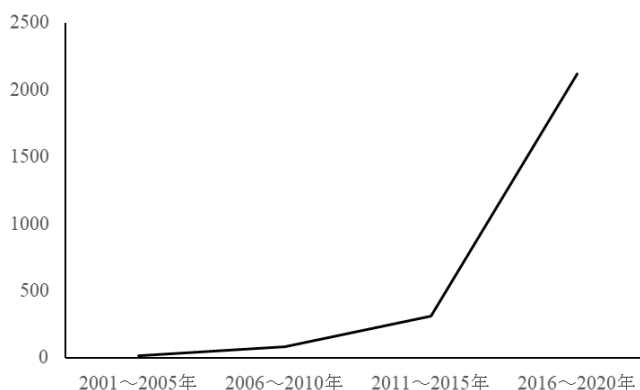


図 1 精神科領域における AI 関連論文数の推移

3. AI について研究の現状

前述の通り、これまで精神科領域における AI 関連論文の報告は相当数存在しており、その全て論ずることは出来ないが、ここに画像や治療予測、あるいはロボット技術といった視点から実際の研究をいくつか見ていくこととする。

まず、患者の訴えにおける心理社会的背景と疾患の割合といった観点からうつ病の患者群と健常群の判別は重要な課題となっていることは、前節で述べたとおりだが、Mwangi らは特に着目した一部の脳領域の核磁気共鳴画像法 (MRI) 所見を機械学習により解析することで、この 2 群の分類を試みている。³⁾ 結果として、2 群を分類する精度は 90.3% (感度 93.3%, 特異度 87.5%) と高く、また、抑うつ症状の重症度を予測できる可能性も併せ示している。この他にも、うつ病と双極性障害、⁴⁾ うつ病と統合失調症、⁵⁾ の分類に関する報告があり、その分類精度について最も高い条件では前者で 69%、後者で 74%とされている。数字だけを見る限り、うつ病の患者群と健常群の判別に高い精度を期待できそうであるが、この患者群の選択基準は、研究参加の 5 年以上前にうつ病の診断を受けていること、抗うつ薬による治療をしっかりと行ったにもかかわらず 3 ヶ月以上の症状持続があること、となっており、実のところ AI を用いなくても診断に苦慮することのない症例を画像の情報だけでも分類が出来たということのように見受けられる。先駆的な報告であることは確かではあるが、臨床の現場で診断に苦慮する症例を判別出来るものであるか否かは、さらなる検証を要すると考える。

精神疾患の診断は、表 1 に示したように、症状に基づいて行われているのが現状であり、現在主流となっている操作的診断についても、ある意味では暫定的なものと思なすことも

出来る。そのため、実際に生物学的な変化が生じているであろう脳の病態に着目し、MRI 所見を機械学習により解析した Drysdale らの報告がある。⁶⁾ ここでは、うつ病の患者が、症状と機能的結合のパターンにより 4 つのサブタイプに分類されただけでなく、反復経頭蓋磁気刺激 (rTMS) への治療反応性が異なっていることが示されており、これが他の集団においても再現性をもって確認されれば、診断から治療へとつながる AI 利活用の先鞭となり得る成果と言えるだろう。また、現在の操作的診断とは異なる新しい診断カテゴリーの創出を期待する一方で、診断根拠の説明を求められた際に、どのように回答を行うべきかといった課題が特に司法精神医学の領域において残されていると考えられる。

もっと直接的に AI を治療に活用する試みも多数なされている。英国では、自閉スペクトラム症 (ASD) の罹患者を対象に、人間型ロボットとのコミュニケーションを行わせて、情動コントロールやコミュニケーション能力の上昇などにつながったという報告がなされている。⁷⁾ このような治療への AI 導入の試みはロボットだけに留まらず、スマートデバイスのカタチでの利用も研究されている。ASD の罹患者に表情認識機能を搭載した眼鏡を装着させ、他者の顔を見た際に、その表情を正答すると得点が得られるというゲーム形式の介入を行い、その結果、不注意、衝動性や多動性に関連した症状が軽減したと報告している。⁸⁾ こうした治療的介入への AI 導入においては、安全性の問題が避けがたく存在しているものの、上述の介入は、患者の人体に対して手術を行うなどの直接的かつ生物学的な侵襲性を有するものではないため、大きな問題とされることはないようだ。

精神科の診療では、基本的に精神科医が患者の心的内界において生じたことを直接観察出来ないため、患者の語った言葉そのものが重要な所見となる。一例として、統合失調症の罹患者では、その話す内容のまとまりが悪く、意味が受け取りにくいとされる連合弛緩、考えが途中で止まってしまうため思考途絶や、自分にしか通用しない言葉を作ってしまう言語操作といった症状があることが知られている。そうした背景から、ヒトの使う『自然言語』を処理する技術である自然言語処理 (NLP) を、患者の語った言葉や言語構造から精神症状の重症度を評価する手法としての応用が期待されており、実際に 83% の精度でハイリスク者の発病を予測することに成功したとする報告もある。⁹⁾

我が国においても、こうした世界的な潮流に合わせ、岸本らがうつ病を対象として、患者の声の音響学的な特徴、体動や表情のデータを定量化し、重症度と相関する指標を導く機械学習アルゴリズムの開発を目指した PROMPT や、¹⁰⁾ 話し言葉や書き言葉からさまざまな精神疾患の症状の特徴を定量的に抽出することを目指す UNDERPIN を推進しており、その成果が待たれる。¹⁰⁾

結論

精神科領域は、病因や臨床上の訴えに心理社会的背景が反映されやすく、AI の導入に対する壁となると考えられていたが。近年、この領域における AI 関連論文数が飛躍的に増加しており、研究の進展がうかがえる。報告された研究としては、診断的研究が多いが、臨床的に判別が難しい症例を峻別できる水準にはないことや、診断根拠の説明を求められた際に、どのように回答を行うべきかといった課題が特に司法精神医学の領域において残され

ている。また、そもそも診断の Gold Standard をどのように考えるかといった問題があるように考えられる。新たな診断カテゴリーの創出を含め現在のヒトの医師を超える機械を開発する方向に進むのか、それとも現在のヒトの医師による診療を代行する水準を目指すのか、現在の研究において参照する診断を操作的診断としている限りは、後者に留まるものの、将来においての課題となりそうだ。また、現在のところ、治療デバイスとしては、患者の人体に対して生物学的介入を行うものは開発されていないため、大きな問題は生じていないものの、将来においては、そうした介入の安全性も課題となるであろう。

精神科領域における AI 導入については、まだ課題が残されているのは確かである。しかし、今後の我が国の精神医療が直面するであろう、人口減少に伴う医療供給サイドの減少、認知症を中心とした患者数の高止まり、地域の医者-患者バランスの不平等性増大といった局面に際して、¹²⁾ 効率的な医療を行うためのツールとして期待されているのは間違いない。

参考文献

- 1) 古茶大樹. 臨床精神病理学 日本評論社 2019 年
- 2) 道喜将太郎. 精神医療と人工知能 可能性と限界 心と社会. 2019; 50 (4): 64-71.
- 3) Mwangi B, Ebmeier KP, Matthews K, Steele JD. Multi-centre diagnostic classification of individual structural neuroimaging scans from patients with major depressive disorder. *Brain*. 2012; 135: 1508-21. DOI: 10.1093/brain/aws084. PMID: 22544901.
- 4) Redlich R, Almeida JJ, Grotegerd D, Opel N, Kugel H, Heindel W, Arolt V, Phillips ML, Dannlowski U. Brain morphometric biomarkers distinguishing unipolar and bipolar depression. A voxel-based morphometry-pattern classification approach. *JAMA Psychiatry*. 2014; 71 (11): 1222-30. DOI: 10.1001/jamapsychiatry.2014.1100. PMID: 25188810.
- 5) Koutsouleris N, Meisenzahl EM, Borgwardt S, Riecher-Rössler A, Frodl T, Kambeitz J, Köhler Y, Falkai P, Möller HJ, Reiser M, Davatzikos C. Individualized differential diagnosis of schizophrenia and mood disorders using neuroanatomical biomarkers. *Brain*. 2015; 138(7): 2059-73. DOI: 10.1093/brain/awv111. PMID: 25935725.
- 6) Drysdale AT, Grosenick L, Downar J, Dunlop K, Mansouri F, Meng Y, Fetcho RN, Zebley B, Oathes DJ, Etkin A, Schatzberg AF, Sudheimer K, Keller J, Mayberg HS, Gunning FM, Alexopoulos GS, Fox MD, Pascual-Leone A, Voss HU, Casey BJ, Dubin MJ, Liston C. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat Med*. 2017; 23(1): 28-38. DOI: 10.1038/nm.4246. PMID: 27918562
- 7) Huijnen CAGJ, Lexis MAS, Jansens R, de Witte LP. Roles, Strengths and Challenges of Using Robots in Interventions for Children with Autism Spectrum Disorder (ASD). *J Autism Dev Disord*. 2019; 49 (1): 11-21. DOI: 10.1007/s10803-018-3683-x. PMID: 30019273.
- 8) Vahabzadeh A, Keshav NU, Salisbury JP, Sahin NT. Improvement of Attention-Deficit/Hyperactivity Disorder Symptoms in School-Aged Children, Adolescents, and Young Adults With Autism via a Digital Smartglasses-Based Socioemotional Coaching Aid: Short-Term,

Uncontrolled Pilot Study. *JMIR Ment Health*. 2018; 5(2): e25. DOI: 10.2196/mental.9631. PMID: 29610109.

- 9) Corcoran CM, Carrillo F, Fernández-Slezak D, Bedi G, Klim C, Javitt DC, Bearden CE, Cecchi GA. Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry*. 2018; 17(1): 67-75. DOI: 10.1002/wps.20491. PMID: 29352548.
- 10) Kishimoto T, Takamiya A, Liang KC, Funaki K, Fujita T, Kitazawa M, Yoshimura M, Tazawa Y, Horigome T, Eguchi Y, Kikuchi T, Tomita M, Bun S, Murakami J, Sumali B, Warnita T, Kishi A, Yotsui M, Toyoshiba H, Mitsukura Y, Shinoda K, Sakakibara Y, Mimura M; PROMPT collaborators. The project for objective measures using computational psychiatry technology (PROMPT): Rationale, design, and methodology. *Contemp Clin Trials Commun*. 2020; 19: 100649. DOI: 10.1016/j.conctc.2020.100649. PMID: 32913919.
- 11) Integrated Innovation Lab for Psychiatry, Keio University School of Medicine. HP, URL: <https://www.i2lab.info/>
- 12) Sugawara N, Yasui-Furukori N, Shimoda K. Projections of psychiatrists' distribution for patients in Japan: a utilization-based approach. *Hum Resour Health*. 2021; 19 (1): 49. DOI: 10.1186/s12960-021-00594-z. PMID: 33836799.

D. 研究発表

なし。

E. 知的財産権の出願・登録状況

(予定を含む)

特になし。