

特集 2 生物学的自殺研究の最前線

2. 機械学習を用いた自殺予防の可能性とその問題点

木下翔太郎* 岸本泰士郎*

抄録：自殺者の90%以上は精神疾患を有しているとされており、精神科医にとって自殺の予防は最重要課題の一つである。しかし、精神神経科領域は、診断や重症度評価の客観的指標となるバイオマーカーに乏しく、複合的な要因によって引き起こされる自殺もまた正確な予測が困難とされてきた。近年、自殺予防において、機械学習(machine learning)を用いた研究が数多く行われるようになってきており、画像データや、自然言語処理(natural language processing)を用いた個人の日記やSNSへの投稿に関する研究など、従来にはない試みも出てきている。本稿では、機械学習を用いた自殺予防に関する新しい知見について紹介しつつ、社会への実装に際し検討すべき倫理的・法的・社会的な課題(ethical legal and social implications: ELSI)についても触れる。

日本生物学的精神医学会誌 31 (3) : 141-146, 2020

Key words : machine learning, artificial intelligence (AI), natural language processing suicide prevention, ethical legal and social implications (ELSI)

はじめに

自殺者の90%以上は精神疾患を有しているとされており²⁾、適切な治療介入によって予防できる可能性があることから、精神科医にとって、患者の自殺リスクの把握は最重要課題の一つである。しかしながら、精神神経科領域は、診断や重症度評価の客観的指標となるバイオマーカーに乏しく⁷⁾、複合的な要因によって引き起こされる自殺もまた正確な予測が困難とされてきた。

近年、情報通信技術(information and communication technology: ICT)の発展を通じて、多種多様で大量のデータが産出されるようになった。その大量のデータに対する解析手法として機械学習(machine learning)が注目されており、医療分野における研究開発も盛んに行われている。自殺予防においても、機械学習を用いた研究が数多く行われるようになってきており、画像データや、自然言語処理(natural language processing)を用いた個人の日記の解析、SNSへの投稿に関する研究など、従来にはない試みも出てきている。

本稿では、機械学習を用いた自殺予防に関する新

しい知見について紹介しつつ、社会への実装に際し検討倫理的・法的・社会的な課題(ethical legal and social implications: ELSI)についても触れる。

1. 機械学習の発展と自殺予防への応用

機械学習とは、人間の「学習」と同じようなプロセスをコンピュータープログラムによって実現するものであり、あらかじめ設定されたアルゴリズムに基づいて訓練用のデータからパターンや規則性を学習することで、新たなデータに対しても識別などを行うことができるようにする手法である。現在、人工知能(artificial intelligence: AI)が一世を風靡しているが、その多くがこれら機械学習の応用によって実現している。機械学習のアルゴリズムには、ナイーブベイズ、ニューラルネットワーク、ランダムフォレスト、ロジスティック回帰、サポートベクターマシンなど多様な種類があり、研究によって使い分けられている。機械学習は、多数の予測因子を用いた複雑な計算や因子間の比較も行うことができるため、単一の因子による識別が困難であったさまざまな領域で、機械学習の応用が期待されている。

Possibilities and problems of suicide prevention using machine learning

*慶應義塾大学医学部 精神・神経科学教室 (〒160-8582 東京都新宿区信濃町35) Shotaro Kinoshita, Taishiro Kishimoto : Department of Neuropsychiatry, Keio University School of Medicine, 35, Shinano-machi, Shinjyuku-ku, Tokyo 160-8582, Japan
【木下翔太郎 E-mail: shotaro.kinoshita@keio.jp】

従来の手法だけでは予測・予防が困難であった自殺の領域でも、機械学習を用いた研究が徐々に報告され始めている。過去5年間に公開された論文を対象にPubMed上で“suicide”と“machine learning”を含むという条件検索を行ったところ(検索実行日:2020年1月4日),107件の論文が該当し、タイトル、abstractの内容から機械学習を用いた自殺の予測・予防に関する研究を行っているとして判断し得たものは39件であった。これらの中から、使用した機械学習モデルや検証方法が明らかであったものをいくつかピックアップし、アプローチ別に紹介する(表)。

2. 人口統計学的データを用いた研究

医療機関や健康保険のデータベースには、結婚や就労の有無などのパーソナルデータや、治療状況に関する情報など個人に紐づく情報が多く含まれている。こうした電子データとして記録されている情報を活用して、自殺の予防・予測を行う研究は多く行われている。

Gradusらは、自殺による死亡者14,103人とそうでない者265,183人を対象に、健康保険データから家族構成、自殺企図の有無、収入、教育歴、就労状況、精神病既往、処方歴など自殺に関連する1,339の項目を抽出し、ランダムフォレストを用いて、自殺の有無を予測するモデルを構築した⁵⁾。構築されたモデルのAUC(area under curve)は、男性では0.80、女性では0.88であった。

3. 臨床検査データを用いた研究

Bhakらは、大うつ病性障害と診断されている自殺企図者56人、大うつ病性障害と診断されているが自殺企図のない39人、健常者87人について、末梢血のメチロームおよびトランスクリプトーム解析データを変数とし、ランダムフォレストを用いてうつ病および自殺のリスクを予測するモデルを構築した³⁾。構築されたモデルは、大うつ病性障害と診断されている者のうち自殺企図者とそうでない者を精度0.926で識別することができた。

4. 文章データを用いた研究

近年、機械学習と合わせて用いられることの多い技術に自然言語処理がある。自然言語処理とは、人が会話や読み書きで用いている「自然」言語を、コンピューターが直接解析する技術である。自然言語処

理により、大量のテキストデータを解析することが可能になったため、カルテや日記などの文章データから自殺を予測する研究も行われている。

Carsonらは、自殺企図のある入院患者27人と自殺企図のない入院患者46人に関して、臨床医の経過記録、家族・他の治療スタッフ・学校・地域サービス機関との連携に関する電子記録情報を収集して自然言語処理を行い、ランダムフォレストを用いて自殺企図の有無を予測するモデルを構築した⁴⁾。構築されたモデルのAUCは0.68であった。

de Ávila Berniらは、英国の自殺した文筆家であるヴァージニア・ウルフによって書かれた手紙・日記について、自殺直前の2か月間とそれ以外の期間のものを抽出して自然言語処理を行ったうえで、ナイーブベイズを用いたテキスト分類によって自殺直前期の物を識別するモデルを構築した¹⁾。結果、“blue”や“books”など、自殺直前期に多く出現する単語群が発見され、また構築したモデルによる自殺直前期とそれ以外の文章との識別のAUCは0.80であった。

5. 音声データを用いた研究

患者の発言内容や声の調子が精神神経科領域の病状評価において有用であることから、音声データを用いた自殺予測に関する研究も報告されている。Pestianらは、130人の自殺企図者、126人の精神疾患患者、および123人の健常者に対して、音声を記録しながら自由回答形式の半構造化インタビューを実施し、発言内容の単語・単語ペアを文字データとして抽出、声の周波数などの特性を音声データとして抽出し、サポートベクターマシンを用いて自殺企図者とそうでない精神疾患患者・健常者を識別するモデルを構築した¹²⁾。構築されたモデルは、自殺企図者と健常者の識別ではAUC 0.93(文字データ単独)、自殺企図者とそうでない精神疾患患者の識別ではAUC 0.82(文字データと音声データの両方)、自殺企図者とそうでない精神疾患患者・健常者の識別ではAUC 0.87(文字データ単独、または文字データと音声データの両方)の値を示し、音声データの併用による識別が有用である可能性を示した。

6. 画像データを用いた研究

画像データの機械学習により疾患の診断を試みる研究は従来より数多く行われており、脳画像データを用いた自殺の予防・予測に関する研究も報告され

ている。Justらは、希死念慮のある参加者17人とコントロール群17人の計34人の成人に対して、“death”や“praise”などの死および生に関する30種類の単語をランダムに提示した際に脳のどこが活動するかを機能的磁気共鳴画像法(fMRI)によるスキャンによって分析し、6つの単語と5つの脳領域が両群間を識別するマーカーになることを見いだしたうえで、ナイーブベイズを用いて希死念慮の有無を識別するモデルを構築した⁶⁾。構築されたモデルは、希死念慮の有無を91%の精度で識別したほか、希死念慮を有する者のうち自殺未遂者とそれ以外の者を94%の精度で識別した。

7. 予測結果への介入を伴った研究

機械学習を用いた自殺予測についてはさまざまな研究が行われているが、その多くが過去の自殺者のデータに基づく後ろ向き研究か、研究時点の希死念慮の有無を予測するものである。自殺という事象が突発的であり対応にも緊急性が求められること、また、後述するような倫理的問題などもあるため、現時点では自殺の予防・予測に関する介入を伴う研究のハードルは高いが、まったく報告がないわけではない。

Liuらは、ソーシャルメディア上の投稿27,007件に対し、“die”や“suicide”などの自殺のリスクが高いと専門家が分析した単語に注釈をつけて一部を教師データとし、サポートベクターマシンを用いて、残りの投稿データから自殺リスクを予測するモデルを構築した⁹⁾。構築されたモデルのF値は0.86であった。その後、モデルが抽出したコメントの投稿者であるインターネットユーザー12,486人に対してダイレクトメッセージを送付し、希死念慮の有無や支援の必要性についてアンケートを実施し、希望者にはカウンセラーがケアを行い、その前後で書き込み内容の比較も行った。アンケートの有効回答が得られた1403人の投稿者のうち89.7%が希死念慮を実際に肯定し、カウンセラーによるケアを1か月受けた者の書き込み内容の変化を観察したところ、死に関する単語の減少と未来に関する単語の増加が有意に見られた。

8. 機械学習を用いた自殺予防に関する倫理的・法的・社会的な課題(ELSI)

ここまで概観してきたように、機械学習はあくまで一定のデータを元に学習した結果から予測する

ツールであり、その精度は人間に比して高くなることはあるが完全ではない。また、学習元のデータの偏りに左右されることもあるなど、結果をそのまま鵜呑みにするべきでないケースが出てくる可能性もある。そのため、機械学習を用いたツールを開発・社会実装するにあたっては、種々の倫理的・法的・社会的な課題について配慮することが求められる。これらの課題は、ELSI (ethical legal and social implications) と総称されている。

機械学習やAIの利用については、各国においてさまざまなルールが検討・提唱されており、我が国においても、各国の議論を踏まえながら、2019年3月に政府統一のAI社会原則として「人間中心のAI社会原則」が策定されている¹¹⁾。同原則の中では、AI利用に関わる最終判断は人が行うとする「人間中心の原則」、AIの誤用・悪用を防ぐために広く教育環境の整備を求める「教育・リテラシーの原則」、個人のパーソナルデータが本人の望まない形で流通・利用されないようにすべきとする「プライバシー保護の原則」、セキュリティの確保および特定AIに対する依存の回避を求める「セキュリティ確保の原則」、支配的地位を利用した不当なデータの収集や不公正な競争があってはならないとする「公正競争確保の原則」、人々の多様性を守るためにAIの利用・データの取得・利用結果について説明責任・透明性を確保すべきとする「公平性、説明責任及び透明性の原則」、データ利用環境の整備や規制改革を求める「イノベーションの原則」の7つの原則が示されている。これらの原則は、現時点での機械学習一般のELSIに関する論点を幅広くカバーしており、あらゆる機械学習を用いたツールを開発・社会実装していくうえで遵守されるべきものであるといえる。しかし、今後の技術発展に伴って新たな論点が出現することも十分考えられるため、常に最新の動向を注視していく必要がある。

また、自殺予防というセンシティブな領域に機械学習を用いる場合のELSIについても、個別にさまざまな議論がなされている。McKernanらは、AIを用いた自殺予防を社会に実装するうえで、自殺に関するスティグマを軽減するために関係者間で議論すること、リスクのある個人のうち何%までAIの監視対象にするのかなどの実践基準を作ることなどを推奨している¹⁰⁾。Linthicumらは、解析の対象となるべき医療データをもたない不法移民などの人々におけるリスク検出について社会的不平等が発生し得ること、機械学習のアルゴリズムは自殺を試みる可能性がある人に警告を発することはできても今ま

表 本稿で紹介した研究の要約

著者 (発表年)	対象	使用したデータ	調査・評価方法	結果
Gradus JL (2020) ⁵⁾	デンマーク 自殺により死亡した 14,103 人 自殺と無関係な 265,183 人	デンマークの国民健康保険 レジストリデータ (1995-2015)	家族構成, 自殺企図, 収入, 教育歴, 就労 状況, 精神病既往, 飲酒歴, 処方歴など, 健康保険データに登録された 1339 の項目を 変数とし, ランダムフォレストを用いて, 自殺の有無を予測するモデルを構築。交差 検証で評価。	AUC: 0.80 (男性), 0.88 (女性)
Bhak Y (2019) ³⁾	韓国 大うつ病性障害と診断されてい る自殺企図者 56 人 (SA) 大うつ病性障害と診断されてい るが自殺企図のない 39 人 (MDD) 健常者 87 人 (Control)	psychiatric clinic of Korea University Anam Hospital in Seoul にてリクルートされ た患者群に関連するデータ (2015-2017)	自殺企図・病歴に関するアンケート, ハミ ルトン評価尺度, Scale for Suicide Ideation (SSI), 末梢血のメチロームおよびトランス クリプトーム解析データを変数とし, ラン ダムフォレストを用いてうつ病および自殺 のリスクを予測するモデルを構築。leave- one-out 交差検証で評価。	①: SA vs MDD ②: MDD vs Control ③: SA vs Control 精度: ① 0.926 ② 0.873 ③ 0.867 感度: ① 0.982 ② 0.59 ③ 0.679 特異度: ① 0.846 ② 1 ③ 0.989 陽性的中位: ① 0.902 ② 1 ③ 0.974
Carson NJ (2019) ⁴⁾	米国 自殺企図のある入院患者 27 人 自殺企図のない入院患者 46 人	psychiatric unit of a community hospital in Massachusetts にてリク ルートされた患者群に関連 するデータ (2012-2016)	入院患者へのアンケート, 臨床医の経過記 録, 家族, 治療チームの他のメンバー, 学校, 地域サービス機関との連携に関する電子健 康記録の情報に対して自然言語処理を行い, それらを元に自殺企図の有無をランダム フォレストを用いて予測するモデルを構築。 交差検証で評価。	AUC: 0.68 感度: 0.83 感度: 0.22 陽性的中位: 0.42
de Ávila Berni G (2018) ¹⁾	英国 自殺直前の 2 か月間に書かれ た 46 の手紙や日記等の文章 他の期間からランダムに選択さ れた 54 の文章	ヴァージニア・ウルフの著 作物 (-1941)	ヴァージニアウルフによって書かれた手紙・ 日記について, 自殺直前の 2 か月間のもの と, 他の期間からのものに対して, ナイー ブベイズを用いたテキスト分類によって自 殺を予測するモデルを構築。	AUC: 0.80 精度: 0.80 感度: 0.69 特異度: 0.92 陽性的中位: 0.90
Pestian JP (2017) ¹²⁾	米国 自殺企図者 130 人 (SA) 自殺企図のない精神病患者 126 人 (NSA) 健常者 123 人 (Control)	Cincinnati Children's Hospital Medical Center, the University of Cincinnati Medical Center, Princeton Community Hospital にてリ クルートされた患者群から 得たデータ (2013-2015)	音声を記録しながら自由回答形式の半構造 化インタビューを実施し, 発言内容の単語・ 単語ペアを文字データとして抽出, 声の周 波数などの特性を音声データとして抽出し, サポートベクターマシンを用いて自殺企図 者とそうでない精神疾患患者・健常者を識 別するモデルを構築。leave-one-out 交差検 証で評価。	SA と Control の識別 AUC 0.93 SA と NSA の識別 AUC 0.82 SA と NSA・Control の識別 AUC 0.87
Just MA (2017) ⁹⁾	米国 希死念慮のある参加者 17 人 (SI) 希死念慮のない参加者 17 人 (Control)	参加者から得たデータ	参加者に対して, death や praise などの死 および生に関する 30 種類の単語をランダム に提示した際に脳のどこが活動するかを機 能的磁気共鳴画像法 (fMRI) によるスキャ ンによって分析し, 6 つの単語と 5 つの脳 領域が両群間を識別するマーカーになるこ とを見いだしたうえで, ナイーブベイズを 用いて希死念慮の有無を識別するモデルを 構築。交差検証で評価。	SI と Control の識別 精度: 91% SI のうち自殺未遂者とそれ以外 の者の識別 精度: 94%
Liu X (2019) ⁹⁾	中国 ソーシャルメディア上の投稿 27,007 件 抽出されたインターネットユー ザー 12,486 人	ソーシャルメディア上の投 稿 (2017-2018)	ソーシャルメディア上の投稿に対し, 死, 自殺, これから, などの自殺のリスクが高 いと専門家が分析した単語に注釈をつけて 教師データとし, 注釈をつけていない書き 込みからロジスティック回帰などを用いて 自殺を予測するモデルを構築。10 分割交差 検証で評価。モデルが抽出したコメントの 投稿者に対してダイレクトメッセージを送 付し, 希死念慮の有無や支援の必要性につ いてアンケートを実施。希望者にはカウ ンセラーがケアを行い, その前後で書き込み 内容を比較。	F 値: 0.86 アンケートの有効回答 1,403 人 のうち希死念慮を肯定した者: 89.73% ダイレクトメッセージに反応 し, カウンセラーによるケアを 1 か月受けた者の書き込み内容 の変化 死に関する単語の減少: P 値 0.03 未来に関する単語の増加: P 値 0.02

さに試みようとしている人には無力であることなどの問題点を指摘している⁸⁾。なお, いずれの報告においても, 機械学習や AI が予測を行った場合であっても, 臨床医が判断しなければならない点は残ることについて言及されていた。

おわりに

機械学習を用いた自殺予防に関する研究について概説した。これまで解析の対象となつてこなかったデータを含め, さまざまなアプローチによる研究が

報告されているが、現時点で実臨床へそのまま応用できるものが多いとは言い難い。例えば、学習データとなるサンプルデータが少ないモデルの場合、より大規模な集団を対象とした際に同じ精度で予測ができるかどうか不透明である。また、fMRI を用いた予測モデルや、過去の文章データの収集など、解析対象となるデータの収集に多くの費用や労力がかかる方法は、スクリーニングとして行う検査方法としては馴染みにくい。加えて、現時点では機械学習で示された自殺リスクに対してどのような介入が有効であるかについての研究も不足している。実臨床において有用な精度・簡便さを有したツールの開発と、示されたリスクに対してどの程度介入すべきかについての更なる研究が待たれるところである。

本稿でみてきたように、機械学習を用いたツールによる自殺予防を社会実装するうえでは、種々の ELSI にも向き合う必要がある。ツールの開発者や使用者は、既に示された機械学習一般に関する法規範に従うだけでなく、ツール固有の ELSI に関して多角的に検討を行い、その妥当性について絶えず検証を行っていくことが求められるだろう。

利益相反：開示すべき利益相反は存在しない。

文 献

- 1) de Ávila Berni G, Rabelo-da-Ponte FD, Librenza-Garcia D, et al (2018) Potential use of text classification tools as signatures of suicidal behavior : A proof-of-concept study using Virginia Woolf's personal writings. *PLoS ONE*, 13 (10) : e0204820.
- 2) Bertolote JM and Fleischmann A (2002) Suicide and psychiatric diagnosis : a worldwide perspective. *World Psychiatry*, 1 : 181-185.
- 3) Bhak Y, Jeong HO, Cho YS, et al (2019) Depression and suicide risk prediction models using blood-derived multi-omics data. *Transl Psychiatry*, 9 (1) : 262.
- 4) Carson NJ, Mullin B, Sanchez MJ, et al (2019) Identification of suicidal behavior among psychiatrically hospitalized adolescents using natural language processing and machine learning of electronic health records. *PLoS ONE*, 14 (2) : e0211116.
- 5) Gradus JL, Rosellini AJ, Horváth-Puhó E, et al (2020) Prediction of sex-specific suicide risk using machine learning and single-payer health care registry data from Denmark. *JAMA Psychiatry*, 77 (1) : 25-34.
- 6) Just MA, Pan L, Cherkassky VL, et al (2017) Machine learning of neural representations of suicide and emotion concepts identifies suicidal youth. *Nat Hum Behav*, 1 : 911-919.
- 7) Kapur S, Phillips AG and Insel TR (2012) Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry*, 17 : 1174-1179.
- 8) Linthicum KP, Schafer KM and Ribeiro JD (2019) Machine learning in suicide science : Applications and ethics. *Behav Sci Law*, 37 : 214-222.
- 9) Liu X, Liu X, Sun J, et al (2019) Proactive suicide prevention online (PSPO) : Machine identification and crisis management for Chinese social media users with suicidal thoughts and behaviors. *J Med Internet Res*, 21 (5) : e11705.
- 10) McKernan LC, Clayton EW and Walsh CG (2018) Protecting life while preserving liberty : Ethical recommendations for suicide prevention with artificial intelligence. *Front Psychiatry*, 9 : 650.
- 11) 内閣官房 (2019) 人間中心の AI 社会原則 (平成 31 年 3 月 29 日 統合イノベーション戦略推進会議決定) <https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/aigensoku.pdf> (accessed 2020.1.4).
- 12) Pestian JP, Sorter M, Connolly B, et al (2017) A machine learning approach to identifying the thought markers of suicidal subjects : A prospective multi-center trial. *Suicide Life-Threat Behav*, 47 (1) : 112-121.

■ ABSTRACT

Possibilities and problems of suicide prevention using machine learning

Shotaro Kinoshita, Taishiro Kishimoto

Department of Neuropsychiatry, Keio University School of Medicine

It is widely acknowledged that over 90% of those who committed suicide had a psychiatric diagnosis at the time of death. Hence, suicide prevention is one of the top priorities for psychiatrists to address. However, the field of neuropsychiatry lacks biomarkers that could serve as objective indicators for diagnosis and severity assessment. Suicide is a complicated result caused by multiple factors and can be difficult to accurately predict. In recent years, research based on data analysis using artificial intelligence (AI) supported by machine learning has attracted attention in the medical field. In suicide prevention, an unprecedented amount of research using machine learning have been reported. In this paper, we will introduce new information about suicide prevention using machine learning as well as comment on the ethical, legal, and social implications (ELSI) that should be considered for implementation into society. There are no potential conflicts of interest to disclose.

(Japanese Journal of Biological Psychiatry 31 (3) : 141-146, 2020)
