

生成AIを用いた治験・臨床研究関連文書のデジタルトランスフォーメーションに向けた研究

研究代表者 浅野 健人 未来医療開発部・特任准教授

#### 研究要旨

##### 目的：

本研究では臨床研究/治験の実施における関連文書にまつわる生成 AI の活用について、完全自動化から人間と生成 AI の協働まで、既存領域知識をどのようにに取り込み、精度を向上させながら安全かつ効果的に利用できるかを検証・実証することを目的とする。

大規模言語モデル（大規模な語句や文章と紐付けされたコーパスで事前訓練された、数百万から数十億以上のパラメータを持つ深層学習モデルを指し、Transformer アルゴリズムを使うことで近年大幅に精度が向上し、幅広い文書生成タスクに優れた汎用モデル）を治験・臨床研究文書に関する業務に応用するにあたり以下の三つの視点で検証、実証を行う。

テーマ A. 提案書から標準化文書を自動で生成する機能の検証

テーマ B. 大規模言語モデルを用いた研究説明文書の平易化による理解力向上の検証

テーマ C. 電子カルテからの自動情報抽出精度の検証

##### 方法：

1. テーマ A：提案書から標準化文書を自動で生成する機能の検証

「人を対象とする生命科学・医学系研究に関する倫理指針」準拠の研究プロトコルを 2 種（胃癌患者対象研究：1 種、食道がん患者対象研究：1 種）、臨床研究法準拠の研究プロトコルの 1 種を、DEEP INTELLIGENT PHARMA(SG) PRIVATE LIMITED が提供している生成 AI ソフトウェア（WritingX）を用いて作成した。

2. テーマ B：大規模言語モデルを用いた研究説明文書の平易化による理解力向上の検証

大規模言語モデルを用いて、大阪大学医学部附属病院で用いている研究説明用の文書（3447 字）を平易化（701 字～1108 字）した。その文書を 10-80 歳代までの一般ネットパネルに提示し、理解度や同意への意向について調査した。

3. テーマ C：電子カルテからの自動情報抽出精度の検証

2015 年 1 月～2023 年 12 月に大阪大学医学部附属病院に通院されていた患者の電子カルテデータを情報源として、DEEP INTELLIGENT PHARMA(SG) PRIVATE LIMITED が提供している生成 AI ソフトウェア（WritingX 等）を用いて、模擬研究 eCRF の自動入力の実証と参加適格者/不適格者判定の実証を実施した。

##### 結果：

1. テーマ A：提案書から標準化文書を自動で生成する機能の検証

WritingX を用いて、3 つの研究に対して、過去の同疾患領域の研究プロトコルを参照する：リファレンスあり、参照しない：リファレンスなしの研究プロトコルを自動作成した。

RAG (Retrieval-Augmented Generation) モジュールを活用することで、章の執筆順序と依存関係を自動的に推論し、関連性に基づいて参照情報を選択することができた。結果として自動生成される内容の品質と正確性に貢献し、ハルシネーション問題の発生を減らし、類似検索内容のノイズを削減することに成功した。

2. テーマ B : 大規模言語モデルを用いた研究説明文書の平易化による理解力向上の検証

研究への同意の意向については原文と平易化文章の間で差が見られなかった。その一方で、平易化文章では概ね全体の理解度が向上し、他の人に内容を説明できるとの回答が多くなった。年代別に詳細を見ると、平易化文書によって 50 代から 70 代ではかえって理解度が低下する傾向があることが示唆された。

3. テーマ C : 電子カルテからの自動情報抽出精度の検証

模擬研究 electric CRF の自動入力の実験では 99% を超える精度を示したが、記載のない情報に関するハルシネーションが発生することを確認した。模擬研究で参加適格者/不適格者の判定を行う実験では、適格基準の判定精度は 98.9% を達成したが、除外基準では特定不可と除外対象外の判定が混在した。

結論 :

1. テーマ A : 提案書から標準化文書を自動で生成する機能の実験

AI が生成する研究プロトコルは、文体や表現の一貫性、内容の整理と明確化、倫理的配慮と同意手続きなどで、改善すべき点が存在することが明らかになった。今後これらの改善により生成 AI によるプロトコル作成がより高度なレベルで実用化できる可能性が示唆された。

2. テーマ B : 大規模言語モデルを用いた研究説明文書の平易化による理解力向上の実験

研究に関する IC 説明用文書を、LLM を用いて平易化することで理解度の向上、あるいは、同意取得の意向への影響などを評価した。文章の平易化によって同意への意向に影響を及ぼすことなく理解度を向上させることができる可能性、また、年代によって好まれる文書に特徴がある可能性を明らかにすることができた。

3. テーマ C : 電子カルテからの自動情報抽出精度の実験

生成 AI を使用して電子カルテデータから eCRF への自動入力や、治験実施計画書からの eCRF 項目の抽出が可能であることが示された。判定結果の信頼性を向上させるためには、その判断根拠もアプリケーション上で明確に提示することが重要となる。

## 研究組織

### 研究代表者

浅野 健人

(国立大学法人大阪大学 医学部附属病院未来医療開発部・特任准教授)

### 分担研究者

土岐 祐一郎

(国立大学法人大阪大学 医学部附属病院消化器外科学・教授)

武田 理宏

(国立大学法人大阪大学 医学部附属病院医療情報学・教授)

川崎 良

(国立大学法人大阪大学 大学院医学系研究科社会医学講座公衆衛生学・教授)

### 研究協力者

和田 聖哉

(国立大学法人大阪大学 医学部附属病院医療情報学・寄附講座助教)

田中 晃司

(国立大学法人大阪大学 医学部附属病院消化器外科学・学部内講師)

山本 晋也

(国立大学法人大阪大学 医学部附属病院未来医療開発部・招へい教授)

中島 悠太

(大阪大学 データビリティフロンティア機構 教授)

大谷まゆ

(株式会社サイバーエージェント)

張培楠

(株式会社サイバーエージェント)

## A. 研究目的

新規医薬品・機器開発のための臨床研究/治験には関連文書作成に膨大な時間と労力がかかる。実際の医療現場ではこのプロセスは非効率的であり、研究開発の高コストは製薬企業の財政にも負担をもたらし、医薬品・機器の価格にも反映されることになる。AI 戦略 2022 では、「医療、創薬、材料科学等の分野における AI 利活用の更なる注力」することが具体目標として記載されており、生成 AI を活用し文書作成の効率化等を図ることにより、日本全体の医薬品・医療機器開発力の向上を促進することは、ドラッグラグ・ロスの解決のために急務である。この課題を解決するため、自然言語の生成 AI 技術が注目され<sup>1</sup>、臨床研究の関連文書作成にも応用が期待される<sup>2</sup>。医療分野での生成 AI は、患者への情報提供<sup>3</sup>、手術記録作成<sup>4</sup>、電子カルテの支援<sup>5</sup>などが試みられ、適切なプロンプト(大規模言語モデルを用いる際の背景情報、事前条件、そして、指示等)作成に加えて、文脈化、ガイドライン参照などのファインチューニング(精度向上のために目的に沿ったデータで追加学習することや参照情報を規定することなどの一連の作業)で精度が向上する報告もある<sup>6,7,8</sup>。一方、医療分野での利用に際しては種々のリスク(バイアス、偽情報、プライバシー、過度な依存)を踏まえ実装前の検証は不可欠である。

AI 戦略 2022 や世界的な生成 AI の利活用の急速な進展等を踏まえ、国内外で生成 AI に関する全般的な議論の機運が高い今こそ、今後の各分野での生成 AI の利用に備えて、生成 AI が保健医療分野にもたらしうる業務効率化や医療の質向上等の可能性を検証し、いち早くこれらの動きを厚労行政の中に取り込むことは喫緊の課題である。

本研究では臨床研究の実施における関連文書にまつわる生成 AI の活用について、完全自動化から人間と生成 AI の協働まで、既存領域知識をどのように取り込み、精度を向上させながら安全かつ効果的に利用できるかを検証・実証することを目的とする。そして大規模言語モデル（大規模な語句や文章と紐付けされたコーパスで事前訓練された、数百万から数十億以上のパラメータを持つ深層学習モデルを指し、Transformer アルゴリズムを使うことで近年大幅に精度が向上し、幅広い文書生成タスクに優れた汎用モデル）を治験・臨床研究文書に関する業務に応用するにあたり以下の三つの視点で検証、実証を行う。

- ・テーマ A : 提案書から標準化文書を自動で生成する機能の検証
- ・テーマ B : 大規模言語モデルを用いた研究説明文書の平易化による理解力向上の検証
- ・テーマ C : 電子カルテからの自動情報抽出精度の検証

## B. 研究方法

### 1. テーマ A : 提案書から標準化文書を自動で生成する機能の検証

「人を対象とする生命科学・医学系研究に関する倫理指針」準拠の研究プロトコルを 2 種（胃がん患者対象研究：1 種、食道がん患者対象研究：1 種）、臨床研究法準拠の研究プロトコルの 1 種を、DEEP INTELLIGENT PHARMA(SG) PRIVATE LIMITED が提供している生成 AI ソフトウェア (WritingX) を用いて作成した。特に、RAG (Retrieval-Augmented Generation) 技術、医学文書構造化解析技術、思考連鎖 (Chain-of-Thought, CoT) プロンプティング技術を活用し、プロトコルを手動で作成する時間と労力を削減

し、自動生成される文書の正確性と一貫性を追求した。また、生成 AI を用いて作成した研究プロトコルのうち、食道がん患者対象の研究プロトコル (リファレンスあり：過去の同疾患領域の研究プロトコルを参照する) を読んだ印象等について、医療分野の有識者 (14 名) に対しアンケートを実施した。

### 2. テーマ B : 大規模言語モデルを用いた研究説明文書の平易化による理解力向上の検証

大規模言語モデルを用いて、大阪大学医学部附属病院で用いている研究説明用の文書 (3447 字) を平易化 (701 字～1108 字) した。その文書を 10-80 歳代までの一般ネットパネルに提示し、理解度や同意への意向について調査した。その結果を集計し、年齢、性別、社会経済背景要因で調整をした上で平易化によって原文に比べて理解度および同意への意向において向上あるいは低下が見られたかを理解度スコアの平均値の差を持って評価した。

### 3. テーマ C : 電子カルテからの自動情報抽出精度の検証

2015 年 1 月～2023 年 12 月に大阪大学医学部附属病院に通院されていた患者の電子カルテデータを情報源として、DEEP INTELLIGENT PHARMA(SG) PRIVATE LIMITED が提供している生成 AI ソフトウェア (WritingX 等) を用いて、模擬研究 electric CRF (eCRF) の自動入力の実証と参加適格者/不適格者判定の実証を実施した。

## C. 研究結果

### 1. テーマ A : 提案書から標準化文書を自動で生成する機能の検証

生成 AI ソフトウェア (WritingX) を用いて、3 種それぞれのリファレンスあり、リフ

アレンスなしの研究プロトコルを自動作成した。RAG モジュールを活用することで、章の執筆順序と依存関係を自動的に推論し、関連性に基づいて参照情報を選択することができた。結果として自動生成される内容の品質と正確性に貢献し、ハルシネーション問題の発生を減らし、類似検索内容のノイズを削減することに成功した。また、CoT プロンプティングという大規模言語モデルを段階的な推論に導く技術により、複雑なタスクを中間推論ステップのシーケンスに分解し、自然言語形式でモデルに提示することで、モデルがタスク目標をよりよく理解し、より正確で論理的な結果を自動生成するのに貢献した。

## 2. テーマ B :大規模言語モデルを用いた研究説明文書の平易化による理解力向上の検証

研究への同意の意向については原文と平易化文章の間で差が見られなかった。その一方で、平易化文章では概ね全体の理解度が向上し、他の人に内容を説明できるとの回答が多くなった。年代別に詳細を見ると、平易化文書によって 50 代から 70 代ではかえって理解度が低下する傾向があることが示唆された。

## 3. テーマ C :電子カルテからの自動情報抽出精度の検証

模擬研究 eCRF の自動入力の実証では 99% を超える精度を示したが、記載のない情報に関するハルシネーションが発生することを確認した。模擬研究で参加適格者/不適格者の判定を行う実証では、適格基準の判定精度は 98.9% を達成したが、除外基準では特定不可と除外対象外の判定が混在した。

## D. 考察

### 1. テーマ A :提案書から標準化文書を自動で生成する機能の実証

生成 AI を用いて作成した研究プロトコルは、その完成度が期待を超えるものである一方、改善を要する部分も多く存在することが分かった。具体的には、文章表現においてより自然で読みやすい文章にすること、内容の明確化として冗長な部分を減らしつつ研究目的を明確にすること、倫理的な配慮と同意手続きに関する記述を充実させることが求められる。

また、作成された研究プロトコルはあくまで「たたき台」として活用するべきであり、研究背景やコンセプト、特に今回検討した手術の介入を評価対象にした研究では、AI が理解しにくい部分を研究者が補完する必要があると考えられた。

### 2. テーマ B :大規模言語モデルを用いた研究説明文書の平易化による理解力向上の実証

LLM を用いて医療現場の研究に関する IC 説明文書を平易化することで、理解度向上と同意取得への影響を評価した。小規模な評価ながら、文章の平易化は同意への意向に影響を与えることなく、理解度を向上させる可能性が示唆された。

ただし、医療従事者において評価の低かった平易化方法は、ネットパネル調査においても理解度向上を示さず、これが文章の難易度によるものか、文章の体裁によるものかはわからず、今後さらなる検討の余地がある。また、年齢層によって好まれる文章の特徴が異なる可能性を示された。今後、LLM を用いることで、個々の患者に合わせた適切な情報提供が出来る可能性が示唆された。

### 3. テーマ C : 電子カルテからの自動情報抽出精度の検証

#### 3-1. 模擬研究 eCRF の自動入力検証

自然言語で書かれた医学情報は AI により高精度 (99.7%) で構造化できることが確認された。医学用語を含む日本語にも対応でき、臨床データの自動抽出における大きな進歩であると思われた。ただし、性別の誤判定(ハルシネーション)も発生し、AI の推論根拠を明確化し、ユーザーによる評価を容易にする仕組みが必要であると考えられる。

#### 3-2. 研究参加適格者/不適格者の仕分け検証

AI は治験実施計画書から適格基準と除外基準を正確に抽出し、高精度な判定を行った。しかし、誤って判定したものとしては、“リンパ節腫大”を“リンパ腫”と誤認したものや、“活動性 B 型肝炎”を“コントロール不良な感染症”と誤認したものがあつた。これらは、医学的解釈を伴うケースでは課題が残ることということを示唆しているのではないかと考えられた。

また、現状では電子カルテに記載のない情報は判定不可とされるべきであつたが、AI はそれが困難であり、何らかの解釈を行い、出力した可能性がある。これは指示プロンプトの改善により、記載のない情報に対して“特定不可”と出力できる可能性があると考えられた。

## E. 結論

### 1. テーマ A : 提案書から標準化文書を自動で生成する機能の検証

AI が生成する研究プロトコルは、文体や表現の一貫性、内容の整理と明確化、倫理的配慮と同意手続きなどで、改善すべき点が存在することが明らかになった。今後これらの改

善により生成 AI によるプロトコル作成がより高度なレベルで実用化できる可能性が示唆された。また、現時点においても、AI が生成するプロトコルは、研究者にとって有用な「たたき台」や「参考資料」となり得るため、有用な支援ツールとなり得ると考えられた。

### 2. テーマ B : 大規模言語モデルを用いた研究説明文書の平易化による理解力向上の検証

研究に関する IC 説明用文書を、LLM を用いて平易化することで理解度の向上、あるいは、同意取得の意向への影響などを評価した。今回は小規模な技術的評価にとどまっているが、文章の平易化によって同意への意向に影響を及ぼすことなく理解度を向上させることができる可能性、また、年代によって好まれる文書に特徴がある可能性を明らかにすることができた。

### 3. テーマ C : 電子カルテからの自動情報抽出精度の検証

本研究により、生成 AI を使用して電子カルテデータから eCRF への自動入力や、治験実施計画書からの eCRF 項目の抽出が可能であることが示された。精度と信頼性を向上させるためには、生成 AI による判断根拠を明確に提示しそれを容易に確認できるユーザーインターフェースの設計が重要である。

## F. 健康危険情報

なし

## G. 研究発表

1. 論文発表  
準備中

## 2. 学会発表

山本 晋也 「生成 AI を用いた治験・臨床研究関連文書のデジタルトランスフォーメーションに向けた研究」日本臨床試験学会第15回学術集会総会 2024/3/8

(テーマ A・テーマ C の研究内容に該当)

## H. 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

## (文献)

1. Betzler BK, Chen H, Cheng CY, Lee CS, Ning G, Song SJ, Lee AY, Kawasaki R, et al. Lancet Digital Health In press 2023.
2. Ghim JL, Ahn S. Transl Clin Pharmacol. 2023;31:131-138.
3. Waisberg, E et al. Ann Biomed Eng 2023; 51, 2353-2355.
4. Kaarre J et al. Sports Traumatology, Arthroscopy 2023;31:5190-5198.
5. Harris JE. JAMA. 2023;330:801-802.
6. Lim DYZ et al. J Gastroenterol Hepatol. 2023 ePub ahead of print.
7. Shah NH et al. JAMA. 2023;330:866-869.
8. Singhal K et al. Nature 2023;620:172-180.