

別紙1

厚生労働科学研究費補助金
厚生労働科学特別研究事業

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

令和5年度 総括・分担研究報告書

研究代表者 浅野健人

令和6（2024）年 5月

目 次

I. 総括研究報告	
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 -----	1
浅野健人	
II. 分担研究報告	
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 -----	10
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 -----	24
3. テーマC：電子カルテからの自動情報抽出精度の検証 -----	49
III. 研究成果の刊行に関する一覧表 -----	55

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 of 検証と参加適格者/不適格者判定の検証を実施した。

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¹、臨床研究の関連文書作成にも応用が期待される²。医療分野での生成 AI は、患者への情報提供³、手術記録作成⁴、電子カルテの支援⁵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^{6,7,8}。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) の自動入力の検証と参加適格者/不適格者判定の検証を実施した。

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.

厚生労働科学研究費補助金（厚生労働科学特別研究事業）
分担研究報告書

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

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
研究分担者 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
研究協力者 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
研究協力者 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

研究要旨

目的：研究の概要情報をもとに、生成 AI を用いて研究プロトコルを自動作成することが可能かどうか、また作成時間や労力の軽減が可能かの proof of concept (POC) 検証を行うことを目的とする。

方法：「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 技術、医学文書構造化解析技術、思考連鎖 (CoT, Chain-of-Thought) 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）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。

結果：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 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。アンケート対象者の属性として、治験/臨床研究の現場に携わる者や、統括/管理するマネージャー等、多職種の医学専門家の協力が得られた。研究プロトコルの印象として、理解のしやすさについては半数（50%）が理解しやすいと回答、また倫理的配慮については半数以上（71%）が適切と回答し、肯定的な回答が多く得られた。一方、読みやすさについては半数（50%）が読みにくいと回答、また自然さについては半数以上（72%）が不自然と回答した。全体完成度については高い・どちらともいえない・低い、何れも同程度の回答数であった。

結論：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。

A. 研究目的

治験・臨床研究ではGCP等の規定を遵守するために記載が必要な項目が詳細かつ精緻に設定されている。試験実施計画書・プロトコル、同意・説明文書、症例報告書(CRF, Case Report Form)、監査・モニタリング計画書、治験薬管理マニュアル、安全性情報報告書など研究実施に必要な文書は多岐にわたる。本研究では、過去に国立大学法人大阪大学医学部附属病院で実施された医師主導の研究プロトコルおよび特定臨床研究の研究プロトコルの概要情報をもとに、生成AIを用いて各研究プロトコルを自動作成することが可能かどうか、また作成時間や労力の軽減が可能かのproof of concept (POC)検証を行うことを目的とする。

B. 研究方法

1. 生成AIを用いた研究プロトコルの作成

「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WritingXを用いて作成した。特に、RAG 技術、医学文書構造化解析技術、思考連鎖 (CoT, Chain-of-Thought)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。尚、プロンプト・リファレンスは以下とし、リファレンスあり、リファレンスなしの研究実施計画書を作成し比較を行った。また、使用する大規模言語モデルは、OpenAI/Microsoft社のGPT-4 turboとした。

・プロンプト：

作成する研究プロトコルの以下の情報を自動抽出しプロンプトとした。

-課題名、目的、主要評価項目、副次評価項目、選択基準、研究方法（治療スケジュール等）、予定参加者数、実施予定期間、研究概要図（研究実施医療機関名）

・リファレンス：

過去に大阪大学医学部附属病院にて実施された研究実施計画書をリファレンスとした。

-胃がん患者対象研究実施計画書のリファレンス数：22件

-食道がん患者対象研究実施計画書のリファレンス数：8件

-特定臨床研究実施計画書のリファレンス数：4件

2. 医療分野の有識者へのアンケート

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）を読んだ印象等について、医療分野の有識者（14名）に対し以下の項目に関するアンケートを実施した。尚、アンケート評価者に対しては、施設基準に沿って謝礼を支払うこととした。

<アンケート者の属性情報>

(1) あなたの所属組織について

1. 臨床研究中核病院
2. 臨床研究中核病院以外

(2) あなたの職種（主な役割）について

1. 治験/試験責任/分担医師
2. CRA
3. データマネージャー

4. 生物統計家
5. メディカルライター
6. プロジェクトマネージャー/スタディー
マネージャー
7. 治験/臨床研究事務局
8. 組織マネージャー
9. その他

(3) あなたの専門分野について
フリーテキスト

<研究プロトコルの印象>

(1) 全体完成度

1. 非常に高い
2. やや高い
3. どちらともいえない
4. やや低い
5. 非常に低い

(2) 読みやすさ

1. 非常に読みやすい
2. やや読みやすい
3. どちらともいえない
4. やや読みにくい
5. 非常に読みにくい

(3) 理解のしやすさ

1. 非常に理解しやすい
2. やや理解しやすい
3. どちらともいえない
4. やや理解しにくい
5. 非常に理解しにくい

(4) 自然さ

1. 非常に自然
2. やや自然
3. どちらともいえない
4. やや不自然
5. 非常に不自然

(5) 倫理的配慮について

1. 非常に適切
 2. やや適切
 3. どちらともいえない
 4. やや不適切
 5. 非常に不適切
- (6) ご意見ご感想
フリーテキスト

(倫理面への配慮)

本研究は個人データ等を扱わないため、倫理面への配慮は不要である。

C. 研究結果

1. 生成 AI を用いた研究プロトコルの作成

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。自動作成過程について図1に示す。まず、医学文書構造化解析モジュールにより、docx, pdf などの構造化および非構造化文書、画像など、様々な形式の医学文書を解析した。章のタイトル、内容などの文書の構造情報、および選択基準/除外基準、試験日、名称などの情報を自動抽出し、プロトコル文書の生成プロセスを導入した。次に、RAG モジュールにより、プロトコル文書の自動生成における参照情報を取得した。RAG 技術を活用し、過去の類似臨床試験や類似文書テンプレート、プロトコル概要の内容やガイダンス等を組み合わせることで、各章の内容を自動生成した。RAG モジュールにより、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。そして、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。CoT は、章の内容生成、章の順序推論、参照情報の選択に適用された。

2. 医療分野の有識者へのアンケート

医療分野の有識者（14名）に対しアンケートを実施した結果について資料1に示す。

アンケート対象者の属性として、治験/臨床研究の現場に携わる者や、統括/管理するマネージャー等、多職種の医学専門家の協力が得られた。研究プロトコルの印象として、理解のしやすさについては半数（50%）が理解しやすいと回答、また倫理的配慮については半数以上（71%）が適切と回答し、肯定的な回答が多く得られた。一方、読みやすさについては半数（50%）が読みにくいと回答、また自然さについては半数以上（72%）が不自然と回答した。全体完成度については高い・どちらともいえない・低い、何れも同程度の回答数であった。フリーコメントとして得られた意見としては、AI に対する知識・理解によるばらつきがあったものの、良い印象のコメントが多く得られた。

D. 考察

AI が生成したプロトコルは、想定以上の完成度が得られている部分もあり、アンケートの結果からは、驚きと期待を持って受け止められていることが分かった。しかしながら、

後述する内容において、改善を要する部分も多く含まれていることも分かった。

1. 文体と表現の一貫性： 文体や表現の一貫性が欠けており、AI が生成する文章の自然さや読みやすさを向上させるための改善が必要と考えられた。
2. 内容の整理と明確化： 同じ内容が何度も繰り返されていることや、本研究で何を収集するのかが読みとりにくいことなどの問題点が指摘された。これは、AI が情報をより効率的に整理し、明確に伝える能力を向上させる必要があることが示唆された。
3. 倫理的配慮と同意手続き： 倫理的配慮や同意手続きに関する記述についても改善の余地があるとの意見が見られた。研究の内容に応じて、倫理的配慮と同意手続きを微調整する必要があるが、プロトコルに求められる記載内容はある程度定型化が可能と思われるため、比較的容易に改善できる可能性があると思われた。
4. AI 生成プロトコルの活用： 生成されたプロトコルそのものを固定版とすることは難しいが、それをもとに修正するプロセスが必要という認識のもとであれば、現時点でも十分に活用できると考えられた。
5. 研究背景やコンセプトの記載： 生成 AI の記載において、記載内容が不正確もしくはわかりにくい点は、研究の発案に至る思考過程や研究背景に関する内容であると思われた。特に、手術術式に関する研究に関する記載においては、生成 AI が収集可能な情報では不十分もしくは、取捨選択が困難な状況があることが推測された。具体的には、解剖学的な情報や、手術術式の多様性、施設のポリシー、歴史的変遷などの情報の取得は困難と思われた。研究者がこれ

らの情報をより重み付けして AI に提供することにより、改善が得られる可能性が示唆された。

E. 結論

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。

F. 健康危険情報

なし

G. 研究発表

1. 論文発表

なし

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

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

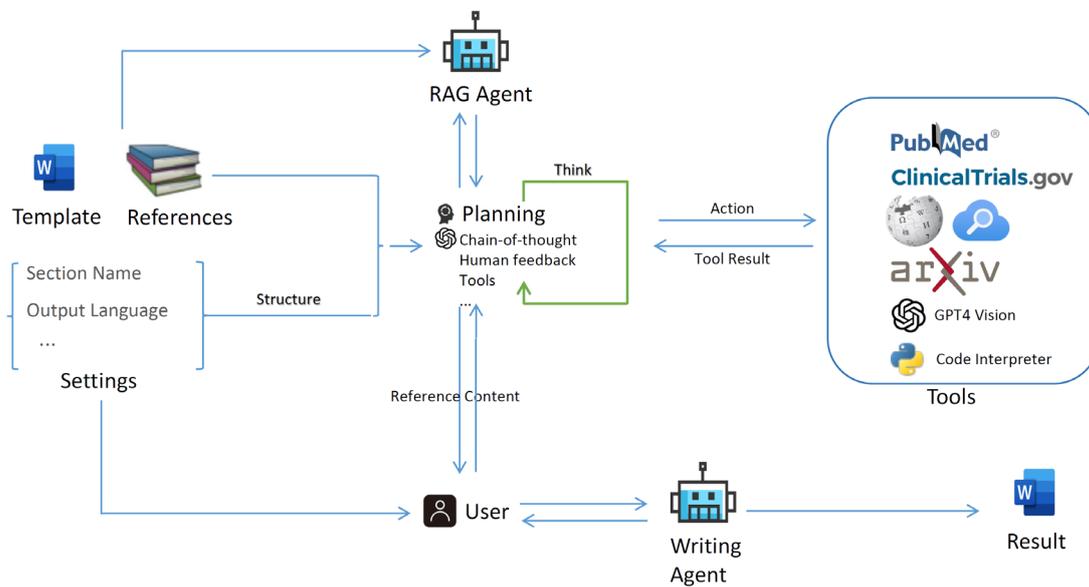
2. 実用新案登録

なし

3. その他

なし

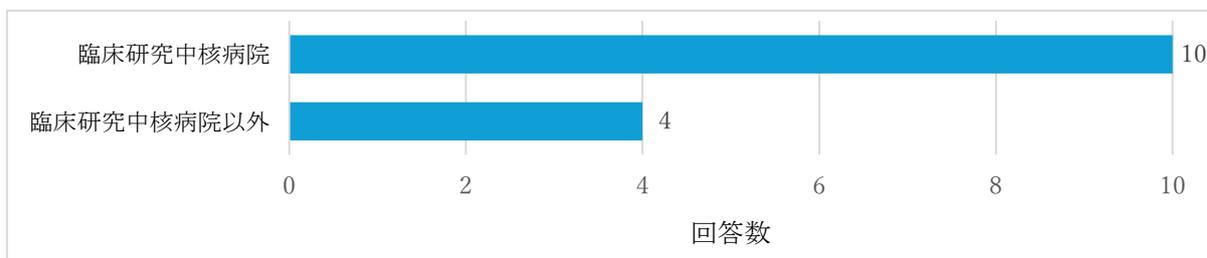
図 1. WritingX を用いた自動作成過程



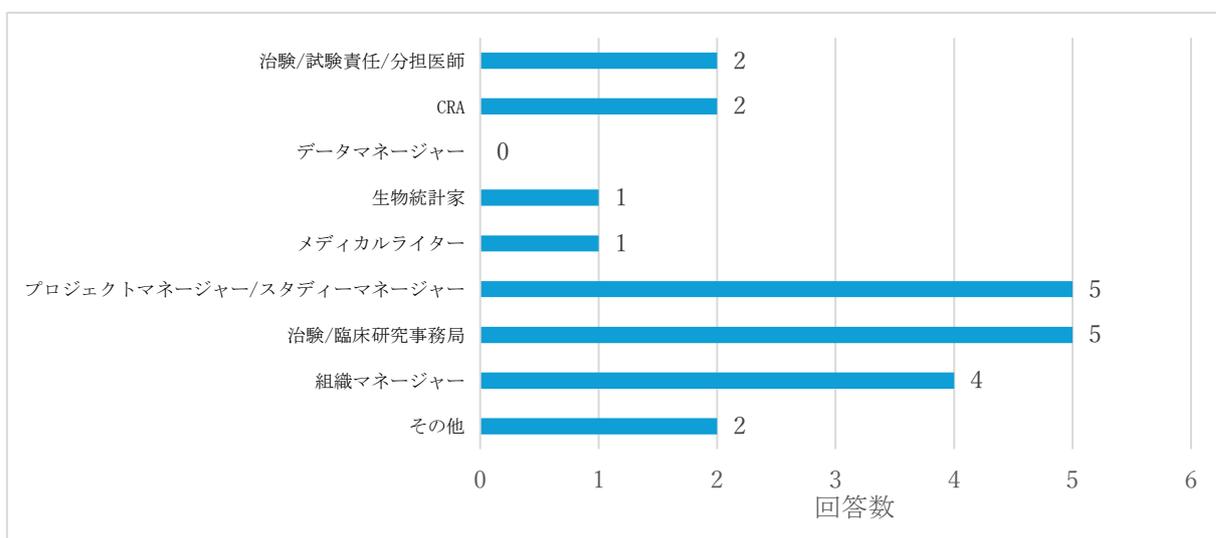
資料1 医療分野の有識者へのアンケート結果

【アンケート者の属性情報について】

1. あなたの所属組織について



2. あなたの職種（主な役割）について ※複数選択可

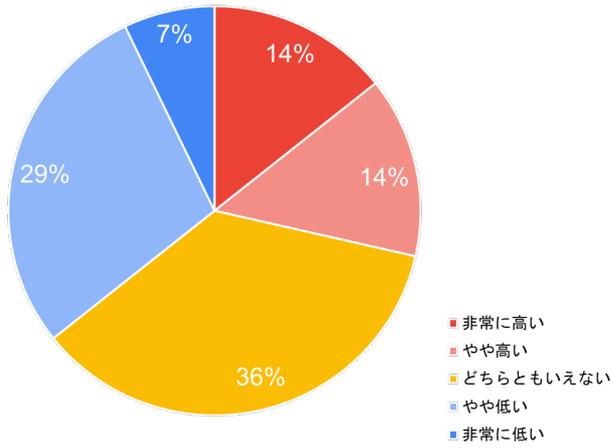


3. あなたの専門分野について

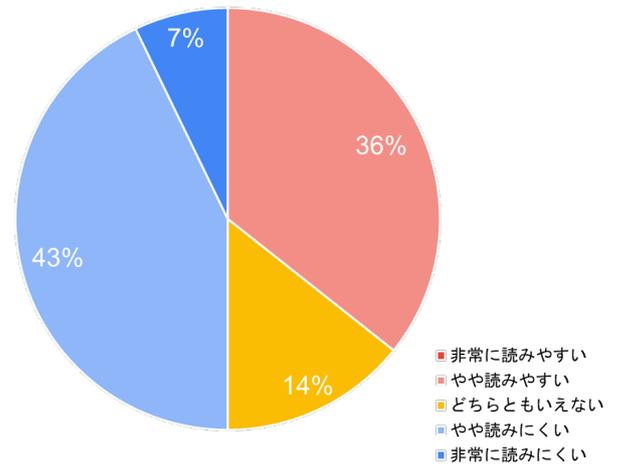
- ・レギュラトリーサイエンス、栄養、胃癌
- ・医学統計学、統計的機械学習
- ・臨床試験方法論、規制、オペレーション
- ・研究倫理、生命・医療倫理
- ・脳神経内科
- ・臨床薬理学、臨床試験管理学
- ・臨床研究支援、医療 Dx、DCT（遠隔医療）
- ・臨床研究専門職
- ・消化器内科
- ・スタディマネジメント

【研究プロトコルの印象について】

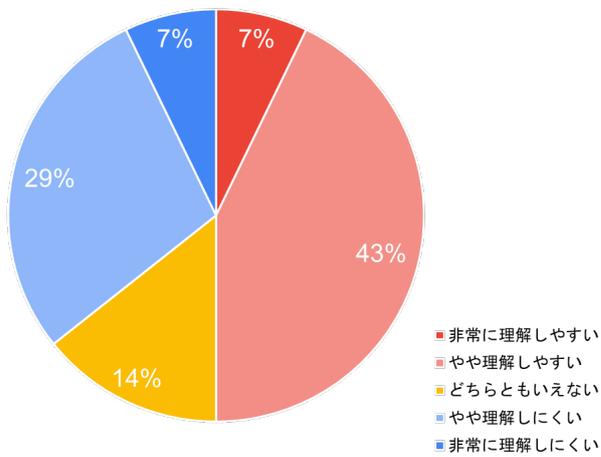
1. 全体完成度



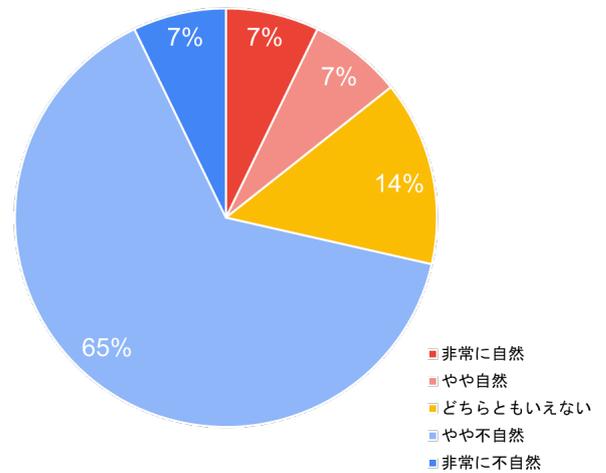
2. 読みやすさ



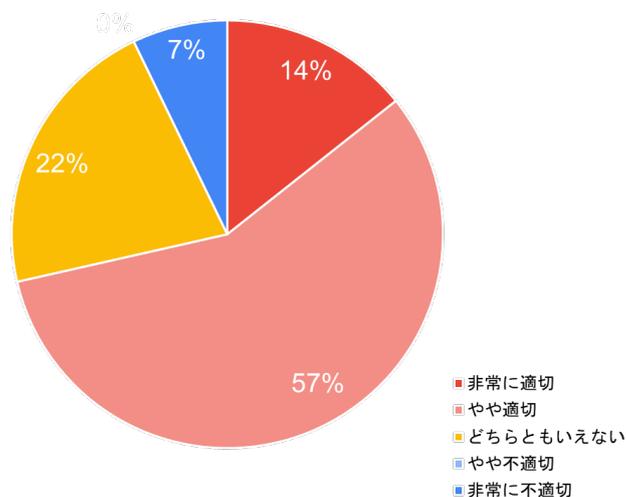
3. 理解のしやすさ



4. 自然さ



5. 倫理的配慮について



6. 意見・感想

回答 No.	内容
1	<p>幽門側胃切除後の食道癌に対して残胃を温存する術式は、残胃に血流障害が生じる可能性が高く、臨床的には不可能ではないかと思えます。そのため、プロトコルの文面というよりも医学的な観点から AI が作成したプロトコルに対する評価が少し厳しめになった可能性があります。</p>
2	<p>文章がですます調とである調が混在していることが気になりました。また、例えば、15 節の「・・・二次利用（メタアナリシスなど）が可能である。」など、プロトコルに記載する内容ではなさそうな文言があったり、逆に、Visit 等のスケジュールがわかりにくかったりしているように思います。一方で、CQ に関する内容は読み取れるなど、校閲前のプロトコルという位置づけでは、これをたたき台にすることは不可能ではないようにも思いました。</p>
3	<p>生成 AI は現時点では今回のように「膨らませる」方向よりも「まとめる」方向に活用するのが現実的。プロトコルから IC 文書を作成したり、学会スライドから lay summary を作成したり、と比較的リッチな情報源から、別の形式に変換する方が向いているのではないかと。</p> <p>例えば背景に記載すべき本試験の rationale は極めて弱く、インターネットから適切にデータを切り取って、input した仮説に対する背景を充実させるということは明らかに発展の余地がある。このあたりはプロンプトで改善できる余地があるのかどうか興味深い。</p> <p>また研究代表者や研究組織など、明らかなハルシネーションと思われる部</p>

	<p>分も散見される。</p> <p>プロトコルでの活用を考えると、改変の必要のない章はテンプレート通りに記載するようプロンプトで指定してもよいかもしれない。</p>
4	<p>全体的に、事前にインプットされた情報は、複数箇所にわたって、幾度も繰り返して用いられる状況になっているため、全体的に冗長なものとなってしまっている印象は否めない。</p> <p>しかし、研究者自身にブラッシュアップさせるための選択肢を、一定の論理性をもって提示してくれていると思えば、ゼロからプロトコルを作成することに比べて、はるかに有用なものであるようにも思われた。例えば、3ページの「5. 統計解析」や「6. 安全性評価」、5ページの「4.3 観察・調査項目」や「4.4 評価項目」の先行研究をふまえた記述は、研究者自身がプロトコルを固めるうえで有用な知見を示してくれている。</p> <p>6ページ下段以降の倫理・社会的な対応に係る記載は、概して妥当な内容であるように思われたが、やはり冗長な箇所・不必要な箇所が多く、加除修正や項目の並び替えなどを行う必要のあるところが相当あるように思われる。また、途中から、研究計画書の記載ぶりと、ICFの記載ぶりとが、混在化してしまっているように思われる箇所も多くみられた。しかし、こちらでも上記同様、研究者向けにブラッシュアップのためのガイドを示してくれていると思えばよいのかもしれない。</p> <p>もしかしたら、わが国の研究規制にある研究計画書の記載事項（例えば、生命・医学系指針第7の25項目や、臨床研究法施行規則第14条の18項目など）の項目、記載順などを含めたひな形となうような情報も（そういうことが可能なら）あらかじめインプットすることで、もうすこし自然なものになったのかもしれない。</p>
5	<p>全体を通して、生成AIのみでここまで出来るのかと驚きを感じました。「0. 概要」は他の多くの研究でみるような自然な流れになっており、完成度が高いと感じました。</p> <p>一方、細かく読み進めていくと、違和感を感じる部分が散見されました。例えば、「0. 概要」は言い切る表現で書かれていますが、「1. 研究の実施体制」からは「ですます調」になります。また、概要内の選択基準の羅列に「文書同意が得られている症例」とありますが、文書同意は選択基準を満たす症例に対してICを行い取得しますので、前後関係が逆転しています。</p> <p>その他、「4.3 観察・調査項目」には参考とした過去の研究に関する内容が多く、本研究で何を収集するのかが読みとりにくい（「4.4に評価項目ありますが、ここは整理してVisit毎にまとめたTableがあるとわかりやすい）、「17. 研究対象者等…」では唐突に遺伝カウンセリングの話が出てくる、</p>

	<p>「20. 侵襲を伴う…」には「臨床研究保険には加入しない」「臨床研究保険に加入する」両方の記載があり、結局どうするのか分かりにくい、ところどころ同じ表現が繰り返されている、など気になりました。1～4の評価はやや厳しくつけていますが、今後の改善を期待するという前向きな評価です。生成AIが医療に役立つ可能性を知ることができ、大変勉強になりました。</p>
6	<p>気づいた範囲ですが、以下コメントします。</p> <p>医学系指針を踏まえた前提でのワーディングはできていない（匿名化、などの使用）。妊娠に関する記載がない。</p> <p>ですます調とそうでないものが混ざっている</p> <p>代諾者での同意を認めていないにもかかわらず、代諾者同意に関する章が設置され、6 インフォームドコンセントとほぼ同じ内容となっている</p> <p>必要に応じて再同意、というのがリスクのところ（9章）に記載されており、要に応じて、になっているのが気になる</p> <p>10章 二次利用について、新たな二次利用と、現段階の二次利用とどちらかがわからない記載となっている</p> <p>副次評価項目で記載されているデータ取得時期に関する情報が記載されていない。また、文章で記載されていてどのようなビジットでどのような検査をするかの記載がないため、研究の再現性が非常に低いと考えられる</p> <p>被験者→研究参加者とするべきである</p> <p>倫理的配慮に関する記載が非常に不適切であり、初学者が使うことを考慮するとこのまま勧めることはできない。ただし、指導者がいる状況での下書きとしてはかなり助かるのではないかと思われる。</p>
7	<p>「過去の〇〇」、「参考とした〇〇」等の記載が散見される。引用を付与せずには余り書かない表現な気がします（chat GPT の回答の特徴がでているか?）。</p> <p>Medical Writer が記載するドキュメントより、やや単語自体の解説が多い気がしました。</p>
8	<p>一部不整合と思われる箇所もありますが、基本的な事項は概ね記載されていると思われ、大変良く作文されていて驚きました。ただ、個人的な意見ですが、基本的にすべて文章で示されているのでわかりにくいと感じました。</p>
9	<p>全般的な印象としては、シノプシス (22151-input.docx) をもとに作成しているなら、個別対応が必要な点を除いては、良い印象。</p> <p>シノプシスにない部分については、生成AI特有のハルシネーションとも思われるような箇所が多数見受けられる。</p>

	<p>また、24 項目ありますが、項目によっては「である調」と「ですます調」が混在していたり、既に記載済みの内容が何度も記載されていたり、項目内にも内容的に同じような段落が繰り返されて冗長な箇所も散見された。</p> <p>この辺りは、プロンプトを調整することで対応可能な印象がありました。</p> <p>シノプシスに記載すべき事項をチェックリスト形式でもよいので、追加することができるなら、より改善できるのではと、素人ながらに感じました。</p>
10	<p>基本的に文章でのみ記述されているので、フロー図や表などを入れられるとわかりやすくなると思われる。</p> <p>代諾者による IC を認めていないにも関わらず、余計な記述によりわかりにくくなっているように、記載内容は取捨選択できるとよいと思う。</p> <p>どの研究にも共通する記述についてはある程度のレベルにあると思われるが、とくに研究治療に関する情報など当該研究に特化した記述については、もともとの読み込ませる情報が少ない影響かもしれないが、その場合は不足情報を要求するなどレベルアップが求められる。</p> <p>生成 AI を用いたプロトコルそのものを固定版とすることはできないが、これをもとに加除修正するプロセスが必要という認識のもとであれば、十分に活用できると考える。</p> <p>今回のレビューコメントは大枠に留めるが、フォーマットをバージョンアップさせることも必要と思われる。</p>
11	<p>繰り返しの記載内容が多いこと、見出しと記載内容が少し一致していない箇所があることが気になりました。</p>
12	<p>規制が分からなかったので、項目漏れがないかのチェックは行っていません。一般的な目線で確認したコメントとなります。</p> <p>表紙も作成されますか？課題名、版数、作成者や作成日がないため気になりました。</p> <p>「0. 概要」は文書より表でプロトコルの内容が要約されているほうが読みやすいです。</p> <p>全体的に文章で記載されているため、研究の実施体制、観察・評価項目、評価項目、研究期間は特に、文章ではない記載のほうが読みやすく間違えないと考えます。</p> <p>言葉がところどころ不自然で気になりました（試験に参加する患者さん、試験代表者など）。用語も統一されていません。</p> <p>「研究の実施体制」に試験の目的や適格基準、症例数の記載が出てきたりと、文書内で同じ内容が繰り返されている印象です。</p> <p>検査・観察項目は表があるほうが分かりやすいです。</p>

	<p>同意に関する記載が不十分です（説明文書の記載事項が 6 にない、本人の署名の記載、代諾者が不在であれば 7 は不要？、金銭の支払いおよび健康被害補償の記載が 6 と 7 で異なる、など）。</p> <p>「12 研究機関の長への報告内容及び方法」については、修正が必要です（いろんな手続きがまとめて書かれているので誤解を受けそうです）。</p> <p>COI の記載も、結局利益相反があるのか無いのかが分かりにくい記載となっているため、修正が必要です。</p> <p>15 が「である調」となっています。</p> <p>18 に謝礼が支払われる場合には金額の記載が必要かと思いました。</p> <p>20 の文章に「臨床研究保険には加入しない」と「臨床研究保険に加入する」の記載があり、理解が難しいです。</p> <p>22 の記載から、結局どの業務が委託されるのかが分かりにくいです（業務内容に記載の内容が委託されるのかと思ったら、「委託先の監督方法」はモニタリングや監査の記載になっています）。</p> <p>23 の記載で、モニタリングは中央とサイトの組み合わせとありますが、計画を立てるのは「試験代表者、統計担当者、データマネージャー」となっていて CRA の記載がありません。</p>
13	<p><レビューに当たり不明だった点></p> <p>word の限られたインプット情報のみから、除外基準や必要な手順等を AI が作成したのか。</p> <p>規制は治験？ 特定臨床研究？</p> <p><意見></p> <p>（気付いた点をその都度挙げていきましたので、とても細かくなってしまいました。）</p> <p>プロトコルの記載順序の検討が必要。研究の流れとして理解しづらい。</p> <p>重要な評価項目は「など」を使用せず、明確にすべきではないか。</p> <p>同じ内容が何度も繰り返し記載されている。まとめて記載できる内容が多いと思うが、何度も記載する必要があるのであれば、項によって記載を工夫してはどうか。</p> <p>文章・箇条書きを使い分け、読みやすく・理解しやすくした方がよい。</p> <p>用語の統一性なし。規制用語に従って使用するのがよい。</p> <p>聞きなれない表現が時々出てくるため分かりづらい。</p> <p>「監査」の使い方が紛らわしい。</p> <p>介入を実施する症例は適格基準に合致しているため、適格基準に記載されている内容を記載する必要はない。</p> <p>プロトコルに記載する必要のない内容（不明な内容）がある。</p>

	<p>除外基準 3 が不明。</p> <p>プロトコル内で齟齬がある。</p> <p>「同意の時期と方法」の記載内容には違和感がある（説明不足？）</p> <p>「7. 代諾者等からの IC を受ける場合の手続き」について、代諾者による同意を認めないのであれば、最初の「代諾者から IC を得る手続きを以下の通り定めます」は不要。紛らわしい。</p> <p>プロトコルの途中から「です・ます」になっている。</p> <p>画像データは生物学的試料ではない。</p> <p>「12. 研究機関の長への報告内容及び方法」は定期報告と発生時報告を明確に分かりやすく記載した方がよい。</p> <p>データセンターがどこまでの役割を担っているのかわからなくなる。</p> <p>モニタリング手法が「中央」、「中央+サイト」の両方記載されている。</p> <p>モニタリング計画の策定にはモニターも必要。</p> <p>評価項目の収集プロセス・具体的な収集データが記載されていない。</p>
14	<p>参考文献の記載（肩番号による紐づけ）が無い場合、本研究のデザインや除外基準設定の科学的合理性について妥当と判断する決定的根拠に欠ける。</p> <p>倫理的配慮については、生成 AI に稀に見られる重大な誤謬は見受けられないと考える。（が私が見落としていたかもしれないので、重大な誤謬が潜んでいた場合後日教えていただけると嬉しいです）</p> <p>以下は些細なことです。</p> <ul style="list-style-type: none"> ・実施体制や評価項目を箇条書きで記載してほしい ・です・ます体とである体の混在 ・項目題と内容があっていない（実施体制に目的の記載がある等） ・同じ文章の繰り返し

厚生労働科学研究費補助金（厚生労働科学特別研究事業）
分担研究報告書

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

研究分担者 川崎良 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 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

研究協力者 大谷まゆ 株式会社サイバーエージェント

研究協力者 張培楠 株式会社サイバーエージェント

研究要旨

目的:現代の医療の重要な要素として、患者が自ら受ける医療について理解し、判断し、納得することが重要である。臨床研究の研究説明文書を題材として、大規模言語モデルによる平易化されたインフォームド・コンセント用説明文書を平易化すること、そして、文章の平易化によって患者視点で説明内容の理解度、信頼度が維持・向上するかを評価する研究を行なった。

方法: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。

結果: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。

結論:研究協力におけるインフォームド・コンセント文書は一般に長く、難解であることが多い。大規模言語モデルによる文書平易化によって理解度を向上させつつ、同意の意向には大きな影響を及ぼさないことが示された。また、年代によって平易化文書による理解度への影響が異なる可能性も示され、年代ごとに提示する文書の型を変えるなどの個別化によってよりきめ細やかな理解度向上を得る可能性もあると思われた。

A. 研究目的

現代の医療の重要な要素として、患者が自ら受ける医療について理解し、判断し、納得することが重要である。そのプロセスは、患者が医療を受ける際に、① まず検査

や治療についての情報を十分に知ること、② その情報（医療の内容、期待される利益と不利益）を適切に理解すること、③ その上で自ら判断し決断を下すこと、から構成され、ヘルス・リテラシーの目指すべき形と言える。このプロセスは、現代の臨床に

においてはインフォームド・コンセント (IC, informed consent) のプロセスとして実装されていると言える。

例えば、検査や治療について、あるいは臨床研究への参加について、医療従事者は文書を用いて患者に対してその内容を説明し、十分な理解を得たことを確認し、本人の同意を確認して医療や研究を進めていく。そのためにはICで用いる説明文書が重要な役割を担っている。その一方、現場で用いられている説明用文書で説明を受けた患者の60%はその内容が十分に理解できていないという報告や、文書のわかりにくさがその後の医療や研究参加に対する不安感の助長、信頼感の低下につながっていることが報告されている。このことは、医療の現場で限られた時間に十分な医療・研究について説明することがいかに難しいかを物語っている。近年、人工知能 (AI, artificial intelligence) 技術の著しい発展と応用によりその応用範囲が医療にも及んでいる。自然言語を解析し、生成するいわゆる大規模言語モデル (LLM, Large Language Model) が開発され、さまざまな用途で用いられている。医療における種々の文書の生成、情報抽出、要約、言い換え、平易化もその応用の一つである。ICで用いる説明文書についてもLLMを用いて、その内容を分かり易くすること、その際には同時に内容の理解度を維持あるいは向上させることができるか、また、根拠に乏しい情報、誤った情報が混入することがないか、流動的な改変文書について誰がどのように責任を取ることができるか、技術的な議論と倫理的・法的・社会的課題の議論の双方が活発化している。

今回、医療現場で用いられるIC説明用文書を、LLMで平易化することの効果を技術的な評価とヘルス・リテラシーの醸成への寄与の評価を目的として研究を行った。具体的には、大阪大学医学部附属病院で行われ公開されている臨床研究の研究説明文書を題材として、LLMによる平易化されたIC説明用文書を平易化すること、そして、文章の平易化によって患者視点で説明内容の理解度、信頼度が維持・向上するかを評価する研究を行った。

B. 研究方法

- (1) 説明文書の選定：本研究では患者説明用の文書として「全ゲノム情報等の高精度かつ迅速な患者還元および新たな創薬等の創出を通じた高度化がんプレジジョン医療の実践」(大阪大学医学部附属病院観察研究等倫理審査委員会 承認番号:22195-6; 研究責任者:大阪大学大学院医学系研究科消化器外科学・教授 土岐祐一郎) (2023年7月8日 第5版) を元文書として用いた。選定の理由としては、複数の診療科に関連する内容であること、治療に関する内容では患者当事者でなければ判断できない要素が関与するため研究的な内容であること、そして、内容の理解が比較的難しい研究協力の依頼であること、から選定した。この説明文書全文は23ページに及ぶため、本検証では「7.様々な目的での利用」(3頁)を用いることとした(資料1)。
- (2) LLMを用いた説明文書の平易化：元の説明文書をChatGPT4 (OpenAI, USA)

および cyberagent/calm2-7b-chat (サイバ
ーエージェント, 日本) の二つの LLM モデ
ルを用いて平易化した。平易化において
は、{対象：一般人、10代、若い人、高齢
者、お年寄り、医学的知識のない方向け
に、}, {平易化指示：分かりやすく、簡潔
に、要点をまとめて、理解しやすいよう
に}, {文字数：1000 字程度で} の指示語の組
み合わせでプロンプトを作成し、それぞ
れのプロンプトを用いて平易化文書を 100 編
作成した。作成された平易化文書につい
て、医学生 3 名が以下の観点から 5 段階評
価を行い、6 編に絞り込んだ。

A) 明らかに誤った情報が含まれていない
か？

1. 明らかな誤りがある (1 点)
2. 誤りの可能性がある (2 点)
3. 誤りはない (3 点)
4. 概ね誤りはない (4 点)
5. 全く誤りはない (5 点)

B) 理解しやすいか？

1. 全くわからない (1 点)
2. ややわかりにくい (2 点)
3. わかりやすい (3 点)
4. とてもわかりやすい (4 点)
5. 大変わかりやすい (5 点)

C) 難しい語句があるか？

1. たくさんある (1 点)
2. 結構ある (2 点)
3. ある (3 点)
4. 少しある (4 点)
5. 一つもない (5 点)

D) 原文を正確に表現しているか？

1. 全く正確ではない (1 点)
2. あまり正確ではない (2 点)
3. 正確である (3 点)
4. とても正確 (4 点)
5. 非常に正確 (5 点)

E) 内容を人に説明できるか？

1. 全く説明できない (1 点)
2. あまり説明できない (2 点)
3. 説明できる (3 点)
4. 概ね説明できる (4 点)
5. 十分に説明できる (5 点)

日本語文章の難易度を示す指標として、日
本語文章難易度判別システム

(<http://jreadability.net/>) を用いた指標化
を行なった。

(3) 医療従事者による平易化説明文書の内
容評価：上記で選別された 6 編について、
医師 7 名、看護師 4 名が医学的な内容の誤
りがないか、そして、上記の評価項目に従
って採点を行なった。

(4) 一般人による平易化された説明文書の
評価：ネット調査を手掛ける委託業者 (イ
ンテージ社) の一般人からなるパネル (対
象者年齢群：15-19 歳、20 代、30 代、40
代、50 代、60 代、70 代、80 代；各 500 名
程度を目標にリクルートを行なった。用意
した原文、平易化された IC 説明用文書 1-
6 の計 7 編を年齢、性別ごとにランダムに
提示した。回答時間に制限は設けなかつ
た。本文を提示後、資料 2 に示す項目につ
いて回答を求めた。

(5) 説明文書生成を可能にするチャットボットの試作：個別化された説明を提供する形のプロトタイプとして、患者向けの説明文書をもとに質問に答えるチャットボットを試作した。大阪大学医学部附属病院で用いている胃がん手術に関する情報を外科医として回答に答えるもので、外部情報の検索を制限し、大阪大学で用いている説明文書のみを参照しつつ回答を生成するように設定した。

C. 研究結果

原文書をもとに作成した 100 編の生成文章のうち、医学生による予備評価によって選定された 6 編を資料 3 から資料 8 に示す。

表 2 には、使用した LLM、文字数、日本語文章難易度判別システムスコアと文章難易度、文章の形式を示す。文字数は 800–1000 字程度で A4 一枚に収まる程度の分量となった。文章の体裁としては、原文のような説明文型に加え、箇条書き型、Q&A 型、会話型などが生成された。これら 6 編について、まず医療従事者が採点した評価の分布を図 1 に示す。平易化 4 が他の文章に比べてスコアが低かった。ネットパネル調査においては、年齢・性の層毎にランダムに原文か平易化文書 6 編のいずれかが提示されることとした。表 1 には性別・年齢層別の回答者人数の分布を示す。性別及び年齢層については均等に回答が得られた。資料 3 から資料 8 を提示した上での回答をもとに 5 段階評価を 1 点から 5 点として平均スコアを算出したものを図 2 に

示す。粗スコア値においては、原文に比べて、平易化 2、平易化 3 については有意な差はなく、平易化 5 は「目的」について、平易化 6 は「国家戦略」についての理解度が悪い方向に、平易化 4 と平易化 7 は「全体の理解度」と「他人へ説明できる」については理解度の向上の方向に差がみられた。

文章の理解においては、年齢や性別、社会経済的要素として居住都道府県、職業（学生含む）、世帯収入などが影響する可能性を考え、これらの変数で調整した上での原文に対する平易化文章の理解度スコアの差を算出したものが表 3 である。その結果、「国家戦略についての理解度」については原文には明記されているものの、平易化の過程で省略されているか明記されているとは言えない内容であることを反映して、理解度が低くなっていた。また、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 5 以外では、全体の理解度、他の人に内容を説明できるなどのスコアが向上していた。逆に言えば、平易化 5 については、原文に比べてかえって理解度が低い結果となった。平易化 5 については、文章の日本語文章難易度判別システムスコアが他の平易化文書に比べて低いこと、また、文書形式が説明文形式であることなどが違いとして示唆された。

図 3 には年代別にスコアを文字数順、日本語文章難易度判別システムスコア順、そして、文書形式別に示した。暖色系は理解度スコアが低下する方向に、寒色系は理解度スコアが向上する方向にあったことを示している。年代別に見ると平易化文

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方向にあることが示唆された。

このように、年代によって平易化文書による理解度への影響が、理解度を上げる方向と下げる方向とに分かれているとすれば、全体としての研究の理解度を向上させるためには、年代ごとに提示する文書を変える、個別化することで全体の理解度を助けるかは今後検証が必要であると考えている。

個別化された研究説明あるいは臨床情報提供の例として、大阪大学医学部附属病院が提供している情報をもとに回答を提供するチャットボットのプロトタイプを作成した（大阪大学医学部附属病院患者向けの胃がん手術説明文書に基づいたチャットボット：<https://chat.openai.com/g/g-TL4Ks53mu-ban-da-hua-zi> (ChatGPTs を用いて作成)；<https://huggingface.co/spaces/peinan/pdfchat> (cyberagent/calm2-7b-chat))。これらのチャットボットで用いた題材は胃がんについての手術説明書であり、外部検索機能を用いず、提供された手術説明書に基づいて回答をすることを指示したものとしている。今回の研究期間中には検証には至らなかったが、このチャットボットについての評価も今後検討していきたい。

D. 考察

今回、医療現場で用いられる研究に関する 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を明らかにすることができた。

医療従事者による評価において評価が低かった平易化 5 についてはネットパネル調査での理解度において原文に比べて理解度が向上することはなかった。かえって目的の理解度においては理解度が低い方向に影響している可能性もあり、これが文章の難易度によるものか、あるいは、文章の体裁（説明文）によるものかについてはさらに検討する余地がある。

また、年代によって好まれる文書に特徴があることが示され、特に 10 歳代から 40 歳代と 80 歳代は平易化文書が理解度を高める方向にある傾向であるのに対して、50 歳から 70 歳代においてはかえって平易化が理解度を低下させている可能性もあり、年代に合わせて好まれる平易化文書の型がある可能性も示された。本来であれば患者に対して丁寧な説明を行う上で、個別にその理解を助けることができれば理想であるが、それが LLM の助けを借りて情報提供を行う可能性が示唆された。

E. 結論

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。

- <https://huggingface.co/spaces/peinan/pdfchat> (cyberagent/calm2-7b-chat を用いて作成)

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

準備中

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

- 大阪大学医学部附属病院患者向けの胃がん手術説明文書に基づいたチャットボット作成例：
<https://chat.openai.com/g/g-TL4Ks53mu-ban-da-hua-zi> (ChatGPTs を用いて作成)

図 1. 生成された平易化文書に対する医療従事者による評価結果

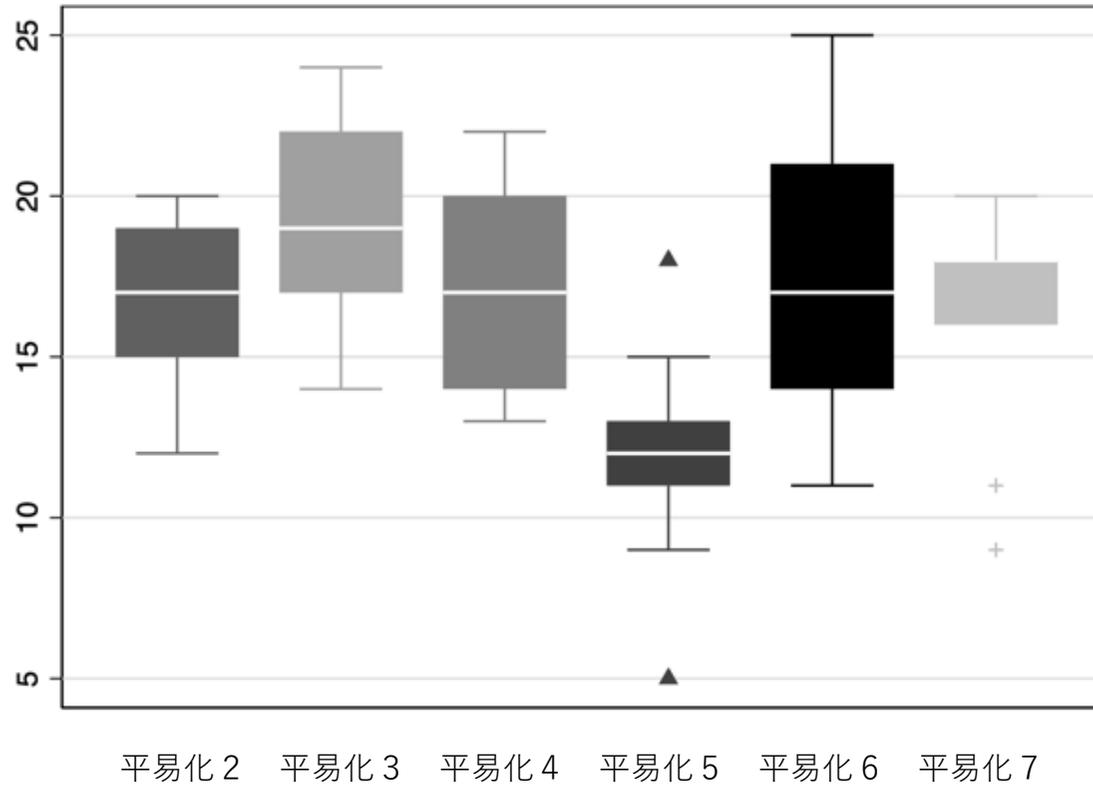


図2. ネットパネル調査での理解度スコア平均値(粗集計)

(* :原文に対して理解度低い; #:原文に対して理解度高い; $p < 0.05$ with One-way ANOVA with Bonferroni correction)

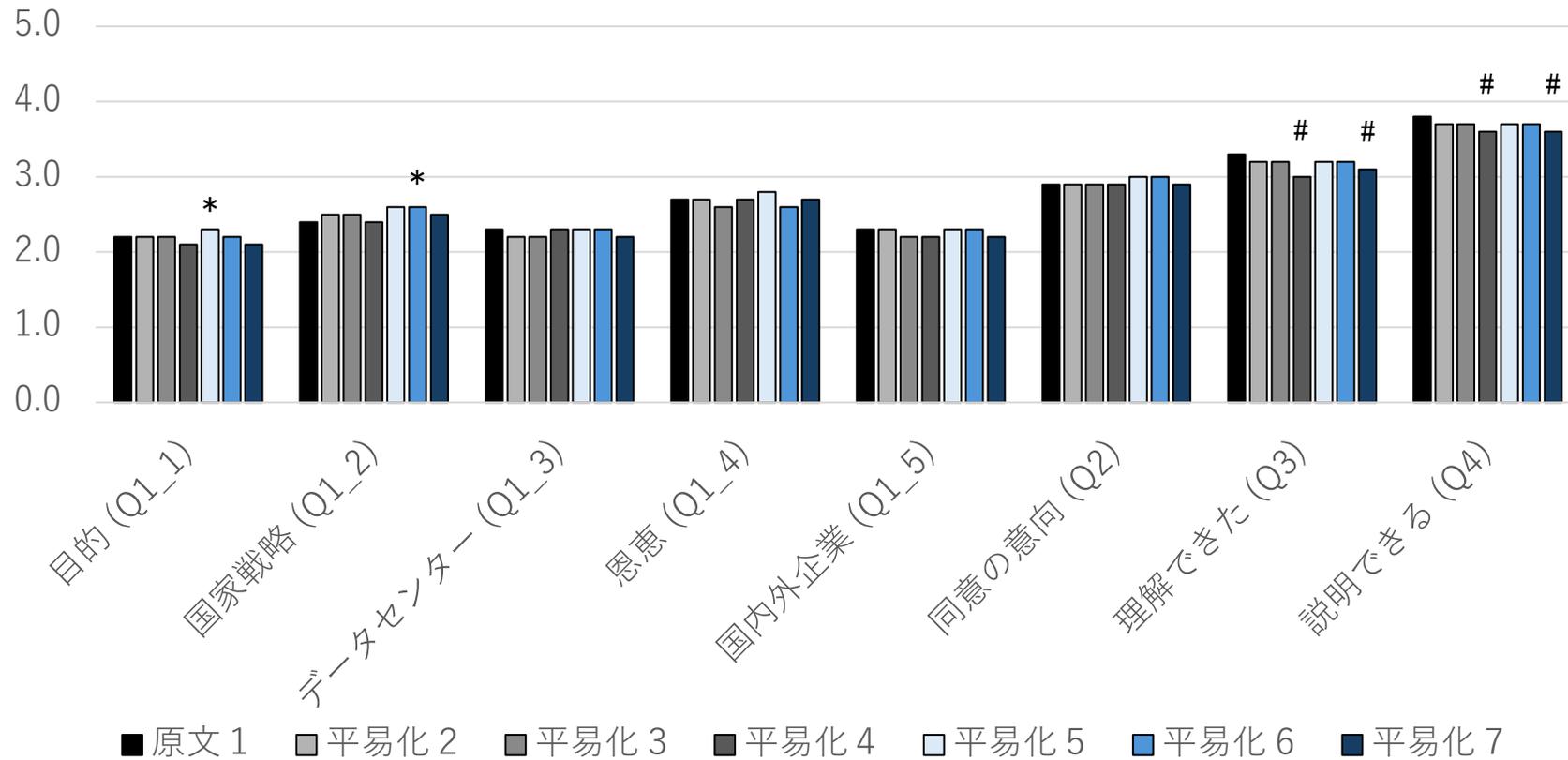


図3.年代別かつ文字数、リーダビリティ、文書形式順に見た理解度スコア平均値のヒートマップ
 (暖色系は原文に比べて理解度が低い方向に、寒色系は原文に比べて理解度が高い方にあることを示す。)

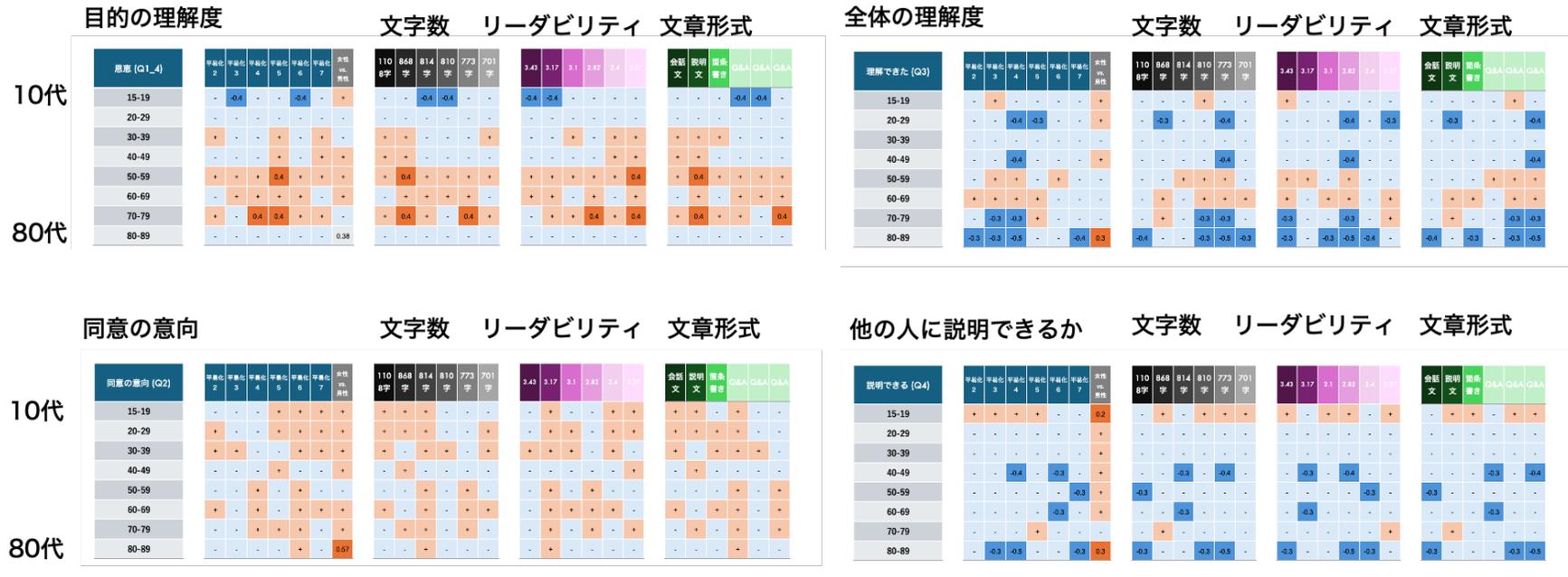


表1. 対象者年齢・性別の分布

		原文1	平易化2	平易化3	平易化4	平易化5	平易化6	平易化7	p-value
N	4193	601	596	603	597	606	598	592	
性別	男性	314 (52.2%)	304 (51.0%)	288 (47.8%)	310 (51.9%)	327 (54.0%)	312 (52.2%)	326 (55.1%)	0.26
	女性	287 (47.8%)	292 (49.0%)	315 (52.2%)	287 (48.1%)	279 (46.0%)	286 (47.8%)	266 (44.9%)	
年代	15-19	72 (12.0%)	72 (12.1%)	72 (11.9%)	73 (12.2%)	74 (12.2%)	72 (12.0%)	73 (12.3%)	1.00
	20-29	72 (12.0%)	71 (11.9%)	73 (12.1%)	71 (11.9%)	73 (12.0%)	71 (11.9%)	71 (12.0%)	
	30-39	75 (12.5%)	78 (13.1%)	78 (12.9%)	76 (12.7%)	79 (13.0%)	75 (12.5%)	72 (12.2%)	
	40-49	76 (12.6%)	75 (12.6%)	78 (12.9%)	75 (12.6%)	76 (12.5%)	76 (12.7%)	75 (12.7%)	
	50-59	77 (12.8%)	75 (12.6%)	78 (12.9%)	76 (12.7%)	76 (12.5%)	75 (12.5%)	73 (12.3%)	
	60-69	76 (12.6%)	76 (12.8%)	74 (12.3%)	78 (13.1%)	78 (12.9%)	76 (12.7%)	78 (13.2%)	
	70-79	78 (13.0%)	75 (12.6%)	75 (12.4%)	75 (12.6%)	74 (12.2%)	76 (12.7%)	76 (12.8%)	
	80-89	75 (12.5%)	74 (12.4%)	75 (12.4%)	73 (12.2%)	76 (12.5%)	77 (12.9%)	74 (12.5%)	

表 2.平易化されたファイルの文字数、難易度

	原文1	平易化2	平易化3	平易化4	平易化5	平易化6	平易化7
平易化に使用した LLM	-		ChatGPT4			calm2-7b-chat	
総文字数	3447 字	701 字	810 字	773 字	868 字	814 字	1108 字
文章難易度	上級前半	中級後半	中級後半	中級後半	上級前半	中級後半	上級前半
リーダビリティ値 [日本語文章難易度判別システム (jreadability.net)]	1.70	3.10	3.43	2.82	2.27	3.17	2.40
文書形式	説明文	箇条書き	Q&A	Q&A	説明文	Q&A	会話文

表3. 原文に対する平均スコア(年齢、性別、居住都道府県、職業(学生含む)、世帯年収で調整した多変量解析)

	目的についての理解度 (Q1_1)				国家戦略についての理解度 (Q1_2)				データセンターについての理解度 (Q1_3)				恩恵についての理解度 (Q1_4)			
	β	95%CI	p 値		β	95%CI	p 値		β	95%CI	p 値		β	95%CI	p 値	
原文 1	(基準)	-	-		(基準)	-	-		(基準)	-	-		(基準)	-	-	
平易化 2	0.04	-0.06	0.13	0.48	0.11	0.01	0.21	0.03	-0.02	-0.12	0.08	0.70	-0.01	-0.12	0.10	0.91
平易化 3	0.04	-0.06	0.14	0.40	0.09	-0.01	0.19	0.09	-0.05	-0.15	0.06	0.37	-0.10	-0.21	0.01	0.07
平易化 4	0.01	-0.09	0.11	0.88	0.05	-0.05	0.15	0.31	0.02	-0.09	0.12	0.76	0.02	-0.09	0.13	0.72
平易化 5	0.19	0.09	0.29	<0.01	0.16	0.06	0.26	<0.01	0.00	-0.10	0.11	0.95	0.10	-0.01	0.21	0.08
平易化 6	0.06	-0.04	0.16	0.25	0.21	0.11	0.31	<0.01	0.05	-0.06	0.15	0.39	-0.10	-0.21	0.01	0.07
平易化 7	-0.02	-0.12	0.08	0.70	0.13	0.03	0.23	0.01	-0.08	-0.18	0.03	0.15	-0.02	-0.13	0.09	0.74
	国内外企業利用について(Q1_5)				同意の意向 (Q2)				全体の理解ができた (Q3)				他の人に説明できる (Q4)			
	β	95%CI	p 値		β	95%CI	p 値		β	95%CI	p 値		β	95%CI	p 値	
原文 1	(基準)	-	-		(基準)	-	-		(基準)	-	-		(基準)	-	-	
平易化 2	0.00	-0.11	0.10	0.96	0.00	-0.11	0.12	0.95	-0.14	-0.23	-0.04	0.01	-0.12	-0.22	-0.02	0.02
平易化 3	-0.07	-0.17	0.04	0.20	-0.06	-0.17	0.06	0.32	-0.13	-0.23	-0.03	0.01	-0.15	-0.25	-0.05	<0.01
平易化 4	-0.04	-0.14	0.07	0.46	-0.05	-0.16	0.07	0.42	-0.24	-0.33	-0.14	<0.01	-0.19	-0.28	-0.09	<0.01
平易化 5	0.02	-0.09	0.12	0.76	0.04	-0.07	0.16	0.44	-0.07	-0.16	0.03	0.19	-0.05	-0.15	0.05	0.29
平易化 6	0.08	-0.03	0.18	0.15	0.05	-0.06	0.16	0.38	-0.11	-0.20	-0.01	0.03	-0.14	-0.24	-0.04	0.01
平易化 7	-0.10	-0.20	0.01	0.07	-0.04	-0.15	0.07	0.46	-0.18	-0.28	-0.08	0.00	-0.20	-0.30	-0.10	<0.01

β : 原文の平均スコアに対する差(負の方向が理解度が高いことを示す) 95%CI: 95%信頼区間

(資料1) 原文1

「全ゲノム解析等実行計画」では、ご提供いただいた試料・情報を用いて「解析・データセンター」に全ゲノム等解析データおよび臨床情報のデータベースを構築し、国内外の研究者や企業の間でデータを広く共有して診療や研究・創薬等に活用するための仕組みを構築します。この仕組みは、日本の患者さんや市民の皆さんのゲノム医療を日本に住む方々のデータを使って実現するための基盤であるとお考えください。そのため、以下の目的で多様な疾患に関する研究・創薬等にデータを活用させていただきます。①全ゲノム解析等の成果を患者さんに還元すること②新たな個別化医療等を実現し、日常診療への導入を目指すこと③全ゲノム解析等の結果を研究・創薬などに活用すること④上記①～③に関わる人材の育成や保健医療政策の検討を行うこと本研究では、データの適正な利用を確保するため、関連法令・指針、AMED との契約および「全ゲノム解析等実行計画」に基づいてデータ共有ルールを定め、第三者も加わったデータ利活用審査委員会が公正な立場からデータの利用を審査・監督する仕組みを設けます。データの利用者には、関連法令・指針、データの利用条件、およびセキュリティに関するガイドラインの遵守が義務づけられます。ゲノム研究は急速に進歩しており、今後さらに研究が進むと、現時点では想定されていない新たな研究に試料や情報を活用させていただくことが必要となる可能性があります。その場合には、新たに研究計画を作成し、関連法令・指針に基づき、データ利活用審査委員会や研究機関の倫理審査委員会による審査で改めて承認を受けた上で研究に活用させていただきます。なお、AMED 研究開発データ統合利活用プラットフォームを通じてデータが利活用される場合があります。その場合も、上記に準じて取り扱われます。「全ゲノム解析等実行計画」ではデータの利活用に関する状況をできる限り公開し、透明性の確保に努めます。本計画で解析・収集されたデータは大変に貴重なものです。そのため、様々な目的で使わせていただきたいと考えております。以下では、その用途として考えられる事項をご説明します。① 将来の研究目的での利用医療機関や研究機関の努力だけで、新しい医薬品や診断技術等を開発することは困難です。また従来、こうした製品・技術の開発に携わる企業が、日本の患者さん等のゲノムデータを収集・解析する環境が整備されておらず、課題とされてきました。そのため、「全ゲノム解析等実行計画」では、研究機関・医療機関だけでなく、医薬品や医療機器等の研究開発に携わる国内外の企業が、研究、医薬品等の開発、科学的なエビデンスに基づく予防等の目的で、データを利用できる環境を整備します。これにより、国内での研究開発を促進することが期待されます。現在はまだ計画されていませんが、今後、さらに研究が進むと、新たな研究が立案される可能性があります。その場合に、あなたからいただいた試料や情報を活用させて頂く可能性があります。その際、国内外の機関（民間企業を含む）との共同研究になる可能性があるほか、共同研究先にあなたから頂いた試料や情報を提供する可能性があります。具体的には、以下のような形でデータを利用させていただきます。a) 「全ゲノム解析等実行計画」に参画する研究機関・医療機関およびこれらの機関との共同研究を行う機関による利用本

研究の実施機関を含む多数の医療機関・研究機関が「全ゲノム解析等実行計画」に参画し、データの収集や解析を行います。得られたデータを効率的に解析するために、参画機関が共同でデータを利用して診療、研究・創薬等を行います。また、参画機関との共同研究を行う機関（企業を含む）がデータを利用させていただく場合もあります。利用を希望する医療機関や学術研究機関は、具体的な研究計画を立案した後、倫理審査委員会の承認を得て、データの利用審査を受けることになっています。利用に際しては、データ利活用審査委員会への報告が義務づけられます。b) 上記 a)以外の機関による利用上記 a)以外の機関は、有償でデータを利用することが可能です（おもに企業による利用が想定されています）。データの利用を希望する研究機関、医療機関、企業等は、具体的な研究計画を立案し、倫理審査委員会の承認を得るとともに、データ利用を申請し、データ利活用審査委員会による審査を受けます。審査により、適切な利用を行うと認められた場合にのみデータの利用が許可されます。ただし、医薬品や医療機器等の開発に携わる企業は、それぞれ競争をしながら開発を進めており、データ利用の詳しい目的や目標等について詳細な情報を公開することはできません。しかし、できる限りの情報を公開し、透明性を担保いたします。

② 国内外の公的データベースへの登録と活用本研究と類似した研究をする無駄を防いだり、解析するデータの量や規模を大きくして研究結果の信頼性を高めるために、公的なデータベースにデータを提供する取り組みも進んでいます。解析されたデータが、データベースを介して他の研究者に利用されることによって、新しい技術の開発が進むとともに、様々な病気について、今まで不可能であった疾患の原因の解明や治療法・予防法の確立が早まる可能性が期待できるためです。「解析・データセンター」で保管されるあなたのデータのうち、全ゲノム等解析データおよび一部の臨床情報は、公的な研究用データベース（国外のものを含む）にも提供・登録されます。公的データベースに登録されたデータは、日本国内の研究機関に所属する研究者だけではなく、国内外の製薬企業等の民間企業や海外の研究機関に所属する研究者も利用する可能性があります。将来、どの国の研究者から利用されるか、現時点ではわかりません。しかし、どの国の研究者に対しても、その国の法令に沿って作成されたデータベースのガイドライン等に準じた利用が求められます。データベースに登録されたデータの利用には、データの種類によって異なるアクセスレベル（制限公開、非制限公開）が設定されます。個人の特定につながらない情報（頻度情報・統計情報等）は非制限公開データとして不特定多数による利用が可能な形で提供され、個人ごとのゲノムデータ等は制限公開データとして、科学的観点と研究体制の妥当性に関する審査を経た上で、データの利用を承認された研究者のみに利用されます。現在、登録を予定しているデータベースには、代表的なものとして以下が挙げられます。a) 国立研究開発法人科学技術振興機構（JST）が運営する、バイオサイエンスデータベースセンター（NBDC）のヒトデータベース NBDC ヒトデータベースでは、インターネットのウェブサイトを通じてデータが公開されています。NBDC にデータを登録する際には、データを 2 種類に分けて取り扱います。（ア）多くの方のデータを集計した結果

は、研究者一般に広く公開します。(イ) 個人ごとの詳しいデータについては、科学的観点と個人情報保護のための体制などについて厳正な審査を受けて、承認された研究者にのみ、利用を許可します。NBDC では、データを登録する研究者、および、データを利用する研究者の双方に、 厳格なセキュリティ環境の整備を義務づけています。NBDC の運営状況や NBDC から公開されているデータの内容をお知りになりたい場合には、下記のウェブサイトをご覧ください。NBDC ヒトデータベース

<http://humandbs.biosciencedbc.jp/>③ 人材育成や保健医療政策など公益性の高い目的での利用ゲノム医療の発展には、大規模なデータを解析できる人材の育成が急務であり、これらのデータを使用した実習や研修などに使わせていただきたいと考えています。また、これらのデータは、国や地域が立案する保健医療の政策の検討においても、大きく役立つものだと考えられます。ぜひ行政機関やその依頼を受けた研究者に使わせていただきたいと考えています。上記以外にも、現在はまだわかりませんが、将来、医療をよりよくするために公益性の高い目的で使用させていただく可能性があります。こうした場合にも、利用を希望する者に対しては、利用目的を明らかにしたうえで審査を受けて頂くとともに、データ利用とセキュリティに関するガイドラインを遵守していただきます。なお、本研究にご協力を頂くにあたっては、本計画で解析・収集されたデータを上記①～③でお示している全ての目的で利用することにご同意頂く必要があります。

(資料2)パネル調査で用いた理解度調査票

次の文章は「ゲノム医療の実現を目指した研究への参加のお願い」の文章の一部です。この文章を読んで以下の質問に回答してください。

「文章の提示」

設問1 この文章の内容について、以下の質問に回答ください。

- ① がんや難病の早期診断や治療法開発に貢献することが期待されている。
- (1) そう思う。
(2) ややそう思う。
(3) どちらでもない。
(4) あまりそう思わない。
(5) そう思わない。
- ② 国家戦略としてゲノム医療が推進されている。
- (1) そう思う。
(2) ややそう思う。
(3) どちらでもない。
(4) あまりそう思わない。
(5) そう思わない。
- ③ 協力することによってあなたの情報は解析・データセンターに保管される。
- (1) そう思う。
(2) ややそう思う。
(3) どちらでもない。
(4) あまりそう思わない。
(5) そう思わない。
- ④ 協力することによって、あなたの治療に直接的な恩恵がある。
- (1) そう思う。
(2) ややそう思う。
(3) どちらでもない。
(4) あまりそう思わない。
(5) そう思わない。
- ⑤ 協力することによって、あなたの解析結果は国内外の企業が利用する可能性がある。
- (1) そう思う。
(2) ややそう思う。
(3) どちらでもない。
(4) あまりそう思わない。
(5) そう思わない。

設問2 この文章を読んで「ゲノム医療」に協力したいと思いましたか？

- (1) 協力したいと思った。
- (2) どちらかと言えば協力したいと思った。
- (3) どちらでもない。
- (4) どちらかと言えば協力したくないと思った。
- (5) 協力したくないと思った。

設問3 この文章の内容を十分理解できたと思いますか？

- (1) とてもよく理解できた。
- (2) 大部分はよく理解できた。
- (3) 概ね理解できた。
- (4) あまり理解できなかった。
- (5) 全く理解できなかった。

設問4 この文章の内容を他の人に説明できると思いますか？

- (1) しっかり説明できると思う。
- (2) 大部分は説明できると思う。
- (3) おおむね説明できると思う。
- (4) あまり説明できそうにない。
- (5) 全く説明できそうにない。

設問5 このような文章の理解を助けるのは以下のどのような要素でしょうか？重要な要素と思われる順に上位三つ選んでください。（複数回答可）

1位 2位 3位

- () () () 文章が短いこと。
- () () () 細かな点まで説明がなされていること。
- () () () Q&A形式であること。
- () () () 箇条書き形式であること。
- () () () 口語調であること。
- () () () その他 (自由記載)
- () () () その他 (自由記載)
- () () () その他 (自由記載)

質問は以上です。ご協力ありがとうございました。

(資料3) 平易化2

1. 計画の概要:

- ・日本中の人々のゲノム（DNA の情報）を解析して、大きなデータベースを作ります。
- ・このデータベースは、病気の治療や新薬の開発に役立てるために、研究者や企業と共有されます。

2. 目的:

- ・ゲノム解析の成果を患者さんに還元する。
- ・新しい個別化医療（人それぞれに合った治療）を実現する。
- ・ゲノムデータを使って、病気の研究や新薬の開発を進める。
- ・関連する研究者や医療関係者の育成、政策の検討を行う。

3. データ利用のルール:

- ・データの使用は法律やガイドラインに基づいて厳しく管理されます。
- ・使う前には、専門の委員会が公正にチェックします。

4. 将来の研究での利用:

- ・新しい医薬品や技術を開発するために、日本だけでなく海外の企業や研究機関もデータを使うことがあります。
- ・これにより、より多くの研究が進むことが期待されます。

5. 公的データベースへの登録:

- ・解析されたデータは、公的な研究用データベースにも登録され、他の研究者も使えるようになります。
- ・これにより、新しい発見が進む可能性があります。

6. 公益性の高い利用:

- ・このデータは、教育や保健医療政策のためにも使われます。
- ・例えば、医療の向上や行政の政策決定に役立つ情報として活用されることがあります。

7. 同意が必要:

- ・この計画に参加するためには、上記のすべての利用目的に同意する必要があります。

この計画は、将来の医療を良くするために、たくさんの人のゲノムデータを使って研究を進めるといふものです。そして、その過程で色々なルールに従ってデータを管理し、安全に利用することが大切にされています。

(資料4) 平易化3

Q1: 全ゲノム解析等実行計画って何ですか？

A1: この計画では、患者さんから提供されたゲノム情報（体の設計図のようなもの）と健康に関するデータを使って、データベースを作ります。このデータベースを使って、日本の研究者や会社が医療の改善や新薬の開発に役立てることができます。

Q2: この計画でデータはどう使われますか？

A2: データは主に4つの目的で使われます。1) 患者さんへの医療の改善、2) 新しい治療法の開発、3) 研究や新薬開発、4) 医療関係者の育成や政策の検討です。

Q3: データはどのように守られますか？

A3: データの安全と適切な利用を守るため、法律やルールに基づき、専門の委員会がデータの使い方を審査し、監督します。データを使う人は、これらのルールを守ることが必須です。

Q4: どんな研究にデータが使われるか、将来変わることはありますか？

A4: はい、研究は進歩するので、新しい研究にもデータが使われることがあります。その場合は、新しい計画を立てて、専門の委員会の承認を得てから使います。

Q5: 企業もこのデータを使えますか？

A5: はい、研究機関や医療機関だけでなく、医薬品や医療機器の研究開発を行う企業もデータを利用できます。ただし、適切な利用計画と承認が必要です。

Q6: このデータはどこに登録されますか？

A6: データは、国内外の公的な研究用データベースに登録されます。例えば、日本の科学技術振興機構が運営するヒトデータベースなどです。

Q7: このデータは他にどんな目的で使われますか？

A7: データは、医療従事者の育成や保健医療政策の検討にも使われます。これは、より良い医療サービスを提供するために重要です。

Q8: データを提供する際、何に同意する必要がありますか？

A8: 提供する際には、この計画で収集されたデータが上述の全ての目的で使われることに同意する必要があります。

(資料5) 平易化4

Q1: 「全ゲノム解析等実行計画」とは何ですか？

A1: 「全ゲノム解析等実行計画」とは、人々の遺伝子に関する大きなデータを集めて、それを使って病気の治療や新しい薬の開発をするための計画です。日本に住む人々の遺伝子情報を集めて、それを研究者や会社と共有して、もっといい医療を作るための基盤を作ります。

Q2: この計画でデータはどう使われますか？

A2: この計画では、集めた遺伝子データを使って、病気の治療法を改善したり、新しい薬を作ったりします。また、このデータは研究者や医師の訓練、健康政策の考え方を作るのにも使われます。

Q3: データの安全はどう守られますか？

A3: この計画では、データを安全に使うために、法律やルールに従って、特別な委員会がデータの使い方を審査して監督します。データを使う人は、セキュリティに関するガイドラインを守る必要があります。

Q4: このデータは将来どう使われるかもわかりますか？

A4: 将来の研究のためにもデータが使われる可能性があります。新しい研究計画があれば、また審査を受けて承認された後に使います。これは国内外の様々な研究機関や会社が参加するかもしれません。

Q5: 他の国の研究者もデータを使えますか？

A5: はい、公的なデータベースにデータを登録することで、日本だけでなく他の国の研究者も利用できます。ただし、どの国の研究者が使っても、その国の法律に基づいたガイドラインに従って使われます。

Q6: この計画に参加するとどのようなメリットがありますか？

A6: この計画に参加することで、あなたのデータが医学の進歩に貢献し、将来的にはより良い医療や新しい薬の開発に役立つかもしれません。また、データは教育や公共政策の検討にも使われ、社会全体の健康向上に貢献する可能性があります。

(資料6) 平易化5

あなたの全ゲノム解析等の結果は、医学の進歩と健康管理に役立つために研究機関・医療機関に提供されます。

本研究では、あなたに提供いただいた試料・情報を活用して「解析・データセンター」に全ゲノム等解析データと臨床情報のデータベースを構築します。このデータベースは国内外の研究者や企業の間で共有され、将来的には日本の患者さんの診療に役立つとされています。

研究機関や医療機関は、データを利用して診療や研究等を行い、データベースを利活用審査委員会などが公正に審査・監督します。研究機関・医療機関は、関連法令やAMEDとの契約に基づいてデータを利用し、その情報を公開することが求められます。本研究では、データの適切な利用を確保するため、関連法令・指針、データの利用条件、セキュリティに関するガイドラインを遵守しなければなりません。また、データの利用には、同意が必要です。

データ提供によって、あなたの健康状態を把握することができ、医学の進歩に貢献できます。ただし、データ利用により、将来的には新たな研究や治療の開発が期待されますが、その際に試料・情報を活用させていただく可能性があるため、研究計画が承認される前に改めて研究計画の審査を受ける必要があります。

研究機関や医療機関は、データの適切な利用を確保するため、関連法令やAMEDとの契約に基づいて利用し、その情報を公表します。

あなたには、本研究の参加に同意していただく必要があります。同意していただけない場合には、本研究の対象外となります。その場合でも、医学の進歩に貢献することはできません。

もし、あなたが提供した試料や情報が、新たな研究や研究結果の公開によって、あなたや他の患者さんに不利益が生じる可能性がある場合には、その提供を控えていただく必要があります。

以上、本研究の目的や利用方法などについてご理解いただいたうえで、同意いただけます場合には、同意書に署名をお願いいたします。同意書にご署名いただきました場合には、研究機関・医療機関に提出をお願いいたします。

(資料7) 平易化6

Q: わたしのゲノムデータは、どのように利用されるのですか？

A: あなたのゲノムデータは、次世代の医療の発展に貢献するために、病気の治療や予防、新薬の開発、健康管理などの研究に使われます。「全ゲノム解析等実行計画」は、データ提供者であるあなたの同意を頂いたうえで、研究に活用します。

Q: 遺伝子検査を受けると、どんなメリットがあるのですか？

A: 遺伝子検査を受けると、将来かかりやすい病気のリスクや体質を知ることができます。また、遺伝的なリスクに基づいた健康管理もできるので、病気の予防や健康維持に役立ちます。ただし、リスクが高い場合でも、生活習慣の改善などでリスクを軽減することができますので、将来の健康管理に役立てることができます。

Q: データを使用する際、わたしの個人情報を守られるのですか？

A: 「全ゲノム解析等実行計画」では、あなたの個人情報やゲノムデータの取り扱いについて、厳重に管理します。あなたのデータは、研究に活用される前に、データ利活用審査委員会によって厳正に審査され、データの安全な利用が確保されます。データの利用に際しては、関連法令・指針に従い、あなたのプライバシーや人権が守られるよう、適切に配慮されます。

Q: わたしが協力を拒否することはできますか？

A: あなたは、提供いただいたデータを研究に活用することについて、自由意思で拒否することができます。ただし、同意をいただいた場合は、上記のようにデータを利用していただきます。

Q: わたしの健康にどんなメリットがあるのですか？

A: あなたの遺伝子情報をもとに、将来かかるかもしれない病気を予防したり、病気にかかりにくい体質に改善することができます。遺伝子検査を受けることで、生活習慣の見直しなど、健康管理にも役立ちます。また、あなたにとってより健康的な方法を選択し、健康的な生活を送ることができるようになるかもしれません。

(資料8) 平易化7

患者：今日はどうしたのですか？

医師：今日は、あなたの全ゲノム解析等実行計画について説明したいと思います。

患者：全ゲノム解析ですか？

医師：そうです、あなたの遺伝子をすべて調べることで、将来の病気の予防や治療法の開発に役立つかもしれないのです。

全ゲノム解析等実行計画では、ご提供いただいた試料・情報を用いて「解析・データセンター」に全ゲノム等解析データおよび臨床情報のデータベースを構築し、国内外の研究者や企業の間でデータを広く共有して診療や研究・創薬等に活用するための仕組みを構築します。

この仕組みは、日本の患者さんや市民の皆さんのゲノム医療を日本に住む方々のデータを使って実現するための基盤であるとお考えください。

データの適正な利用を確保するため、関連法令・指針、AMED との契約および「全ゲノム解析等実行計画」に基づいてデータ共有ルールを定め、第三者も加わったデータ利活用審査委員会が公正な立場からデータの利用を審査・監督する仕組みを設けます。

データの利用者には、関連法令・指針、データの利用条件、およびセキュリティに関するガイドラインの遵守が義務づけられます。

患者：なるほど、では、私のデータはどのように使われるのですか？

医師：まず、将来の研究目的での利用です。具体的には、研究機関・医療機関が、医薬品や医療機器等の研究開発に携わる国内外の企業が、研究、医薬品等の開発、科学的なエビデンスに基づく予防等の目的で、データを利用できる環境を整備します。

次に、国内外の公的データベースへの登録と活用です。解析されたデータが、データベースを介して他の研究者に利用されることによって、新しい技術の開発が進むとともに、様々な病気について、今まで不可能であった疾患の原因の解明や治療法・予防法の確立が早まる可能性が期待できるのです。

患者：では、私のデータはどこで使われるのですか？

医師：公的な研究用データベース（国外のものを含む）に、解析されたデータが提供・登録されます。公的データベースに登録されたデータは、日本国内の研究機関に所属する研究者だけではなく、国内外の製薬企業等の民間企業や海外の研究機関に所属する研究者も利用する可能性があります。

患者：なるほど、私のデータは、将来的には海外の研究者にも使われるかもしれないですね。

医師：はい、そのため、将来的に医療がよりよくなる可能性が期待できます。

患者：わかりました。自分のデータが使われることで、将来の医療がよりよいものになるな

ら、協力する意義があると思います。

医師：ありがとうございます。あなたの協力が、将来の医療の発展につながります。

厚生労働科学研究費補助金（厚生労働科学特別研究事業）
分担研究報告書

テーマC：電子カルテからの自動情報抽出精度の検証

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
研究分担者	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
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
研究協力者	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

研究要旨

目的：本研究では、自然言語およびテンプレート形式、検査記録などを含む電子カルテのデータを用いて、生成 AI により模擬研究 electric CRF (eCRF) の入力項目に自動的にデータを抽出することができるかどうかとその正確性の検証、また自動抽出された eCRF の入力内容から模擬研究の参加適格者/不適格者の仕分けを適切に実行できるかどうかを検証することを目的とする。

方法：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 WritingX 等の生成 AI ソフトウェアを用いて、模擬研究 eCRF の自動入力の実験と参加適格者/不適格者判定の実験を実施した。

結果：模擬研究 eCRF 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：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 項目の抽出する可能性を示すことができた。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。

A. 研究目的

臨床研究を行う際に、研究に関する情報を研究用のデータとして収集する際には医師、臨床研究コーディネーターが電子カルテを閲覧しながら electric CRF (eCRF) に必要な情報を入力する作業を担っている。研究を担当する医療従事者はその情報入力作業の多くを手作業で行う必要がある。電子的に情報入力を行う electric data capture (EDC) が普及しているとはいえ、電子カルテに保管された情報をそのまま電子的に抽出、引用することは困難であり、入力作業は未だ手作業

で行われている。本研究では自然言語およびテンプレート形式、検査記録などを含む電子カルテのデータを用いて、生成 AI により模擬研究 eCRF の入力項目に自動的にデータを抽出することができるかどうかとその正確性の検証、また自動抽出された eCRF の入力内容から模擬研究の参加適格者/不適格者の仕分けを適切に実行できるかどうかを検証することを目的とする。

B. 研究方法

情報源としては、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用いる。事前に個人情報や個人が特定できる情報、診療に関する機微な情報などが含まれていない形に限定した。

1. 模擬研究 eCRF の自動入力検証

生成 AI が電子カルテの診療録から指定した項目に従って正確に情報抽出を行うことが可能かどうかを検証するために、大阪大学医学部附属病院で採用されているダイナミックテンプレートを用いて記録された電子カルテ情報を利用した。本テンプレートから入力すると、その内容が構造化データとして格納される。また、診療記録上で表示される際には違和感のない自然な文（自然言語文）に変換して出力されるため、これらを組み合わせると自然言語文と構造化データのペアが準備できる。本検証では、生成 AI に自然言語文と構造化データの項目名を与え、構造化データの値が正しく抽出できるかを評価することで、eCRF の自動入力を擬似的に検証することとした。

対象患者は心不全の増悪で入院し、患者背景や既往歴などを含む下記 42 項目に欠損のない 50 名を対象とした。各項目の詳細とデータ型は以下の通りであった：

- 日付：3 項目
入院日、身長測定日、体重測定日
- 数値：10 項目
入院時年齢、身長、体重、Body Mass Index、来院時収縮期血圧、来院時拡張期血圧、心拍数、体温、経皮的動脈血酸素飽和度、心胸郭比
- 文字列：5 項目
性別、入院目的、入院形式、NYHA、胸部エ

ックス線体位

- カテゴリ（あり/なし/不明）：24 項目
急性心不全による入院、過去の心不全入院治療歴、植込みデバイス、在宅酸素療法、持続陽圧呼吸療法、自動調節サーボ換気、補助人工心臓、維持透析、経皮的冠動脈インターベンション/経皮的冠動脈形成術、カテーテルアブレーション、経皮的心臓手術、開心術、高血圧、糖尿病、脂質異常症、慢性閉塞性肺疾患、心房細動/心房粗動、慢性腎臓病、肝疾患、膠原病、喫煙歴、多量飲酒歴、入院時 6 分間歩行テスト、入院時 Fibroscan

本テンプレートではこれら以外にも現病歴やその他の既往歴、治療方針などが入力されており、それらも含めて自然言語文に出力される。本検証ではこれらの情報を誤って抽出対象としないかどうかとも評価に含まれる。また、上記 42 項目のうち、性別は自然言語文の出力には含まれない。本項目の出力を他の 41 項目とは区別して評価し、生成 AI が記述されていない情報を出力してしまうハルシネーションについて検討した。

本検証では生成 AI として東京リージョンの Amazon Bedrock で利用できる Claude 2.1 を採用した。プロンプトでは自然言語文と各項目名を与え、JSON 形式で解釈可能な文字列を生成する指示を構築した。なお、各項目の医学的説明は指示プロンプト内では与えていない。

2. 研究参加適格者/不適格者の仕分け検証

DEEP INTELLIGENT PHARMA(SG) PRIVATE LIMITED が提供している WritingX の生成 AI アプリケーションを用いて、治験実施計画書から模擬研究の参加適格基準/除外基準の自

動抽出を試みる。また、電子カルテデータ(診療録、検査報告書、検体検査結果、入院時バイタルサインなど)を書き下した自然言語文を生成 AI アプリケーションに与え、自動抽出された各基準項目について適格/不適格/判定不可の判定をその理由とともに出力し、適切に判定できているかを検証する。

・電子カルテデータ提供のプロセス

対象患者は大阪大学医学部附属病院で過去に実施された治験(腹膜播種を評価するための審査腹腔鏡検査に関する治験)から選択し、15名を選出した。模擬研究の条件は同治験の条件を参考にして以下の項目を設定した:

【適格基準】

- 1) 同意取得時の年齢が、20歳以上の患者。
- 2) 進行胃癌のうち大型3型(最大径8cm以上)、4型胃癌もしくはCT画像上で腹膜播種が疑われる胃癌患者。
- 3) ECOG Performance Statusが0~1の患者。
- 4) 心、肺、肝、腎、骨髄機能が保たれており、臨床検査値が以下の条件をすべて満たす患者。(a~c、eについては、審査腹腔鏡施行前14日以内の検査値を用いる。dについては、同意取得前14日以内の検査結果の利用も可能とする。) a. 腎機能: BUN \leq 施設正常値上限の1.5倍 血清クレアチニン値 \leq 施設正常値上限の1.5倍 b. 肝機能: AST(GOT) \leq 施設正常値上限の1.5倍 ALT(GPT) \leq 施設正常値上限の1.5倍 血清総ビリルビン値 \leq 施設正常値上限の1.5倍 c. 血液: 好中球 1,500/mm³ 以

上血小板 100,000/mm³ 以上 Hb 8.0g/dL

以上 d. 心機能心電図: 重篤な異常を認めない e. 肺機能: 肺機能: 動脈血酸素飽和度(SpO₂) \geq 94% (酸素非投与下)

- 5) 他に活動性の重複癌(血液悪性腫瘍を含む)がない患者。
- 6) 本人に病名・病態の告知がなされており、かつ本人に同意能力があり、文書同意が得られた患者。

【除外基準】

- 1) コントロール不良な感染症(活動性の結核を含む)を有する患者。
- 2) 重篤な合併症(悪性高血圧、重症のうつ血性心不全、重症の冠不全、同意取得前3ヶ月以内の心筋梗塞、末期肝硬変、重症の腎不全あるいは血液透析を要する慢性腎不全、コントロール不良な糖尿病、重症の肺線維症、活動性の間質性肺炎等)を有する患者。
- 3) 同種骨髄移植後の患者。
- 4) 重症の精神障害を有する患者。
- 5) ポルフィリン類縁物質に対して過敏症又は既往のある患者。
- 6) 他の治験(臨床研究を含む)に参加している患者。
- 7) 妊娠、授乳婦である患者。
- 8) その他、治験責任医師又は治験分担医師が不適格と判断した患者。

これらの情報が記載されている可能性のある電子カルテデータとして、画像診断報告書、患者プロフィール、患者基本/患者属性、看護サマリ、電子カルテの記載、上部内視鏡検査報告書、熱型表バイタル記録リスト、臨床検査値一覧を抽出対象とした。一定のフォ

ーマットで自然言語文に変換できるようにするために、各情報源をシート名にして、構造をできる限り保持したxlsx ファイルを患者単位で作成した。xlsx ファイルを研究協力者へ転送する前に、同ファイルをアジャイル暗号化して提供した。

- ・生成 AI アプリケーションの詳細説明

プロトコル解析:プロトコルの文書情報を構造化解析し、選択基準/除外基準、試験日、名称、基本情報などを抽出した。

データ前処理:被験者データに対してプライバシー保護処理を実施し、個人を特定できる情報をチェック、マスキングすることでデータの安全性を確保した。また、Day 0 以前の情報のみを参照するようにデータをフィルタリングした。

自動判定処理:大規模言語モデル (LLM: GPT-4 turbo) とカスタムプロンプトを用いて、被験者データとプロトコル情報を分析し、被験者が選択基準/除外基準を満たすか否かを自動的に判定した。また、ユーザーが判定結果を最適化するために、プロンプトや LLM 関連パラメータ (モデル選択、温度設定、最大応答長など) をカスタマイズできるように UI を設計した。そして、被験者が選択基準/除外基準を満たすか否かを出力し、その判定理由も併せて提供した。また、結果の出力を可能にすることで、その後の分析やアプリケーションに活用できるように機能を追加した。

- ・倫理面への配慮

本研究は、電子カルテ情報を用いて実施する観点から、本研究実施計画書を作成し、大阪大学医学部附属病院倫理審査委員会によ

る審査を受け実施した (承認番号: 23477)。

C. 研究結果

1. 模擬研究 eCRF の自動入力検証

50 名の自然言語文の平均文字数は 2,941 ±802 であった。性別を除いた 41 項目の精度は 99.7% (2,044/2,050) を達成した。一方で、性別に関しては「記載なし」もしくは空白が出力されたのは 12% (6/50) にとどまった。また、7% (7/50) で逆の性別が出力された。

2. 研究参加適格者/不適格者の仕分け検証

治験実施計画書は A4 全 38 ページで構成され、そのうちの 2 ページにわたって適格基準と除外基準が記載されていた。生成 AI アプリケーションは治験実施計画書から適格基準 6 項目、除外基準 8 項目の全てを正しく自動抽出できた。

適格基準の判定精度に関して、15 名のうち 14 名 (93.3%) で 6 項目全てを正しく判定した。1 名のみ、「他に活動性の重複癌 (血液悪性腫瘍を含む) がない患者」の 1 項目で基準を満たさないと判定された (全体精度 98.9%)。

除外基準に関しては、対象患者の判定項目合計数 120 のうち、除外基準を満たすと判定された項目が 1 (0.8%)、除外基準を満たさないと判定された項目が 107 (89.2%)、判定不可の項目が 12 (10.0%) であった。

D. 考察

1. 模擬研究 eCRF の自動入力検証

自然言語文において項目名や値が明確に記載されている場合、自動構造化は極めて高い精度 (99.7%) で実行可能であることを確

認できた。また、今回の検証で欠落した6項目は身長もしくは体重の測定日であり、いずれも生成AIの出力は身長・体重の測定値の後方に測定日が記述された形式となっており、完全な見落としではなかった。一般的に生成AIでは、学習に用いたテキストのドメインやタイプと異なるタスクに対してはパフォーマンスが低下する可能性が懸念されている。しかしながら、本検証結果は、専門的な医学用語が含まれる日本語のプロンプトでも生成AIが有効に機能して情報抽出を達成したことを示している。これは、臨床データの自動抽出において重要な進歩を示唆する。

一方で、性別項目におけるハルシネーション現象が観察され、生成AIが文脈から性別を推測する際に不正確な情報を生成しうることが明らかになった。この問題に対処するためには、生成AIの推論根拠を明確にするとともに、ユーザーが生成AIの判断を容易に評価できるような仕組みの導入が重要である。

本検証の限界は、入力が自然言語文とはいえ、特定のフォーマットに沿って記述された文章であり、再構造化が比較的容易である点にある。より自由度の高い文章に対して、各判定項目に関して医学的解釈を要する場合の検討は今後の課題である。

2. 研究参加適格者/不適格者の仕分け検証

生成AIアプリケーションは治験実施計画書から適格基準および除外基準を正確に抽出し、またそれらの項目毎に判断根拠を提示しながら高精度に判定を出力することができた。しかし、適格基準で誤って不適格判定された患者が一例存在した。この患者の判断

根拠として「リンパ腫の合併」が挙げられていたが、実際にはそのような併存症はなかった。腹部造影CTレポートに記載された「リンパ節腫大」や「リンパ節転移」が誤判定の原因だと推察される。

一方、除外基準に該当すると判定された一例では、「コントロール不良な感染症（活動性の結核を含む）を有する患者」が原因であった。具体的には術前の臨床検査でHBs抗原陽性、HBs抗体陰性、HBc抗体陽性と判定された活動性B型肝炎の患者であった。「コントロール不良な」という表現の曖昧さが誤判定を引き起こした可能性がある。

なお、除外基準に関して、今回の対象者は実際の治験に全員参加できているため、臨床的には除外基準を満たしていないという判定は正しいと言える。しかし、これらの項目を判定するための具体的な情報は電子カルテの構造化データには存在せず、しばしば電子カルテの記録にも記載されないこともある。「妊娠、授乳婦である患者」に関しては、性別から明確に除外基準を満たさないと判定できる場合もあるが、本検証ではそれ以外の除外基準については「判定不可」と出力されることが望ましい。除外基準を満たさないと判断された項目の判断根拠を確認すると、「no mention」や「indicated」のような表現が使用されていた。生成AIアプリケーションは直接的な記載や情報がない場合の解釈も行っているようにみえるが、それを「特定不可」と明確にするのは困難な問題だと考えられる。

本検証の限界として、模擬研究で参照した治験に参加した患者以外での評価が行えていないことが挙げられる。適格基準を明らかに満たさない患者や、除外基準の疾患を合併

している患者に対して生成 AI アプリケーションが下した各項目の判定と判断根拠を精査することで、記載のない情報に対して高精度に「特定不可」と出力させることのできる指示プロンプトが構築できる可能性が残されている。

E. 結論

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
なし
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

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

別紙 4

研究成果の刊行に関する一覧表

書籍

著者氏名	論文タイトル名	書籍全体の編集者名	書 籍 名	出版社名	出版地	出版年	ページ
該当なし							

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
該当なし					

2024年4月18日

厚生労働大臣
—(国立医薬品食品衛生研究所長) 殿
—(国立保健医療科学院長)—

機関名 国立大学法人大阪大学

所属研究機関長 職名 医学部附属病院 病院長

氏名 野々村 祝夫

次の職員の令和5年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

- 研究事業名 厚生労働科学特別研究事業
- 研究課題名 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
- 研究者名 (所属部署・職名) 医学部附属病院 未来医療開発部・特任准教授 (常勤)
(氏名・フリガナ) 浅野 健人 (アサノ ケント)

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
人を対象とする生命科学・医学系研究に関する倫理指針 (※3)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	大阪大学医学部附属病院	<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること (指針の名称:)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他 (特記事項)

(※2) 未審査に場合は、その理由を記載すること。

(※3) 廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> (有の場合はその内容:)

(留意事項) ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

2024 年 4 月 18 日

厚生労働大臣
—(国立医薬品食品衛生研究所長)— 殿
—(国立保健医療科学院長)—

機関名 国立大学法人大阪大学

所属研究機関長 職 名 大学院医学系研究科長

氏 名 熊ノ郷 淳

次の職員の令和5年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

1. 研究事業名 厚生労働科学特別研究事業

2. 研究課題名 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

3. 研究者名 (所属部署・職名) 大学院医学系研究科 ・ 教授

(氏名・フリガナ) 土岐 祐一郎 (ドキ ユウイチロウ)

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
人を対象とする生命科学・医学系研究に関する倫理指針 (※3)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	大阪大学医学部附属病院	<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること (指針の名称:)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他 (特記事項)

(※2) 未審査の場合は、その理由を記載すること。

(※3) 廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> (有の場合はその内容:)

(留意事項) ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

厚生労働大臣
—(国立医薬品食品衛生研究所長) 殿
—(国立保健医療科学院長)—

機関名 国立大学法人大阪大学

所属研究機関長 職名 医学部附属病院長

氏名 野々村 祝夫

次の職員の令和5年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

- 研究事業名 厚生労働科学特別研究事業
- 研究課題名 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
- 研究者名 (所属部署・職名) 医学部附属病院医療情報部 / 大学院医学系研究科医療情報学・部長/教授
(氏名・フリガナ) 武田 理宏 (タケダ トシヒロ)

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
人を対象とする生命科学・医学系研究に関する倫理指針 (※3)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	大阪大学	<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること (指針の名称:)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他 (特記事項)

(※2) 未審査の場合は、その理由を記載すること。

(※3) 廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> (有の場合はその内容:)

(留意事項) ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。

2024 年 4 月 18 日

厚生労働大臣
—(国立医薬品食品衛生研究所長)— 殿
—(国立保健医療科学院長)—

機関名 国立大学法人大阪大学

所属研究機関長 職 名 大学院医学系研究科長

氏 名 熊ノ郷 淳

次の職員の令和5年度厚生労働科学研究費の調査研究における、倫理審査状況及び利益相反等の管理については以下のとおりです。

1. 研究事業名 厚生労働科学特別研究事業

2. 研究課題名 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

3. 研究者名 (所属部署・職名) 大学院医学系研究科 ・ 教授

(氏名・フリガナ) 川崎 良 (カワサキ リョウ)

4. 倫理審査の状況

	該当性の有無		左記で該当がある場合のみ記入 (※1)		
	有	無	審査済み	審査した機関	未審査 (※2)
人を対象とする生命科学・医学系研究に関する倫理指針 (※3)	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input checked="" type="checkbox"/>	大阪大学医学部附属病院	<input type="checkbox"/>
遺伝子治療等臨床研究に関する指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
厚生労働省の所管する実施機関における動物実験等の実施に関する基本指針	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>
その他、該当する倫理指針があれば記入すること (指針の名称:)	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>		<input type="checkbox"/>

(※1) 当該研究者が当該研究を実施するに当たり遵守すべき倫理指針に関する倫理委員会の審査が済んでいる場合は、「審査済み」にチェックし一部若しくは全部の審査が完了していない場合は、「未審査」にチェックすること。

その他 (特記事項)

(※2) 未審査の場合は、その理由を記載すること。

(※3) 廃止前の「疫学研究に関する倫理指針」、「臨床研究に関する倫理指針」、「ヒトゲノム・遺伝子解析研究に関する倫理指針」、「人を対象とする医学系研究に関する倫理指針」に準拠する場合は、当該項目に記入すること。

5. 厚生労働分野の研究活動における不正行為への対応について

研究倫理教育の受講状況	受講 <input checked="" type="checkbox"/> 未受講 <input type="checkbox"/>
-------------	---

6. 利益相反の管理

当研究機関におけるCOIの管理に関する規定の策定	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究機関におけるCOI委員会設置の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合は委託先機関:)
当研究に係るCOIについての報告・審査の有無	有 <input checked="" type="checkbox"/> 無 <input type="checkbox"/> (無の場合はその理由:)
当研究に係るCOIについての指導・管理の有無	有 <input type="checkbox"/> 無 <input checked="" type="checkbox"/> (有の場合はその内容:)

(留意事項) ・該当する□にチェックを入れること。
・分担研究者の所属する機関の長も作成すること。