

2026年6月30日

治験候補患者抽出の精度向上・効率化に向け、 リアルワールドデータと生成 AI (LLM) を用いた共同研究を開始

学校法人近畿大学
中外製薬株式会社
NTT 株式会社
株式会社 NTT データ

近畿大学病院 腫瘍内科(以下、近畿大学病院)、中外製薬株式会社(以下、中外製薬)、NTT 株式会社(以下、NTT)、株式会社 NTT データ(以下、NTT データ)は、実臨床で蓄積されるリアルワールドデータ^(注1)と AI 技術である大規模言語モデル(LLM)^(注2)を活用し、治験候補患者抽出の精度および抽出プロセスの効率を検証する4者共同研究(以下、本研究)を2026年6月より開始しました。本研究では近畿大学病院が保有する電子カルテデータ等を対象に、中外製薬が策定した治験実施計画書^(注3)の適格基準^(注4)に基づき、従来の抽出手法と LLM を組み合わせた抽出手法を比較・評価します。医師および治験コーディネーター(CRC)^(注5)の判定結果を比較基準とし、実運用における有効性や作業負荷低減、治験参加者組み入れまでのリードタイム短縮への寄与を多面的に検証します。

【背景】

新薬の臨床開発において、治験開始までの期間および治験参加者の組み入れ期間は上市^(注6)時期に大きく影響します。なかでも治験候補患者の抽出は、治験実施計画書に定められた適格基準をもとに医師や CRC が個別に診療情報を確認する必要があり、多くの時間と労力を要してきました。その結果、治験参加者の組み入れが計画通りに進まず、治験全体のスケジュールに影響を及ぼすケースが指摘されています。

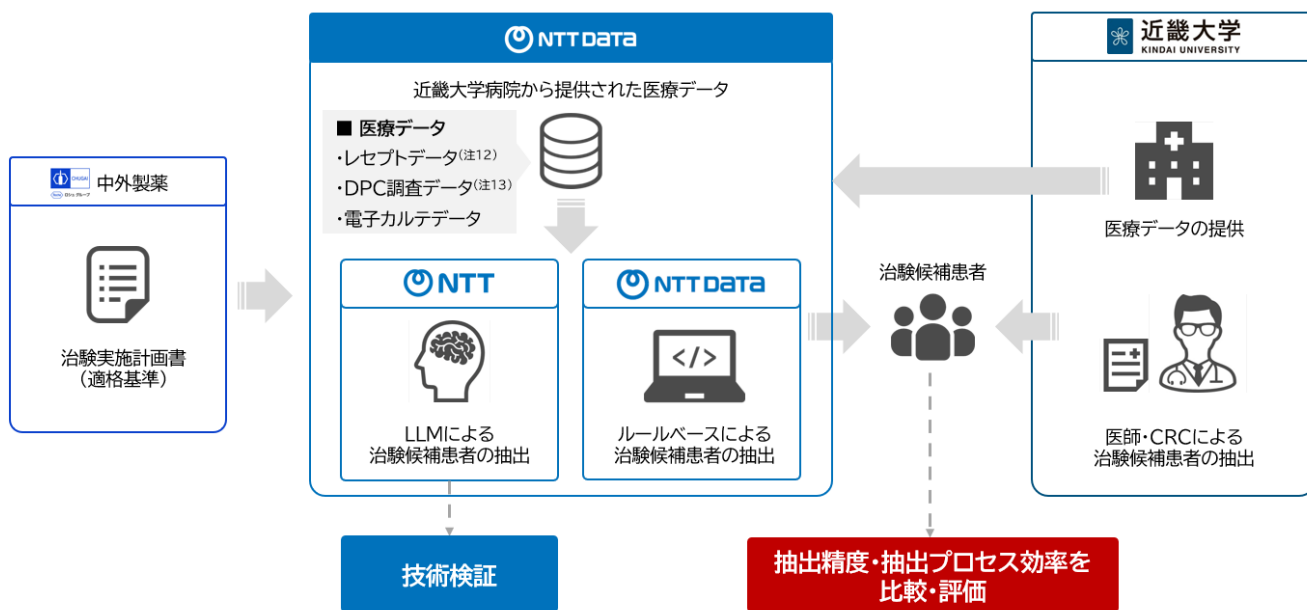
近年、実臨床で蓄積されるリアルワールドデータの利活用が進む中で、非構造化データを含む患者情報を横断的に解釈できる LLM の活用が注目されています。これらの技術を活用することで、治験候補患者抽出の精度の向上や、抽出プロセスの効率化が期待されています。

【本研究の概要】

本研究では、近畿大学病院の電子カルテデータ等を活用し、中外製薬が策定した治験実施計画書に定められた適格基準に基づき LLM を活用した治験候補患者の抽出を行います。抽出にあたっては、NTT データが、これまで医療情報活用基盤「千年カルテ[®]」^(注7)の運用で培ってきた安全な情報管理およびデータ運用設計の実績に加え、医療データの活用・解析に関する知見を生かし、LLM およびルールベース手法^(注8)による治験候補患者抽出の技術検証を実施します。本技術検証においては、NTT が独自開発した純国産の大規模言語モデル「tsuzumi 2」^(注9)を用います。データ統制を重視した設計のもと、機微情報を含むデータへの対応を想定した運用が可能な点を特長としています。さらに医学的公開論文等の事前学習を継続的に行った医療特化型「tsuzumi 2」として検証を行う予定です。

抽出手法として、①Python^(注10)および SQL^(注11)によるルールベースの抽出手法、②LLM を活用した抽出手法、また①と②を組み合わせた手法、それぞれを実施し、それらの結果を医師および CRC による判定結果と比較することで、治験候補患者抽出の精度を評価します。あわせて、治験候補患者の抽出に要する時間や、医師・CRC の作業量・内容の変化についても確認し、治験候補患者抽出の精度および抽出プロセスの効率化の両面から、治験参加者組み入れまでのリードタイム短縮につながるかを検証します。

なお本研究は、近畿大学病院の近畿大学医学部等倫理委員会の承認を得たうえで、2027年3月までの実施を予定しています。



本研究の概要

【本研究の意義】

本研究は、LLM 活用による治験候補患者抽出の精度向上と抽出プロセスの効率化が、治験参加者組み入れまでのリードタイム短縮につながるかを定量的に検証するものです。これにより、治験全体の期間短縮、製薬企業の開発スピード向上、医療機関における治験実施体制のスリム化を図るとともに、患者が新たな治療選択肢へより迅速にアクセスできる環境の整備に貢献します。なお、AI の出力結果は医師の判断支援を目的としたものであり、最終的な診療判断は医師が行うことを前提とします。

【各機関の役割】

- ・近畿大学病院:医療データの提供、治験候補患者の抽出および抽出精度・抽出プロセス効率の比較・評価
- ・中外製薬:治験実施計画書(適格基準)の提供および評価協力
- ・NTT:LLM による治験候補患者抽出に関する技術検証^(注14)の実施
- ・NTT データ:ルールベースによる治験候補患者抽出の実施および、抽出精度・抽出プロセス効率の比較・評価

なお、本研究は、2026年5月に近畿大学病院、NTT データが発表した技術検証^(注15)の成果で得られた知見を基盤としています。具体的には、治験参加者組み入れが進行中の治験実施計画書に定められた適格基準の下でリアルワールドデータを用い、治験候補患者抽出の有効性および実運用への適用可能性を評価するものです。

本研究の成果を踏まえ、関係各社は医療機関および製薬企業とのさらなる連携を通じたリアルワールドデータと AI を活用した治験候補患者抽出基盤の社会実装の可能性について検討を進めていきます。

【注釈】

注1 リアルワールドデータ:実際の診療現場において蓄積される医療データで、臨床試験や新薬開発、医療の質向上などへの活用が進んでいる。

注2 大規模言語モデル(LLM):大量のテキストデータを学習し、文章の理解や生成を可能とする人工知能技術。自然言語処理の分野で幅広く活用されている。

注3 治験実施計画書:新しい医薬品の効果や安全性を調べるために、治験の目的や方法、参加条件などをあらかじめ定めた計画書。

注4 適格基準:治験の安全性や科学的な信頼性を確保するために、参加できる方の条件を年齢や症状などにより定めたもの。

注5 治験コーディネーター(CRC):医師や治験参加者、製薬企業などの関係者間を調整し、治験が適切に実施されるよう支援する医療専門職。

注6 上市:新薬が薬事承認を受けた後に販売され、医療機関で使用できるようになること。

- 注7 千年カルテ:「医療分野の研究開発に資するための匿名加工医療情報及び仮名加工医療情報に関する法律」(次世代医療基盤法)に基づき、一般社団法人ライフデータイニシアティブと株式会社 NTT データが国から認定を受けて、実施している事業の名称。
- 注8 ルールベース手法:治験実施計画書に定められた条件をあらかじめプログラム化し、治験候補患者を抽出する手法。
- 注9 tsuzumi 2:NTT 株式会社が開発した大規模言語モデル。日本語を含む多様な言語処理に対応している。
- 注10 Python:データ解析やシステム開発などに広く利用されているプログラミング言語。
- 注11 SQL(Structured Query Language):データベースを操作するための言語。条件に基づくデータ抽出や集計に使用される。
- 注12 レセプトデータ:医療機関が診療報酬を請求するために作成するデータで、診療内容や処方された薬剤などの情報が含まれる。
- 注13 DPC 調査データ:入院医療において診断名や治療内容を整理したデータで、患者の病状や治療の概要を把握するための情報が含まれる。
- 注14 LLM による治験候補患者抽出に関する技術検証:NTT コンピュータ&データサイエンス研究所が開発した技術を用いて検証を行う。
- 注15 2026 年 5 月に近畿大学病院、NTT データが発表した技術検証:リアルワールドデータを活用した治験候補患者抽出の技術検証を実施し、その有効性を検証した取り組み。(参照:<https://www.nttdata.com/global/ja/news/topics/2026/052500/>)

*「千年カルテ」は日本国内における一般社団法人ライフデータイニシアティブの登録商標です。

*「tsuzumi」は NTT 株式会社の登録商標です。

*その他の商品名、会社名、団体名は、各社の商標または登録商標です。

【本件に関するお問い合わせ先】

学校法人近畿大学
医学部・看護学部・病院運営本部 総務広報課
担当:森井、橋本
TEL:072-288-7222 FAX:072-288-6303
E-mail:med.koho@itp.kindai.ac.jp

中外製薬株式会社
・報道関係者の皆様:
広報 IR 部メディアリレーションズグループ
TEL:03-3273-0881
E-mail:pr@chugai-pharm.co.jp
・投資家の皆様
広報 IR 部インベスターリレーションズグループ
TEL:03-3273-0554
E-mail:ir@chugai-pharm.co.jp

NTT 株式会社
広報部門
E-mail:ntt-pr@ntt.com

株式会社 NTT データ
インダストリ統括本部 第二インダストリ事業本部 製薬・化学事業部
担当:込山、池ヶ谷
E-mail:patient-recruitment@am.nttdata.co.jp