

【公開日】 2026 年 10 月 8 日

作成日 2026 年 8 月 25 日
(最終更新日 2026 年 9 月 29 日)

「情報公開文書」

受付番号：2026-4-102

課題名：薬剤知識ベースを用いた LLM による服薬情報正規化手法の開発と評価に関する研究

研究責任者：東北メディカル・メガバンク機構 教授 木下 賢吾

1. 研究の対象

東北メディカル・メガバンク計画の地域住民コホート調査（宮城、岩手）・三世代コホート調査に参加された方

2. 研究目的・方法

【研究期間及び試料・情報の利用又は提供を開始する日】

研究期間：2026 年 10 月（研究実施許可日）～2031 年 3 月

試料・情報の利用又は提供を開始する日：研究実施許可日

【研究目的】

東北メディカル・メガバンク計画のコホート調査では、参加された方がお使いのお薬について、調査票等に文字で記入していただいています。お薬に関する情報は、病気の治療の状況や健康状態を知るための重要な情報であり、健康と病気の関係を調べる研究に欠かせません。しかし、実際に記入された内容には、お薬の名前（商品名や一般名など）、成分名、剤形（錠剤、粉薬など）、量について混ざって書かれていたり、書き間違いや略称、複数のお薬がまとめて書かれていたりすることがあります。そのままでは、同じお薬をまとめ、お薬の種類ごとに分類して解析することが難しいため、これまでは専門的な知識を持つ担当者が内容を確認して整理してきました。

この研究では、こうしたお薬に関する記入内容を、正しいと考えられるお薬の名前・一般名・成分・剤形・量・お薬の分類などに自動的に対応付けて整理する（正規化する）方法を、文章を扱う人工知能の一つである大規模言語モデル（Large Language Model：LLM）を用いて開発することを目的とします。あわせて、LLM がお薬に関する専門的な内容をどの程度正確に扱えるのか、どのような使い方をすれば正確かつ効率よく整理できるのかについても評価します。この研究で行うのは、お薬に関する記述を正規化する方法をつくり、その正確さを確かめるところまでです。正規化の対象とするのは調査票等に記入されたお薬に関する記述のみであり、この記述をお一人おひとりの他の情報（病気や検査の結果など）と結び付けて調べることはしません。この研究により、コホート調査で集められたお薬の情報を、研究に利用しやすい形に整えることを目指します。

【研究方法】

この研究では、東北メディカル・メガバンク計画のコホート調査で既に収集された情報のみを用います。新たに採血や検査、問診を行うことはありません。お薬に関する記述は自由に文章で書いていただいたものですので、お名前や住所など特定の個人が分かる情報が含まれていないことをあらかじめ確認したうえで用います。研究は、大きく次の二つの段階に分けて行います。

（１）お薬に関する知識をまとめたデータベース（薬剤知識ベース）を作成します。公開されている薬剤データベースや分類情報と、これまで当機構で整備してきたお薬の対応表を用いて、お薬ごとに一般名・商品名・剤形・規格含量・お薬の分類等を結び付けた辞書を作成します。一定のルールに基づいて自動的に対応できない部分については、LLM を用いた対応付けを行い、その正確さを確認します。

（２）作成した薬剤知識ベースと LLM を用いて、調査票等に記入されたお薬に関する記述の正規化を行います。具体的には、該当する可能性のあるお薬の候補の提示、複数のお薬がまとめて書かれている場合の切り分け、剤形や量の抽出、判断が難しい場合のその理由の出力を行います。LLM には薬剤知識ベースの内容を参照させる方法を用いることで、LLM がもっている知識のみに頼ったもっともらしい誤り（ハルシネーション）が生じにくい仕組みとします。

正規化した結果は、既にあるお薬の対応表との一致の程度や、専門的な知識を持つ研究者による確認により評価します。また、複数の LLM（企業が提供するもの、公開されているもの、日本語やお薬の分野に特化したもの）を比較し、正確さ、処理に必要な時間や費用、同じ処理を繰り返したときの結果の安定性等を評価します。

この研究は、お薬に関する記述を整理する方法を開発し、評価するものであり、個々の参加者の方の体質や病気との関連を調べる解析は行いません。そのため、成果の論文発表や知的財産化を除いては、個別にお知らせするような結果はありません。

3. 研究に用いる試料・情報の種類

対象者の基本情報（年齢・性別）、調査票情報（服薬情報）

4. 外部への試料・情報の提供

該当なし

5. 関係研究組織

該当なし

6. お問い合わせ先

本研究に関するご質問等がありましたら下記の連絡先までお問い合わせ下さい。

ご希望があれば、他の研究対象者の個人情報及び知的財産の保護に支障がない範囲内で、研究計画書及び関連資料を閲覧することが出来ますのでお申出下さい。

東北大学 東北メディカル・メガバンク機構 ゲノム解析部門 木下賢吾
〒980-8573 宮城県仙台市青葉区星陵町 2-1 TEL : 022-274-6040

東北大学の東北メディカル・メガバンク事業に協力された方で、本研究に限って試料・情報の利用を希望されない方は、下記までご連絡下さい。

東北大学 東北メディカル・メガバンク機構 地域住民コホート担当
〒980-8573 宮城県仙台市青葉区星陵町 2-1 TEL : 022-718-5161

東北大学 東北メディカル・メガバンク機構 三世代コホート担当
〒980-8573 宮城県仙台市青葉区星陵町 2-1 TEL : 022-718-5162

◆個人情報の利用目的の通知

保有個人情報の利用目的の通知に関するお問い合わせ先：「6. お問い合わせ先」

※注意事項

以下に該当する場合にはお応えできないことがあります。

＜人を対象とする生命科学・医学系研究に関する倫理指針 第9章第18の1＞

＜個人情報の保護に関する法律第21条の4＞

- ①利用目的を本人に通知し、又は公表することにより本人又は第三者の生命、身体、財産その他の権利利益を害するおそれがある場合
- ②利用目的を本人に通知し、又は公表することにより当該個人情報取扱事業者の権利又は正当な利益を害するおそれがある場合
- ③国の機関又は地方公共団体が法令の定める事務を遂行することに対して協力する必要がある場合であって、利用目的を本人に通知し、又は公表することにより当該事務の遂行に支障を及ぼすおそれがあるとき。
- ④取得の状況からみて利用目的が明らかであると認められる場合

◆個人情報の開示等に関する手続

東北大学が保有する個人情報のうち、本人の情報について、開示、訂正及び利用停止を請求することができます。

保有個人情報とは、東北大学の役員又は職員が職務上作成し、又は取得した個人情報です。

保有する個人情報については、所定の請求用紙に必要事項を記入し情報公開室受付窓口へ提出するか又は郵送願います。詳しくは当機構HPよりプライバシーポリシーを確認の上、請求手続きのホームページをご覧ください。（※手数料が必要です。）

【東北大学東北メディカル・メガバンク機構プライバシーポリシー】

<https://www.megabank.tohoku.ac.jp/contact/privacypolicy>

【東北大学情報公開室】

<https://www.bureau.tohoku.ac.jp/kokai/disclosure/index.html>

※注意事項

以下に該当する場合には全部若しくは一部についてお応えできないことがあります。

＜人を対象とする生命・医学系研究に関する倫理指針 第9章第18の1＞

＜個人情報の保護に関する法律第33条の2＞

- ①本人又は第三者の生命、身体、財産その他の権利利益を害するおそれがある場合
- ②個人情報取扱事業者の業務の適正な実施に著しい支障を及ぼすおそれがある場合
- ③他の法令に違反することとなる場合