

令和6年度 成果報告書

基本情報(公開)

事業名 プログラム名		戦略的イノベーション創造プログラム(SIP)「統合型ヘルスケアシステムの構築」
研究開発課題名		A-3 臨床情報プラットフォームと連携した PHR によるライフレコードデジタルツイン開発
研究開発 担当者＊ 1	機関名	国立大学法人東京大学
	所属	医科学研究所
	役職	特任教授
	氏名	鈴木 亨

実施期間＊2	令和5年9月 22 日～令和7年3月 31 日
--------	-------------------------

＊1 委託研究開発契約書に定義

＊2 年度の契約に基づき、本委託研究開発を行った期間又は中止までの期間

1. 研究開発テーマ概要

1.1 研究開発内容

本研究では個人のヘルス・健康・医療情報をもとにライフレコード型デジタルツインを開発し、実用化する。

個人のヘルスケア・医療情報が集約された情報基盤上の蓄積データに疾患リスクの予測アルゴリズムを適用し、疾患予測と経過予測を算出して、個人のヘルス・デジタルツインが形成される。このデジタルツインに、さらに個人のリアルタイムでのヘルスケア・医療情報(PHR 等)を追加することで、疾病予測や疾病予防法や治療法がアップデートされ、個人や医療者にとって必要な情報を適切に提供できるシステムがライフレコード型デジタルツインである。

まず未病・健常な段階から疾病発症さらにその後の経過・転帰に関する経時的なデータ(=ライフレコード)を整備し、次にライフレコードを用いて疾病リスクの可視化から予防的手法から治療法選択などに関する情報を提供するアルゴリズムを算出し、デジタルツインを開発する。

本研究におけるライフレコード整備は、産学連携パートナーである日本電信電話株式会社(NTT)グループ社員を対象に、入社時からの縦断的な健康・医療情報と生体情報に加え、本研究で収集していく疾患発症後の病院等での治療経過や追跡情報を融合して実施する。このライフレコードに対し、データサイエンスに基づくデータ解析手法を用いて疾病リスクを分析・可視化するアルゴリズムを搭載し、研究代表者が兼任・担当している英国レスター大学ならびに附属病院において進めている疾患経過を中心としたライフレコード型の医療デジタルツイン開発計画と連携し、国際的な標準化と普遍化も視野に入れた国際的人的ネットワークを形成する。

1.2 研究開発実施体制

体制を以下に示す。



図1 体制図

【経理様式1 別添】【R6】

1.3 研究推進のスケジュール

スケジュールを以下に示す。

表1 スケジュール

項目	初年度 (R5.9～ R6.3)	2年度 (R6.4～ R7.3)	3年度 (R7.4～ R8.3)	4年度 (R8.4～ R9.3)	最終年度 (R9.4～ R10.3)
【1】コホート構築と得られた 経時的な医療・ヘルスケア データのデータベース化	← 設計・構築	← データ収集・改良	← データ収集・改良	← データ収集・改良	← データ収集・改良
【2】データサイエンスによる 疾病リスクのアルゴリズム演算	← サーバ構築・実データ導入	← アルゴリズム開発	← アルゴリズム開発	← アルゴリズム開発	← アルゴリズム開発
【3】デジタルツインの システム化と ユーザインタフェースの開発	← 設計・構築	← 設計・構築	← 改良	← 改良	← 改良
【4】デジタルツインの実用化と エコシステム構築	← 設計・開発	← 設計・開発	← ビジネスモデル検証・調査	← ビジネスモデル検証・調査	← ビジネスモデル検証・調査
【5】医療デジタルツインの データサイエンティスト人材の 育成	← 人材募集	← 人材募集	← デジタルツイン構築	← デジタルツイン構築	← デジタルツイン構築

2. 本年度の成果・進捗の概要

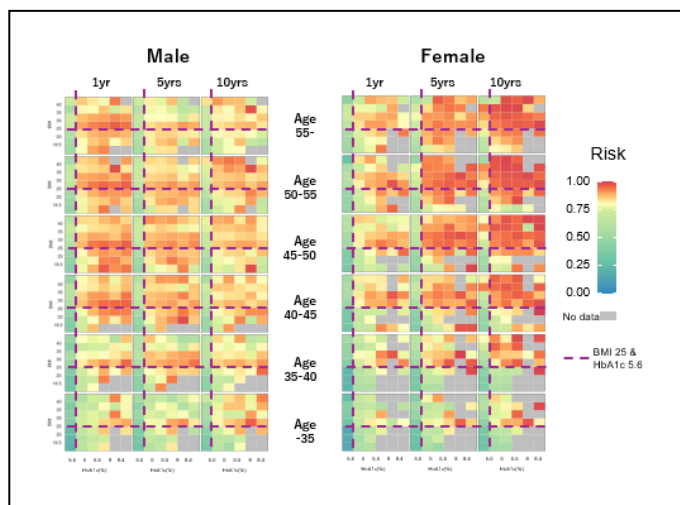


図2 リスクマップ(耐糖能異常、1年後、5年後、10年後)



図3 プレサービス・アプリ画面

令和5年度に作成したユーザインタフェースをもとに、本年度はプレサービスのシステム開発を行った。電子カルテおよび健診データを統合した機微情報を扱うことから、NTT グループの情報セキュリティガイドラインに則り、セキュリティの担保されたデータセンタ内にオンプレミスサーバを設置している。プレサービスのシステムに関する情報セキュリティ対策として、ウェブアプリケーションファイアウォール(ウェブ攻撃の監視・遮断)を設置し、情報セキュリティ対策の維持・管理・更新と緊急時対応に向けた体制を整備、昨今増加傾向にあるサイバー攻撃への対策を実施した。情報セキュリティに関するルールの文書化をおこない、継続的なチェックと必要な更改を維持している。

疾患リスク予測モデルとして、約5万人で平均17.7年の蓄積のある健診データを用い、1、5、10年後に耐糖能異常、高血圧、脂質異常の3疾患について、特定保健指導対象および要精密検査(人間ドック学会基準)となるリスクを機械学習(LightGBM)より算出した。算出されたリスクについて、ユーザへの説明と妥当性の学会等での議論に向け、各疾患における指標(耐糖能異常:HbA1c、高血圧:収縮期血圧、脂質異常:LDL濃度)とBMIを軸としヒートマップ形式で表現した(図2)。このリスクマップは、20～65歳までの5歳刻み男女別で作成しており、各ヒートマップのセルを分ける値については、厚労省からのNDBオープンデータをもとに設定し、セルごとに含まれる人(群)におけるリスクの中央値をヒートマップの色で表現している。3疾患のリスクマップについては、2025年3月の日本循環器学会学術集会で発表し、医師・専門家向けのリスクマップ活用に向けたコンセンサス構築を図っている。また、プレサービスの疾患リスク予測サービスにユーザ登録した者は、自身の健診データから1、5、10年後の各疾患におけるリスク値をグラフ表示および根拠となるリスクマップの両方で閲覧できる(図3)。

ユーザ個人への行動変容に向けた指導を人的リソースのコストも低減しながら効率よく実施するため、AI保健指導の開発において、令和6年度は生成AIの適用性について検討した。30名の対象者への保健師の指導内容とアウトカムについて、個人情報を削除した上で生成AIに学習させ、保健師

と同レベルの指導文を生成可能か検証した(Llama 3.1 Swallow 8B Instruct のベースモデルを使用)。特定保健指導に準じた食事指導の実施を当面の目標とし、指導対象者の問題意識レベルを抽出し、レベルに応じた指導文を生成する過程の生成 AI の適用性を評価した。AI が生成する指導文に対して、保健師が判断する「評価指標」の作成も本検討の範囲である。指導対象者の問題意識に対応した指導文について、今回作成した「評価指標」に基づき 85%の割合で正解と判定された。一方で、専門的な質問(例:断食)に対して、適切な指導文を生成できていなかった。このように保健や医療に関する基礎知識不足が判明したことから、本知見を令和7年度以降のAI保健指導の開発に反映する(図4)。

医療デジタルツイン構築と社会実装を担うデータサイエンティスト育成に向けて、令和6年度からはアカデミア人材を確保した。また、ライフレコード型デジタルツイン研究の加速のため、東京大学医科学研究所内に健診、診療データを扱う生成 AI 含めた開発環境として、GPU サーバを設置し Llama 3.1 Swallow 70B Instruct ベースモデルをインストールし、AI 検討を開始している。

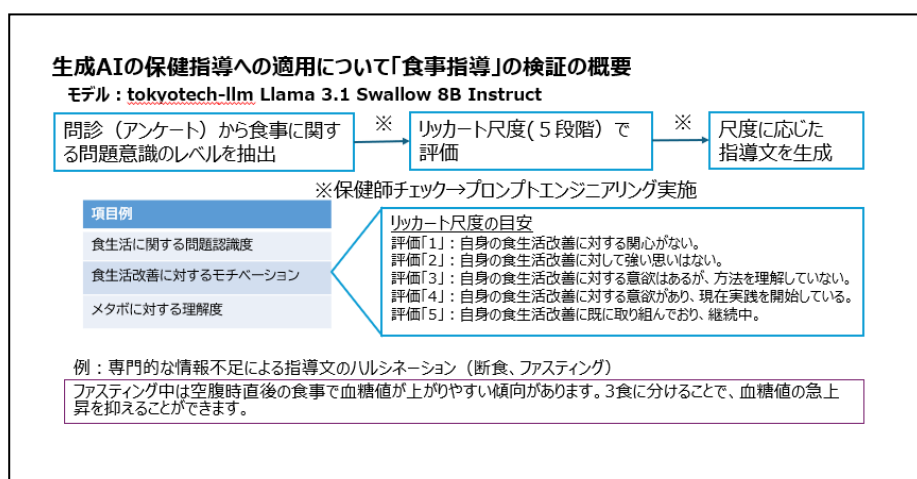


図4 生成 AI を用いた保健指導 AI の検討の概要

3. 成果物の公表

3.1 論文など(原著論文、学位論文、プロシーディングス、総説、解説、速報など)

現時点なし。

3.2 学会発表など((国内・国際)学会口頭発表・ポスター発表、招待講演で成果を公表したもの)

1. Predictive Map of Lifestyle-related Disease Risk from Big Data Analysis of 50,000 Employee's Health Checkup Data, 齊藤剛仁, 鈴木亨, 瀬山倫子, 第 89 回日本循環器学会学術集会, 2025/3/30, 国内, ポスター発表

3.3 診療ガイドライン、省令、基準、日本薬局方、添付文書改訂、国の技術文書(通知)等への反映

現時点なし。

3.4 研修プログラム、カリキュラム、シラバス、教材、e-learning 等の公表

現時点なし。

3.5「国民との科学・技術対話」に対する取り組み

現時点なし。

3.6 その他

現時点なし。

以上