

厚生労働科学研究費補助金
(臨床研究等ICT基盤構築・人工知能実装研究事業)
令和5年度 総括研究報告書

ナショナルデータベースや介護保険総合データベース等を活用した
医療・介護特性を総合的に捉えたAIプロトタイプの開発と分析結果を
根拠とした医療介護特性格別の最適介入 (22AC1005)

研究代表者 福井小紀子 (東京医科歯科大学 大学院保健衛生学研究科・教授)

研究分担者 廣岡佳代 (東京医科歯科大学 大学院保健衛生学研究科・准教授)
岡田就将 (東京医科歯科大学 大学院保健衛生学研究科・教授)
浜野淳 (筑波大学 医学医療系 講師)
高橋邦彦 (東京医科歯科大学 大学院保健衛生学研究科・教授)
佐藤一樹 (名古屋大学 大学院・教授)
菅野雄介 (東京医科歯科大学 大学院保健衛生学研究科・講師)

研究要旨

本研究課題は、病院・施設・在宅という場を超えたケアの質の担保、及び、医療経済的観点を統合した住民個々への医療介護資源の効率的な再分配、健康・医療分野における行政政策に資する科学的根拠の創出を目指し、医療と介護特性を総合的に考慮したAI (artificial intelligence) のプロトタイプを活用した国等が保有する公的データの有用性の検証と課題を明確化することである。今年度は、以下3点、①レセプト情報・特定健診等情報データベース (以下、NDB) と介護保険総合データベース (以下、介護DB) の利用申出書を作成し、審査に向けた利用申請、②予測モデルの予測精度向上に向けた統計解析研究計画の精緻化、③静岡県の国保データベース (KDB) を用いたフィージビリティスタディを実施することを目的とした。研究計画は、1) NDBと介護DBの利用申出書を作成と利用申請、2) 予測モデルの予測精度向上に向けた統計解析研究計画の精緻化、3) 静岡県KDBを用いたフィージビリティスタディを挙げた。本年度は、令和4年度に医療や介護現場で臨床や研究を専門とする分担研究者と協議を重ね立案した研究疑問18個をもとに、NDBと介護DBの利用申出書を作成し、審査に向けた利用申請を行った。通常、利用申出の承認日から提供までの期間が300日程度を要するため、待機期間には立案した研究疑問を研究者間で再協議し、また新たな研究疑問を立案し、統計解析計画書の精緻化を進めた。さらに、AIプロトタイプ開発に向け解析手法や予測精度を検討するため、作成した統計解析計画書にもとに、静岡県の国保データベース (KDB) を用いたフィージビリティスタディを実施した。この結果をもとに、NDBと介護DBを突合した場合の症例数や解析結果等を考慮し、予測精度を高める変数の見直し等を行った。なお、介護DBでは2024年2月に提供を受け、データクリーニングを実施中である。NDBでは2024年6月頃に提供予定である。今後は、NDBが提供され次第、すでに提供されている介護DBとの連結を行い、AI分析用データセットを作成する。データセットをもとに、生存率や再入院等をアウトカム指標とした予後予測等のAI分析をすることで、予後予測精度の検証や、公的データの横断的な活用可能性や有用性を検証する。

A. 研究目的

AI戦略2022では研究データ基盤等へのAI (artificial intelligence) の利活用を掲げている。それを踏まえ、「保健医療分野AI開発加速コンソーシアム」では、AIを用いることで保健医療サービスの高度化・現場の負担軽減を目指すことが掲げられている。そのデータ源の一つとして、NDBと介護DBを活用することが提案されている。

レセプト情報・特定健診等情報データベース (NDB) と介護保険総合データベース (介護DB) 等を連結し、主疾患、併存疾患、要介護度やIADL (Instrumental Activities of Daily Living) 等の医療と介護特性を考慮した予後予測等のAI分析プロトタイプを開発することで、個々の特性に応じた適時適切な介入にあわせ、効果的な医療介護人材の活用につなげられる。NDB等を用いてAI分析を行い、予後予測等の改善や科学的根拠に基づく医療介護サービス提供のあり方を示していくことが、超高齢人口減少社会が延伸する我が国にとって不可欠である。

本研究課題は、病院・施設・在宅という場を超えたケアの質の担保、及び、医療経済的観点を統合した住民個々への医療介護資源の効率的な再分配、健康・医療分野における行政政策に資する科学的根拠の創出を目指し、医療と介護特性を総合的に考慮したAIのプロトタイプを開発し、国等が保有する公的データでの有用性の検証と課題を明確化することである。今年度は、以下3点、①NDBと介護DBの利用申請書を作成し、審査に向けた利用申請、②予測モデルの予測精度向上に向けた統計解析研究計画の精緻化、③静岡県国保データベース (KDB) を用いたフィージビリティスタディを実施した。

B. 研究方法

上記の今年度実施した3点の研究方法は以下の通りである。

1) NDBと介護DBの利用申請書作成と提出 (代表者：福井、分担者：廣岡、菅野)

令和4年度に立案した、研究疑問18個をもとに、文献レビューやがん、在宅医療などを専門とする研究分担者との協議のうえ、解析に必要な変数を選定し、NDBと介護DBの2つの利用申請書一式を作成した。作成にあたり、NDBと介護DBの利用

申請書の作成の経験を持つ専門家の助言を受けながら進めた。

2) 予測モデルの予測精度向上に向けた統計解析研究計画の精緻化 (代表者：福井、分担者：廣岡、高橋、佐藤、岡田、浜野、菅野)

通常、利用申出の承認日から提供までの期間が300日程度を要するため、待機期間で、立案した研究疑問を研究者間で再協議し、また新たな研究疑問を立案し、AIプロトタイプ開発に向けた統計解析計画書の精緻化を進めた。研究疑問の再検討においては、各々の研究分担者の専門性を活かし、医療と介護の臨床家 (浜野)、在宅・緩和ケアを専門とする看護学研究者 (福井・廣岡)、がん・緩和ケアを専門とする看護学研究者 (佐藤、菅野)、医療政策 (岡田)、統計の専門家 (高橋) など、多角的な視点で行った。

3) 静岡県KDBを用いたフィージビリティスタディの実施

次年度、NDBと介護DBを用いたAIプロトタイプを開発するため、フィージビリティスタディを実施し、その解析結果をもとに変数の見直しと解析手法を検討した。

データベースは、静岡県のKDBを用いた。このデータベースは、浜松市、静岡市清水区、伊東市を除く、静岡県内の国保・後期高齢者医療制度の医科・DPC・調剤・介護レセプト (2012年4月~2018年9月診療分) と被保険者台帳情報が格納されていた。

解析は、立案した研究疑問から2個選定し、統計解析書に基づき、要介護度、性別、年齢、医薬品、傷病名、訪問看護等の診療行為、介護サービスの利用状況を予測変数として、ロジスティック回帰モデルと4つの機械学習のモデル (random forest、neural network、eXtreme Gradient Boosting、Stochastic Gradient Boosting) を作成し、モデル別に予測精度を検証した。

倫理面への配慮として、本研究は、NDBや介護DBを用いて解析するため、倫理指針に基づきインフォームド・コンセントを要しない研究である。ただし、利用する各種データベースは匿名化処理が施されているものの、患者個人の年齢や性別など基本的な情報や診療に関する情報も含まれるた

め、他の情報と照合することにより患者の特定可能性が高められることを否定できない。また、AI開発において産学連携も視野に入れていることから、情報の機密保持が一層求められる。以上のことから、本研究では、各種データに対し、厚労省が提示する「匿名レセプト情報・匿名特定健診等情報の提供に関するガイドライン」を遵守し、また産学連携においては、本学の産学連携研究センターを通じ秘密保持契約を締結した上で共同研究を進めることとする。本年度は、上記のガイドラインを遵守することを明記した研究計画書を作成し、研究代表者の研究倫理審査委員会に提出し承認を得た。

C. 研究結果

本年度は、医療と介護特性を総合的に考慮したAIのプロトタイプを開発するために、主に、1) NDBと介護DBの利用申出書を作成し、審査に向けた利用申請、2) 予測モデルの予測精度向上に向けた統計解析研究計画の精緻化、3) 静岡県国保データベース（KDB）を用いたフィージビリティスタディを行った。

1) NDBと介護DBの利用申出書の作成

本年度は立案した研究計画の通り、医療や介護現場で臨床や研究を専門とする分担研究者と協議を重ね、NDBと介護DBの利用申出書を作成し、審査に向けた利用申請を行った。申請期間中、厚生労働省の担当課と連絡調整し、申請状況を定期的に確認した。

介護DBの利用状況は、介護DBでは2024年2月に提供を受け、データクリーニングを実施中である。NDBでは2024年6月に提供見込みの予定である。

2) 予測モデルの予測精度向上に向けた統計解析研究計画の精緻化

各研究分担者が専門的立場から、立案した研究疑問に対し意見交換を行い、収集する変数やアウトカムの設定等、統計解析研究計画の精緻化を進めた。具体的には、認知症やがん、循環器疾患等、わが国の主要疾患に関し、臨床や介護現場の現状に沿った研究疑問として、NDBと介護DBで取得可能な変数の選定（例：疾患、薬品名、要介護度など）を行った。その際、AIプロトタイプ開発

のユーザーを誰に設定するのか、開発後の実装も含め検討を行った。

3) 静岡県KDBを用いたフィージビリティスタディの実施

本研究では、以下の2つの研究疑問について、静岡KDBを基にフィージビリティスタディを実施した。

(1) 要介護度3以上の認知症を有するがん外科手術を受けた患者における手術月の翌月から12か月時点の要介護度2段階以上の悪化を予測する

(2) 新規に要介護認定を受けた要介護2以下の利用者における認定月の翌々月から12か月時点の要介護度の悪化を予測する

(1) の解析対象集団は、221名であった。要介護度、性別、年齢、医薬品、傷病名、訪問看護等の診療行為、介護サービスの利用状況を予測変数として、ロジスティック回帰モデルと4つの機械学習のモデル（random forest、neural network、e Xtreme Gradient Boosting、Stochastic Gradient Boosting）を構築した。C統計量の値は、0.683~0.780に分布しており、良好な予測精度（識別能）であった。

(2) の解析対象集団は、20,154名であった。上記と同様の予測変数と統計モデルにより予測モデルを構築した。C統計量の値は、0.721~0.738に分布しており、良好な予測精度（識別能）であった。以上の2つの研究課題の結果から、医療レセプトと介護レセプトを用いて、一定の精度の予測モデルを構築できることが確認された。1県のデータの場合、特に研究課題（1）では対象数が十分でなく、その予測精度が低下することから、今後、NDB・介護DBでの解析が求められることが明らかになった。

これらの結果について多角的な視点から、さらに分担研究者間で協議を重ね、データ提供後の解析準備に向けて進める予定である。

なお、介護DBは提供済みであり、今後提供予定のNDBとの連結、AI解析に向けてデータクリーニングを実施済みである。

D. 考察

本年度は、前年度に協議し立案した研究疑問に対し、次年度以降の準備期間として、静岡県KDB

を用いたフイージビリティスタディを実施し、訪問看護、在宅医療、緩和ケア・総合診療、医療政策等の背景を持つ研究者とともに多角的な視点から、予後予測モデルのAIプロトタイプ開発に向けて協議し、研究疑問や統計解析計画書の精緻化等を進めることができた。また、NDBと介護DBの利用申請に向け、厚生労働省の担当課の協力を得ながら、早期に取得できるように働きかけ、介護DBを取得することができた。今後は、NDBが提供され次第、すでに提供されている介護DBとの連結を行い、AI分析用データセットを作成する。生存率や再入院等をアウトカム指標とした予後予測等のAI分析をすることで、予後予測精度の検証や、公的データの横断的な活用可能性や有用性を検証する。データ利用が可能となる期間はAIプロトタイプ開発に向けた統計解析計画書の精緻化を進める予定である。

E. 結論

本年度は、医療と介護特性を総合的に考慮したAIのプロトタイプを開発するために、医療や介護現場で臨床や研究を専門とする分担研究者と協議を重ね、立案した研究疑問の測精度向上に向けた統計解析研究計画の精緻化に加え、静岡県KDBを用いたフイージビリティスタディを実施した。

今後は、NDBが提供され次第、すでに提供されている介護DBとの連結を行い、AI分析用データセットを作成する。生存率や再入院等をアウトカム指標とした予後予測等のAI分析をすることで、予後予測精度の検証や、公的データの横断的な活用可能性や有用性を検証する。データ利用が可能となる期間はAIプロトタイプ開発に向けた統計解析計画書の精緻化を進める予定である。

F. 健康危険情報

該当なし

H. 知的財産権の出願・登録状況（予定を含む。）

1. 特許取得

該当なし

2. 実用新案登録

該当なし

G. 研究発表

該当なし