

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)4年度～7年度 総合研究報告書

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別の最適介入 (22AC1005)

研究代表者 福井小紀子 (東京科学大学 大学院保健衛生学研究科・教授)

研究分担者 菅野雄介 (東京科学大学 大学院保健衛生学研究科・准教授)
高橋邦彦 (東京科学大学 大学院保健衛生学研究科・教授)
佐藤一樹 (名古屋大学 大学院医学系研究科・教授)
岡田就将 (東京科学大学 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 准教授)
廣岡佳代 (東京科学大学 医学部・准教授)

研究協力者 坂野朋未 (東京科学大学 大学院保健衛生学研究科・講師)

研究要旨

本研究課題は、匿名医療保険等関連情報データベース (NDB) および介護保険総合データベース (介護DB) 等の公的データを活用し、医療と介護の特性を総合的に考慮したAI (artificial intelligence) 分析プロトタイプを開発するとともに、公的データの有用性と課題を明確化することを目的とした。具体的には、主疾患、併存疾患、要介護度、医療サービス利用、介護サービス利用等の情報を統合し、予後予測や医療・介護費予測に資する機械学習モデルを構築することを目指した。これにより、病院・施設・在宅という場を超えたケアの質の担保、医療経済的観点を踏まえた医療介護資源の効率的な再配分、ならびに厚生労働行政に資する科学的根拠の創出を図ることを目指した。

研究期間全体を通じて、本研究は3つの段階で進めた。第一段階では、臨床医、看護学研究者、医療政策専門家、統計学者等からなる学際的研究チームを構築し、入退院支援看護師や介護支援専門員等の実践家へのヒアリングを通じて、現場の課題に基づく臨床研究疑問を整理した。その結果、がん、認知症、骨折、循環器疾患、透析、高額医療・介護費等に関する20個の臨床研究疑問を立案し、NDBおよび介護DBを用いた解析に向けて統計解析計画を精緻化した。

第二段階では、国レベルのNDB・介護DB解析に先立ち、KDBを用いたフィージビリティスタディを実施した。要介護度3以上の認知症がん患者や新規要介護認定者等を対象として、Random Forest、eXtreme Gradient Boosting (XGBoost) 等の機械学習モデルを用いた予測モデル構築の実行可能性を検証した。その結果、一部の解析ではAUC 0.7前後の一定の予測精度が得られた。一方で、地域限定データでは、希少疾患や特定の病態において症例数が不足し、一般化可能性にも課題があることが確認された。この結果は、全国規模のNDB・介護DB連結データを用いた解析の必要性を示す予備的知見となった。

第三段階では、提供を受けたNDBおよび介護DBを連結し、当初立案した20個の臨床研究疑問のうち、

症例数、実行可能性、臨床的優先度、政策的意義を踏まえて7つの重点研究疑問を選定した。これらの重点課題について、機械学習を用いた予後予測モデルおよび医療・介護費予測モデルを構築し、AUC、感度、特異度、陽性的中率、陰性的中率等により性能を評価した。その結果、がん・認知症領域、慢性疾患・透析領域、骨折・費用予測領域のいずれにおいても、個別患者の1年程度先の予後や高額費用発生を、臨床現場での意思決定に直ちに活用できる水準で予測することには課題が残ることが明らかとなった。特に、AUCが一定程度得られる場合であっても、感度または陽性的中率に課題が残り、ハイリスク者を抽出するスクリーニングモデルとしては実用化が推奨される水準には至らない研究疑問が多かった。

研究期間の最終段階では、政策的・実務的価値が高く、早期の介入による医療・介護資源の適正化効果が期待される課題として、大腿骨近位部骨折後の要介護度悪化または死亡の予測、および高齢者の高額医療・介護費予測の2課題を重点的に検討した。大腿骨近位部骨折後の要介護度悪化または死亡の予測では、整形外科専門医へのヒアリングに基づき、併存疾患、薬剤、術式、入院日数、外来受診回数等の変数を精緻化し、対象集団を手術実施日に要支援1から要介護2までの認定を受けていた軽度要介護者に限定した。しかし、精緻化後のAUCは0.67～0.69にとどまり、精緻化前のAUC 0.71～0.72と比較して改善は認められなかった。高齢者の高額医療・介護費予測では、傷病名、診療行為、薬剤情報、要介護認定情報を精緻化し、対象を要支援認定者に限定したが、感度は11.1～11.8%にとどまり、精緻化前の9.3～10.4%からの改善は限定的であった。

以上より、本研究では、現行のNDB・介護DB連結データのみを用いて、1年程度先の要介護度悪化、死亡、高額医療・介護費発生を、個人単位で臨床現場の意思決定に直ちに活用できる水準で予測することには課題が残ることが明らかとなった。この背景には、1年程度先の転帰が、ICU入室中などの短期予後と比べて患者の個別性や時間経過中の出来事の影響を受けやすく、予測困難性が高いことがある。加えて、これらの転帰は、疾患や医療利用だけでなく、身体機能、認知機能、栄養状態、生活環境、家族介護力等の多様な因子に影響される一方で、現行の公的データベースにはこれらの情報が十分に含まれていないことも、予測性能を制約した可能性がある。また、傷病名情報や介護サービス情報が月単位で把握されるなど、時間情報の粒度にも制約がある。

一方で、本研究により、特定患者群における要介護度悪化、死亡、高額医療・介護費発生の実態を定量的に把握できたことは重要な成果である。公的データベースの強みは、個別患者の1年程度先の転帰を高精度に予測することよりも、全国規模で患者群の規模、転帰の発生割合、医療・介護サービス利用、費用分布を記述できる点にある。したがって、これらの知見は、個別予測AIとして直ちに社会実装するための根拠としては限定的である一方、記述疫学的知見に基づく支援対象群の可視化、ポピュレーション・アプローチの検討、地域差・施設差の把握、政策介入対象の優先順位付けに資する基礎資料として活用可能である。

今後は、NDB・介護DBに加えて、DPCデータ、科学的介護情報システム（LIFE）、特定健診・後期高齢者健診データ等との連結可能性を検討する必要がある。ただし、これらのデータを追加しても、1年程度先の個別予測精度が直ちに大幅に改善するとは限らない。むしろ、ADL、認知機能、栄養状態、排泄、口腔、褥瘡、リハビリテーション実施状況等の生活機能やケアの質に近い情報を含めることで、死亡や要介護度悪化だけでなく、より実践的なアウトカムを設定し、患者群の状態や医療・介護サービス利用の実態をより精緻に記述できる可能性がある。したがって、時点情報の標準化、構造化項目の整備、データ品質の確保、アウトカム定義の精緻化を並行して進める必要がある。本研究の成果は、医療DXおよび全国医療情報プラットフォームを通じた公的データ利活用基盤を検討するうえで、現行データの有用性と課題を示す基礎資料となる。

A. 研究目的

わが国では、超高齢社会の進展に伴い、医療と介護の双方を必要とする高齢者が増加している。病院、施設、在宅という療養の場を超えて、医療サービスと介護サービスを適時適切に提供するためには、疾患、併存疾患、要介護度、医療サービス利用、介護サービス利用等の情報を統合し、対象者の状態に応じた支援や資源配分を検討する必要がある。また、医療・介護需要の増大と社会保障費の伸びを踏まえると、限られた医療介護資源を、支援を必要とする対象者へ効率的に配分するための科学的根拠が求められる。

近年、保健医療分野では、AI (artificial intelligence) や機械学習を活用し、予後予測、リスク層別化、医療・介護サービス利用の把握、政策介入対象の検討に資する分析を行うことが期待されている。匿名医療保険等関連情報データベース (NDB) および介護保険総合データベース (介護DB) は、全国規模で医療と介護の利用実態を把握できる公的データベースである。これらを連結して分析することにより、医療と介護の特性を総合的に捉えた予後予測や医療・介護費予測を行い、行政施策や現場の意思決定に資する知見を創出できる可能性がある。一方で、公的データベースを用いて、1年程度先の要介護度悪化、死亡、高額医療・介護費発生といった転帰を個人単位で高精度に予測できるかについては、十分に検証されていない。

本研究課題は、NDB、介護DB等の公的データを活用し、医療と介護の特性を総合的に考慮したAI分析プロトタイプを開発するとともに、公的データの有用性と課題を明確化することを目的とした。具体的には、主疾患、併存疾患、要介護度、医療サービス利用、介護サービス利用等の情報を統合し、要介護度悪化、死亡、高額医療・介護費発生等の予測に資する機械学習モデルを構築することを目指した。これにより、医療・介護をまたぐ対象者の状態変化を把握し、病院・施設・在宅という場を超えたケアの質の担保、医療経済的観点を踏まえた医療介護資源の効率的な再配分、ならびに厚生労働行政に資する科学的根拠の創出を図ることを目指した。

本研究では、まず、臨床医、看護学研究者、医療政策専門家、統計学者等からなる学際的な研

究体制を構築し、医療・介護現場の課題に基づく臨床研究疑問を整理することを目的とした。次に、国レベルのNDB・介護DB解析に先立ち、地域データを用いたフィージビリティスタディにより、機械学習モデル構築の実行可能性と課題を検討することを目的とした。さらに、NDB・介護DB連結データを用いて、複数の重点課題に対する予測モデルを構築し、AUC、感度、特異度、陽性的中率、陰性的中率等の指標に基づいて性能を評価することを目的とした。

加えて、本研究では、単に予測モデルの精度を評価するだけでなく、AI分析プロトタイプを医療・介護現場や行政施策にどのように活用できるかを検討することを目的とした。具体的には、個別患者の予後や高額医療・介護費発生を、臨床現場での意思決定に直ちに活用できる水準で予測できるかを検証した。また、予測精度に課題が残る場合でも、特定患者群の予後実態、要介護度悪化、死亡、高額費用発生の分布を定量的に示すことで、支援を要する患者群の可視化、ポピュレーション・アプローチの検討、政策介入対象の優先順位付けに資する知見を得ることを目指した。

最終的に、本研究は、NDB・介護DB等の公的データを用いた個別予測AIの実装可能性と課題を明らかにし、今後の医療DX、公的データベース連結、全国医療情報プラットフォームの整備に向けた示唆を得ることを目的とした。特に、現行の公的データベースのみで捉えられる情報と、生活機能、認知機能、栄養状態、家族介護力等の未把握情報との関係を整理し、今後追加を検討すべきデータ項目、連結可能性を検討すべきデータベース、設定すべきアウトカムを明確化することを目指した。さらに、公的データベースの強みを踏まえ、個別予測だけでなく、集団レベルでの実態把握や政策支援に資する記述疫学的活用可能性を検討することも目的とした。

B. 研究方法

本研究は、NDB、介護DB、および国保データベース (KDB) 等の既存公的データを用いた後ろ向き観察研究として実施した。医療と介護の特性を総合的に考慮したAI分析プロトタイプの開発可能性を検討するため、記述疫学的分析、機械学習を用いた予測モデルの構築、モデル性能の評価、

ならびに重点課題におけるモデル精緻化を段階的に行った。

研究開始時には、臨床医、看護学研究者、医療政策専門家、統計学者等からなる学際的な研究チームを構築した。研究班では、入退院支援看護師、介護支援専門員等の実践家へのヒアリングを行い、医療・介護現場における意思決定や支援上の課題を整理した。これらの意見を踏まえ、がん、認知症、骨折、循環器疾患、透析、高額医療・介護費等に関する20個の臨床研究疑問を立案した。各臨床研究疑問について、対象集団、基準日、説明変数、アウトカム、追跡期間、解析方法を検討し、NDBおよび介護DBを用いた解析に向けた統計解析計画を作成した。

NDBおよび介護DBの利用にあたっては、データ提供申出に必要な申出書類を作成し、専門委員会による審査に対応した。NDBおよび介護DBは、利用申出から提供までに一定の期間を要するため、この待機期間を活用して、臨床研究疑問の優先順位付け、変数定義、アウトカム定義、解析対象集団の包含・除外基準、モデル評価指標の検討を進めた。また、各研究疑問について、実行可能性、症例数、臨床的意義、政策的意義を研究班内で協議した。

国レベルのNDB・介護DB解析に先立ち、2012年から2018年までの静岡県32市町のKDBを用いてフィージビリティスタディを実施した。具体的には、要介護度3以上の認知症がん患者や新規要介護認定者等を対象に、機械学習による予測モデルを構築し、予測モデル作成の実行可能性を検証した。予測モデルには、Random Forest、eXtreme Gradient Boosting (XGBoost) 等を用いた。モデル性能は、AUC、感度、特異度、陽性的中率、陰性的中率等により評価した。この予備的検討により、地域データを用いた予測モデル構築の可能性と、地域限定データにおける症例数および一般化可能性の課題を確認した。

その後、提供を受けたNDBおよび介護DBを連結し、全国規模の公的データを用いた解析を行った。NDBからは、医科レセプト、DPCレセプト、調剤レセプト等に含まれる傷病名、診療行為、薬剤、入院・外来利用等の情報を用いた。介護DBからは、要介護認定情報、認定調査情報、介護レセプトに含まれる介護サービス利用情報等を用い

た。これらの情報を基準日前の説明変数として整理し、基準日以降の要介護度悪化、死亡、高額医療・介護費発生等をアウトカムとして定義した。

当初立案した20個の臨床研究疑問について、症例数、データ取得可能性、実行可能性、臨床的優先度、政策的意義を踏まえて精査した。その結果、症例数が極端に少なくモデル構築が困難と判断された課題や、他の研究疑問と内容が重複し統合可能と判断された課題を整理し、最終的に7つの重点研究疑問を選定した。これらの重点研究疑問について、Random Forest、XGBoost等の複数の機械学習モデルを構築し、予測性能を比較した。モデル評価では、AUCや正解率のみならず、ハイリスク者抽出やスクリーニングへの活用可能性を評価するため、感度、特異度、陽性的中率、陰性的中率を含めて総合的に判断した。

解析対象となる各研究疑問では、疾患、治療、介護状態、医療・介護サービス利用状況に基づいて対象集団を定義した。基準日は、手術日、退院日、治療開始日、初回要介護認定月、サービス開始月等、各研究疑問の臨床的文脈に応じて設定した。追跡期間は、基準日以降の一定期間とし、主に1年程度の転帰を評価した。アウトカムには、要介護度の悪化、死亡、要介護度悪化または死亡の複合アウトカム、高額医療・介護費発生等を用いた。あわせて、予測モデルの性能評価だけでなく、対象集団の転帰発生割合、医療・介護サービス利用、費用分布等を把握する記述疫学的分析も行った。

最終段階では、前段階の解析結果を踏まえ、政策的・実務的価値が高く、早期の介入による医療・介護資源の適正化効果が期待される2課題を重点課題として選定した。

第一の課題は、大腿骨近位部骨折後の要介護度悪化または死亡の予測である。この課題では、整形外科専門医へのヒアリングに基づき、高血圧、心疾患、脳血管疾患、糖尿病等の併存疾患、内服薬の種類、術式、入院日数、外来受診回数等を精緻化した。また、予測モデル実装後に予後改善可能性がある集団を特定することを想定し、対象集団を手術実施日に要支援1から要介護2までの認定を受けていた軽度要介護者に限定した。基準日は手術実施日とし、アウトカムは術後365日時点における要介護度の2段階以上の悪化または死亡

に再設定した。

第二の課題は、高齢者の高額医療・介護費予測である。この課題では、医療・介護資源の重点的な配分が必要となる対象者を把握することを目的として、傷病名、診療行為、薬剤情報等の医療情報を精緻化し、要介護認定情報を予測変数として追加した。解析対象は、介護予防や早期支援の観点から、要支援認定者に限定した。基準日は初回の認定有効期間の開始日とし、アウトカムは基準月の翌月から12か月間の年間医療・介護費が上位5%となることとして定義した。精緻化後のモデルについても、AUC、感度、特異度、陽性的中率、陰性的中率等を用いて性能を評価した。

解析にあたっては、請求情報由来の傷病名の解釈、死亡を含む複合アウトカムの扱い、初期要介護度による悪化余地の違い、イベント不均衡下における評価指標の解釈に留意した。特に、高額医療・介護費発生のようにイベント割合が限られるアウトカムでは、正解率や特異度が高くても感度が低い場合がある。そのため、モデルの実用化可能性は、単一の指標ではなく、想定される用途に応じて複数の評価指標を組み合わせて検討した。また、1年程度先の転帰は、短期予後に比べて患者の個性や時間経過中の出来事の影響を受けやすいことを踏まえ、モデル性能の解釈には慎重を期した。なお、予測モデルの課題を踏まえ、研究期間中に因果推論を取り入れた新たな解析も計画し、データ変更申出を行った。しかし、データ提供時期の関係により、研究期間内に最終的な因果効果の推定までは至らなかった。そのため、本総合研究報告書では、主に予測モデル構築、モデル精緻化、ならびに公的データ活用の可能性と課題の検討結果を報告する。

倫理面への配慮として、本研究は、匿名化されたNDB、介護DB、KDB等を用いて実施する研究であり、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。提供されるデータには、厚生労働省が審査支払機関等から提供を受ける時点で個人情報に含まれておらず、ハッシュ化により復元不可能な形で管理されている。本研究では、個人の特定可能性を低減するため、詳細な地域情報を使用しないなどの措置を講じた。また、各種データの利用にあたっては、厚生労働省が提示する匿名レセプト情報等に関するガイド

ラインを遵守した。研究計画は、研究倫理審査委員会の承認を得て実施した。

C. 研究結果

本研究では、NDBおよび介護DB等の公的データを活用し、医療と介護の特性を総合的に考慮したAI分析プロトタイプの開発可能性を検討した。研究期間全体を通じて、研究基盤の構築と臨床研究疑問の立案、KDBを用いたフィージビリティスタディ、NDB・介護DB連結データを用いた重点研究疑問の解析、ならびに最終段階における重点課題のモデル精緻化を行った。

1) 研究基盤の構築と臨床研究疑問の立案

研究開始後、臨床医、看護学研究者、医療政策専門家、統計学者等からなる学際的な研究チームを構築した。また、入退院支援看護師および介護支援専門員等の実践家へのヒアリングを行い、医療・介護現場における意思決定や支援上の課題を整理した。これらの検討を踏まえ、がん、認知症、骨折、循環器疾患、透析、高額医療・介護費等に関する20個の臨床研究疑問を立案した。

NDBおよび介護DBは、利用申出からデータ提供までに長期間を要することから、データ提供までの期間を活用して、各臨床研究疑問の統計解析計画を精緻化した。具体的には、対象集団、基準日、説明変数、アウトカム、追跡期間、解析方法、モデル評価指標を検討し、NDB・介護DB連結データを用いた予測モデル構築に向けた準備を行った。

2) KDBを用いたフィージビリティスタディ

国レベルのNDB・介護DB解析に先立ち、2012年から2018年までの静岡県32市町のKDBを用いてフィージビリティスタディを実施した。要介護度3以上の認知症がん患者および新規要介護認定者等を対象として、Random Forest、eXtreme Gradient Boosting (XGBoost) 等の機械学習モデルを用いた予測モデル構築の実行可能性を検証した。

KDBを用いた解析では、要介護度3以上の認知症がん患者および新規要介護認定者を対象とした一部の予測モデルにおいて、AUC 0.7前後の一定の識別能が得られた。一方で、地域限定データ

では、希少疾患や特定の病態において症例数が不足し、モデル構築や精度評価に制約が生じることが確認された。また、地域データで得られた結果を全国レベルに一般化するには限界があることも明らかとなった。これらの結果は、全国規模のNDB・介護DB連結データを用いた解析の必要性を裏付ける予備的知見となった。

3) NDB・介護DB連結データを用いた重点研究疑問の選定と解析

2024年2月に介護DB、同年8月にNDBの提供を受け、両データベースを連結した。連結後、当初立案した20個の臨床研究疑問について、症例数、実行可能性、データ取得可能性、臨床的優先度、政策的意義を踏まえて精査した。その結果、症例数が極端に少なくモデル構築が困難と判断された課題や、他の研究疑問と内容が重複し統合可能と判断された課題を整理し、最終的に7つの重点研究疑問を選定した。

選定した7つの重点研究疑問について、Random Forest、XGBoost等の機械学習モデルを用いて予測モデルを構築し、AUC、感度、特異度、陽性的中率、陰性的中率等により性能を評価した。その結果、個別患者の1年程度先の予後や高額医療・介護費発生を、臨床現場での意思決定に直ちに活用できる水準で予測することには課題が残ることが明らかとなった。多くの課題では、AUCが一定程度得られる場合であっても、感度または陽性的中率に課題が残り、ハイリスク者を抽出するスクリーニングモデルとしては十分な性能に至らなかった。

4) がん・認知症領域における解析結果

がん・認知症領域では、要介護認定を受けた認知症がん患者における外科的手術後の要介護度悪化、および認知症を有する要介護中重度者における肺炎発症入院後の要介護度悪化等を対象に予測モデルを構築した。これらの解析では、術後1年または退院後1年の要介護度悪化や死亡をアウトカムとして評価した。その結果、予測モデルのAUCはおおむね0.60～0.65程度にとどまり、個別予測モデルとして十分な識別能を得ることは困難であった。一方で、記述疫学的には、認知症がん患者では術後1年で約30%が要介護度悪

化または死亡となり、肺炎で入院した認知症を有する要介護中重度者では退院後1年で約40%が要介護度悪化または死亡となることが確認された。これらの結果は、個別予測AIとしての実装には課題が残る一方で、手術適応、退院支援、ACP、ケアプラン見直し等が必要となりやすい患者群を把握する基礎資料として有用であると考えられた。

5) 慢性疾患・透析領域における解析結果

慢性疾患・透析領域では、肝硬変による難治性腹水患者におけるトルバプタン導入後の死亡、ならびに慢性維持透析患者における要介護度悪化または死亡を対象として予測モデルを構築した。これらの解析では、治療開始後または基準日以降の一定期間における死亡や要介護度悪化をアウトカムとして評価した。

モデル性能を検証した結果、透析領域では特異度が高い一方で感度が10%前後にとどまり、支援を必要とする対象者を十分に抽出することは困難であった。一方、肝疾患領域では感度が高い場合でも特異度が40%未満にとどまり、ハイリスク者の抽出や臨床現場でのスクリーニング用途には課題が残った。したがって、慢性疾患・透析領域においても、現行のNDB・介護DB連結データのみを用いて、個別患者の予後を実用的な水準で予測することには課題が残ると考えられた。

6) 骨折・費用予測領域における解析結果

骨折・費用予測領域では、股関節手術後の要介護度悪化、および介護予防・日常生活支援総合事業の開始1年後の高額医療・介護費発生を対象として予測モデルを構築した。股関節手術後の要介護度悪化予測では、1年後の要介護度1段階以上の悪化または死亡をアウトカムとして評価した。その結果、AUCは0.72前後、感度は約85%と比較的良好な傾向がみられた。しかし、特異度が低く、実用化には課題が残った。

介護予防・日常生活支援総合事業の開始1年後の高額医療・介護費発生予測では、医療・介護費が上位5%となることをアウトカムとしてモデルを構築した。モデルの正解率は約95%と高い値を示したが、AUCは約0.75、感度は約10%にと

どまった。高額費用発生者の割合が限定されるアウトカムでは、正解率が高くても、実際に支援対象となる高額費用発生者を十分に抽出できないことが確認された。

7) 重点課題におけるモデル精緻化と検証

以上の結果を踏まえ、最終段階では、政策的・実務的価値が高く、早期の介入による医療・介護資源の適正化効果が期待される2課題を重点課題として選定し、モデルの精緻化を行った。第一の課題は、大腿骨近位部骨折後の要介護度悪化または死亡の予測であり、第二の課題は、高齢者の高額医療・介護費予測である。

大腿骨近位部骨折後の要介護度悪化または死亡の予測では、整形外科専門医へのヒアリングに基づき、臨床現場で重要と考えられる変数を精緻化した。具体的には、高血圧、心疾患、脳血管疾患、糖尿病等の併存疾患や内服薬の種類を細分化した。さらに、医療プロセスが予後に与える影響を考慮し、術式、入院日数、外来受診回数等の診療実績データを追加した。また、予測モデル実装後に予後改善可能性がある集団を特定することを想定し、解析対象を手術実施日に要支援1から要介護2までの認定を受けていた軽度要介護者に限定した。基準日は手術実施日とし、アウトカムは術後365日時点における要介護度の2段階以上の悪化または死亡に再設定した。しかし、詳細な医療情報を追加しても、精緻化後のAUCは0.67～0.69にとどまり、精緻化前のAUC 0.71～0.72と比較して改善は認められなかった。陽性的中率も十分ではなく、現行のNDB・介護DB連結データのみでは、実用化が推奨される水準の予測性能には至らないと判断した。

高齢者の高額医療・介護費予測では、傷病名、診療行為、薬剤情報等の医療情報を精緻化するとともに、要介護認定情報を追加し、解析対象を要支援認定者に限定した。基準日は初回の認定有効期間の開始日とし、アウトカムは基準月の翌月から12か月間の年間医療・介護費が上位5%となることとして定義した。しかし、感度は11.1～11.8%にとどまり、精緻化前の9.3～10.4%と比較して改善は限定的であった。したがって、高額医療・介護費発生者を事前に抽出するスクリーニングモデルとしては、実用化が推奨される水

準には至らないと判断した。

8) 研究期間全体を通じた成果のまとめ

研究期間全体を通じて、NDB・介護DB等の公的データを用いたAI分析プロトタイプの開発可能性を段階的に検証した。KDBを用いたフィージビリティスタディでは一定の予測精度が得られ、全国規模のNDB・介護DB連結解析の必要性が確認された。一方で、NDB・介護DB連結データを用いた7つの重点研究疑問の解析、および2つの重点課題におけるモデル精緻化では、個別患者の1年程度先の予後や高額医療・介護費発生を、臨床現場での意思決定に直ちに活用できる水準で予測することには課題が残った。この背景には、1年程度先の要介護度悪化、死亡、高額医療・介護費発生は、短期予後と比べて患者の個別性や時間経過中の出来事の影響を受けやすく、予測困難性が高いことがある。加えて、現行のNDB・介護DBには、身体機能、認知機能、栄養状態、生活環境、家族介護力等の情報が十分に含まれていないことも、予測性能を制約した可能性がある。

一方で、本研究により、特定患者群における要介護度悪化、死亡、高額医療・介護費発生の実態を定量的に把握できたことは重要な成果である。公的データベースの強みは、個別患者の長期予後を高精度に予測することよりも、全国規模で患者群の規模、転帰の発生割合、医療・介護サービス利用、費用分布を記述できる点にある。これらの知見は、個別予測AIとして直ちに社会実装するための根拠としては限定的である一方、支援を要する患者群の可視化、ポピュレーション・アプローチの検討、地域差・施設差の把握、政策介入対象の優先順位付けに資する基礎資料として活用可能である。

D. 考察

本研究では、NDBおよび介護DB等の公的データを活用し、医療と介護の特性を総合的に考慮したAI分析プロトタイプの開発可能性を段階的に検討した。研究期間全体を通じて、学際的研究チームの構築、20個の臨床研究疑問の立案、KDBを用いたフィージビリティスタディ、NDB・介護DB連結データによる7つの重点研究疑問の解

析、ならびに2つの重点課題におけるモデル精緻化を実施した。その結果、現行のNDB・介護DB連結データのみを用いて、1年程度先の要介護度悪化、死亡、高額医療・介護費発生を個人単位で臨床現場の意思決定に直ちに活用できる水準で予測することには課題が残ることが明らかとなった。一方で、特定患者群の予後実態や医療・介護費発生の構造を定量的に把握できたことは、医療・介護資源の適正配分や政策介入対象の検討に資する重要な成果である。

KDBを用いたフィージビリティスタディでは、一部の対象集団においてAUC 0.7前後の一定の識別能が得られた。この結果は、医療・介護関連データを用いた機械学習モデル構築が一定程度実行可能であることを示した。一方で、地域限定データでは、希少疾患や特定病態において症例数が不足し、モデル構築や精度評価に制約があることも確認された。また、地域データで得られた結果を全国レベルに一般化するには課題があった。したがって、この段階の成果は、地域データの有用性を示すとともに、全国規模のNDB・介護DB連結解析の必要性を裏付ける予備的知見と位置づけられる。

NDB・介護DB連結データを用いた解析では、20個の臨床研究疑問から7つの重点研究疑問を選定し、複数の機械学習モデルを用いて予測性能を検証した。しかし、多くの課題において、個別患者の1年程度先の予後や高額医療・介護費発生を実用的な水準で予測することには課題が残った。がん・認知症領域では、術後1年または肺炎退院後1年の要介護度悪化や死亡をアウトカムとしたモデルのAUCはおおむね0.60～0.65程度にとどまった。慢性疾患・透析領域では、感度または陽性的中率に課題が残り、ハイリスク者抽出には十分ではなかった。骨折・費用予測領域では、一部でAUCや感度が比較的良好な傾向を示したものの、陽性的中率または感度の低さにより、スクリーニング用途としての実装には課題が残った。

最終段階で実施した2つの重点課題におけるモデル精緻化でも、予測性能の明確な改善は得られなかった。大腿骨近位部骨折後の要介護度悪化または死亡の予測では、整形外科専門医の知見を取り入れ、併存疾患、薬剤、術式、入院日

数、外来受診回数等の変数を詳細化した。また、予測モデル実装後に予後改善可能性がある集団を特定することを想定し、対象集団を手術実施日に要支援1から要介護2までの認定を受けていた軽度要介護者に限定し、アウトカムを術後365日時点における要介護度の2段階以上の悪化または死亡に再設定した。しかし、精緻化後のAUCは0.67～0.69にとどまり、精緻化前のAUC 0.71～0.72と比較して改善は認められなかった。この結果は、大腿骨近位部骨折後の生活機能、要介護度の変化、死亡といった転帰が、レセプト情報や要介護認定情報に含まれる診療内容だけでは十分に説明できない可能性を示している。また、本研究で用いたデータでは、治療内容の詳細や、治療直後の身体状態・生活機能の変化を十分に捉えることが難しく、これらの情報不足が予測性能を制約した可能性がある。

高齢者の高額医療・介護費予測でも、傷病名、診療行為、薬剤情報、要介護認定情報を精緻化し、対象を要支援認定者に限定したが、感度は11.1～11.8%にとどまった。高額医療・介護費発生者の割合に限られる状況では、正解率や特異度が高くても、感度が低ければ、実際に支援を必要とする対象者を十分に抽出できない。そのため、高額費用発生者を事前に把握し、介入対象を選定するスクリーニングモデルとしては、現行モデルの実用化には慎重な検討が必要である。AIモデルの性能評価では、AUCや正解率だけでなく、想定する用途に応じて、感度、特異度、陽性的中率、陰性的中率を組み合わせる必要がある。

本研究で予測精度が十分に向上しなかった背景には、アウトカム自体の予測困難性がある。要介護度悪化、死亡、高額医療・介護費発生といった1年程度先の転帰は、ICU入室中などの短期予後と比べて、患者の個別性や時間経過中の出来事の影響を受けやすい。そのため、仮に生活機能やケアの質に近い情報を追加したとしても、個人単位の長期予後予測の精度が直ちに大幅に改善するとは限らない。加えて、これらの転帰は、疾患や医療利用だけでなく、身体機能、認知機能、栄養状態、フレイル、生活環境、家族介護力、経済状況など、多様な因子に影響される。現行のNDB・介護DBには、これらの情報が十分に含まれていない。また、傷病名情報や介護サービス情報

は月単位で把握される場合があり、急性期イベント後の状態変化や生活機能の変化を細かく捉えることにも制約がある。したがって、現行の公的データベースのみで長期的な個別予後を、臨床現場での意思決定に直ちに活用できる水準で予測することには構造的な課題があると考えられる。

一方で、本研究で得られた記述疫学的知見には重要な意義がある。たとえば、認知症を有するがん患者や肺炎で入院した要介護高齢者において、1年後に要介護度悪化または死亡が一定の割合で発生することを定量的に示したことは、手術適応、退院支援、ACP、ケアプラン見直し等を検討するうえで有用である。また、高額医療・介護費発生の分布を把握することは、医療・介護資源の配分や重点的支援の対象を検討するための基礎資料となる。したがって、本研究成果は、個別予測AIとして直ちに社会実装するための根拠としては限定的である一方、支援を要する患者群の可視化、地域差・施設差の把握、ポピュレーション・アプローチ、政策介入対象の優先順位付けに活用することが現実的である。

公的データを用いたAI・データサイエンスの活用においては、個別予測と集団レベルの政策支援を区別して設計する必要がある。個別患者の意思決定支援に用いる予測モデルでは、見逃しを避けるための感度、不要な介入を減らすための陽性的中率、想定される利用場面に応じた評価指標の妥当性が重要となる。一方、政策立案や資源配分に用いる分析では、個人単位の高精度予測よりも、患者群の規模、イベント発生率、地域差、サービス利用パターン、費用分布を把握することが重要である。本研究の結果は、現行のNDB・介護DB連結データが、個別予測AIの直接実装よりも、集団レベルの実態把握や政策支援において有用性を発揮しやすいことを示している。

今後のデータ基盤整備に向けては、NDB・介護DBに加えて、DPCデータ、科学的介護情報システム（LIFE）、特定健診・後期高齢者健診データ等との連結可能性を検討することが重要である。ただし、これらのデータを追加しても、1年程度先の個別予測精度が直ちに大幅に改善するとは限らない。むしろ、DPCデータ、LIFE、健診デー

タ等を連結する意義は、予測精度の改善だけではなく、生活機能やケアの質に近いアウトカムを設定し、患者群の状態や医療・介護サービス利用の実態をより精緻に記述できる点にもある。DPCデータには、入院時重症度、救急搬送の有無、主要処置、合併症、退院時状態等の急性期臨床経過に関する情報が含まれる。LIFEには、ADL、認知機能、栄養状態、排泄、口腔、褥瘡、リハビリテーション実施状況等、生活機能やケアの質に近い情報が含まれる。これらの情報を連結することで、死亡や要介護度悪化だけでなく、ADL低下、栄養状態の悪化、排泄状態の変化、褥瘡の発生・治癒など、より実践的なアウトカム設定が可能になると考えられる。

ただし、データ項目を追加すれば直ちに予測精度が向上するわけではない。データの測定時点、更新頻度、入力の実質、欠測、施設間差、アウトカム定義の一貫性が確保されなければ、モデル性能や解釈可能性は十分に高まらない可能性がある。したがって、今後は、データベース間の連結拡大に加えて、時点情報の標準化、構造化項目の整備、生活機能に近いアウトカムの定義、データ入力負担と利活用可能性の均衡を検討する必要がある。特に、医療と介護をまたぐ状態変化を捉えるためには、基準日、評価時点、サービス利用時点、転帰発生日の関係を明確に定義できるデータ構造が重要である。

また、AI開発に資するデータ提供体制の整備も重要な課題である。本研究では、NDBおよび介護DBのデータ提供までに長期間を要し、モデル構築後に変数定義、対象集団、アウトカムを見直して再解析する反復的な開発プロセスには制約があった。AIモデルの開発では、初回解析の結果に基づき、特徴量を再設計し、モデルを再学習し、評価指標を再確認する過程が不可欠である。したがって、研究目的に応じて安全に再解析を行えるオンサイトセンター等の解析環境、追加変数や再抽出の審査を効率的に行う仕組み、研究者と行政が迅速に協議できる体制が求められる。

本研究にはいくつかの限界がある。第一に、解析に用いた情報は、主に請求情報および要介護認定情報に基づいており、生活機能、認知機能、栄養状態、住環境、家族介護力等の詳細な情報を十分に扱うことができなかった。第二に、多くの

解析で1年程度先の転帰をアウトカムとしたため、短期転帰に比べて多様な要因の影響を受けやすく、予測の困難度が高かった。第三に、アウトカムによってはイベント不均衡が存在し、正解率や特異度のみではモデルの実用性を判断できなかった。第四に、因果推論を取り入れた新たな解析も計画したが、データ提供時期の関係により、研究期間内に最終的な因果効果の推定には至らなかった。

以上より、本研究は、NDB・介護DB等の公的データを用いたAI分析プロトタイプの開発可能性と課題を明確化した点に意義がある。現行のNDB・介護DB連結データのみでは、個別患者の長期予後や高額医療・介護費発生を、臨床現場での意思決定に直ちに活用できる水準で予測することには課題が残る。一方で、全国規模の公的データを用いて、特定患者群の予後実態、医療・介護サービス利用、費用発生構造を可視化することは、医療・介護資源の適正配分、医療DX、公的データベース整備に向けた政策的検討に資する。今後は、個別予測AIの実装可能性を慎重に評価しつつ、集団レベルの意思決定支援とデータ基盤整備を並行して進める必要がある。

E. 結論

本研究では、NDB、介護DB、KDB等の公的データを活用し、医療と介護の特性を総合的に考慮したAI分析プロトタイプの開発可能性を検討した。研究期間全体を通じて、学際的研究チームの構築、20個の臨床研究疑問の立案、KDBを用いたフィジビリティスタディ、NDB・介護DB連結データを用いた重点研究疑問の解析、ならびに重点課題におけるモデル精緻化を段階的に実施した。その結果、KDBを用いた予備的検討では一部の対象集団において一定の予測性能が得られ、医療・介護関連データを用いた機械学習モデル構築の実行可能性が確認された。一方で、地域限定データでは症例数や一般化可能性に課題があり、全国規模のNDB・介護DB連結データを用いた解析の必要性が示された。

NDB・介護DB連結データを用いた解析では、20個の臨床研究疑問から7つの重点研究疑問を選定し、要介護度悪化、死亡、高額医療・介護費発生等を対象とした予測モデルを構築した。し

かし、多くの課題において、1年程度先の個別予後や高額医療・介護費発生を実用的な水準で予測することには課題が残った。さらに、大腿骨近位部骨折後の要介護度悪化または死亡の予測、および高齢者の高額医療・介護費予測において、専門家の知見を取り入れた変数精緻化や対象集団の再定義を行っても、予測性能の明確な改善は得られなかった。

以上より、現行のNDB・介護DB連結データのみを用いて、個別患者の1年程度先の要介護度悪化、死亡、高額医療・介護費発生を、臨床現場での意思決定に直ちに活用できる水準で予測するAIを直接社会実装することには課題が残ることが明らかとなった。この背景には、これらの転帰が短期予後と比べて患者の個別性や時間経過中の出来事の影響を受けやすく、予測困難性が高いことがある。加えて、要介護度悪化、死亡、高額医療・介護費発生といった転帰は、疾患や医療利用だけでなく、生活機能、認知機能、栄養状態、住環境、家族介護力等の多様な因子に影響される一方で、現行の公的データベースにはこれらの情報が十分に含まれていないことも関係していると考えられる。

一方で、本研究により、特定患者群における要介護度悪化、死亡、高額医療・介護費発生の実態を定量的に把握できたことは重要な成果である。公的データベースの強みは、個別患者の長期予後を高精度に予測することよりも、全国規模で患者群の規模、転帰の発生割合、医療・介護サービス利用、費用分布を記述できる点にある。したがって、これらの知見は、個別予測AIとして直ちに社会実装するための根拠としては限定的である一方、支援を要する患者群の可視化、ポピュレーション・アプローチの検討、地域差・施設差の把握、医療・介護資源配分、政策介入対象の優先順位付けに資する基礎資料として活用可能である。

今後は、NDB・介護DBに加えて、DPCデータ、科学的介護情報システム（LIFE）、特定健診・後期高齢者健診データ等との連結可能性を検討する必要がある。ただし、これらのデータを追加しても、1年程度先の個別予測精度が直ちに大幅に改善するとは限らない。むしろ、ADL、認知機能、栄養状態、排泄、口腔、褥瘡、リハビリテーショ

ン実施状況等の生活機能やケアの質に近い情報を含めることで、より実践的なアウトカム設定と、集団レベルでの実態把握や意思決定支援が可能になると考えられる。また、時点情報の標準化、構造化項目の整備、データ提供体制の迅速化、研究目的に応じて安全に再解析できる環境整備を進めることが重要である。

本研究の成果は、公的データを用いたAI・データサイエンスの社会実装において、個別予測AIの課題と、集団レベルの政策支援に向けた活用可能性を明確にした点に意義がある。今後は、個別予測の精度向上のみを目指すのではなく、医療DX、全国医療情報プラットフォーム、公的データベース連結の推進とあわせて、医療・介護資源の適正配分に資するデータ利活用基盤を整備していく必要がある。

F. 健康危険情報

該当なし

G. 研究発表

1. 論文発表

1) Sakano T, Hirooka K, Kanno Y, Itoh S, Sagawa M, Nitta S, Fujita J, Umezu C, Fukui S. Home-visit nursing service use by older adults in Japan using insurance claims data. *Geriatric Nursing*. 2025;67:103711.

2) Hirooka K, Kanno Y, Itoh S, Sagawa M, Sakano T, Takahashi K, Anzai T, Fukui S. Evaluation of claims-based frailty measurements in older patients with cancer: a retrospective cohort study. *Age and Ageing*. 2025;54(2):afaf015.

3) Sagawa M, Noguchi-Watanabe M, Fukui S. Patterns of long-term care service use and multimorbidity in older Japanese adults: a cross-sectional study of claim data. *Geriatrics & Gerontology International*. 2024;24(3):283-289.

4) Itoh S, Sakano T, Sagawa M, Yanagiya R, Kanno Y, Hirooka K, Fukui S. Home Care Service in the Last Year of Life by Type of Care Management. *Studies in Health Technology and Informatics*. 2024;315:685-686.

5) Sakano T, Anzai T, Takahashi K, Fukui S. Impact of home-visit nursing service use on costs in the last three months of life among older adults. *Journal of Nursing Scholarship*. 2024;56(1):191-201.

2. 学会発表

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

3. その他

該当なし