

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7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。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目指した。

2025年度は、前年度までに実施したNDB・介護DB連結データを用いた機械学習モデルの構築と検証結果を踏まえ、政策的・実務的価値が高く、早期の介入による資源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。具体的には、第一に、大腿骨近位部骨折後の要介護度悪化または死亡の予測、第二に、高齢者の高額医療・介護費予測を対象とした。これらの課題について、臨床専門家の知見を取り入れ、併存疾患、薬剤、術式、入院日数、外来受診回数、診療行為、要介護認定情報等の予測変数を精緻化し、対象集団およびアウトカム定義の再検討を行った。そのうえで、Random Forest、eXtreme Gradient Boosting(XGBoost)等の機械学習モデルを用いて、予測精度の改善可能性を検証した。

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、股関節手術を受けた要支援1から要介護2までの高齢者217,482名を解析対象とした。基準日は手術実施日とし、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を予測対象とした。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併存疾患、内服薬、術式、診療プロセスに関する情報を追加した。しかし、詳細な医療情報を追加しても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。また、陽性的中率は63.1～65.1%であり、陽性と判定された対象者の一定割合には実際のアウトカムが認められな

かった。そのため、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、新規に要支援認定を受けた65歳以上の者1,243,877名を解析対象とした。基準日は初回の認定有効期間の開始日とし、基準月の翌月から12か月間の年間医療・介護費が上位5%となることをアウトカムとして定義した。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た。年間医療・介護費の分布では、医療・介護費が一部の対象者に集中する傾向がみられ、年間医療・介護費が上位5%の対象者が、年間医療・介護費総額の約24.7%を占めていた。一方、機械学習モデルの予測性能では、AUCはおおむね0.76であったが、感度は11.1～11.8%にとどまり、精緻化前の9.3～10.4%と比較してわずかな改善に限られた。したがって、高額医療・介護費発生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と判断した。

以上より、2025年度の検討では、専門家の知見を取り入れた変数の追加、対象集団の再定義、アウトカム設定の見直しを行っても、NDB・介護DB連結データのみを用いた個別予後予測および高額費用予測には課題が残ることが明らかとなった。この背景には、1年程度先の転帰が患者の身体状態、生活機能、認知機能、栄養状態、家族介護力、社会環境など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が関係していると考えられる。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、特定の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。現時点では、個別予測AIを臨床現場へ直接実装するには慎重な検討が必要であり、まずは記述疫学的知見に基づいて、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段階的に活用することが現実的である。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し、生活機能やケアの質に近いアウトカムを含めた公的データの利活用基盤を整備する必要がある。

A. 研究目的

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ものである。

前年度までの研究では、NDBおよび介護DBの利用申出、統計解析計画の精緻化、KDBを用いたフィージビリティスタディ、ならびにNDB・介護DB連結データを用いた複数の臨床研究疑問に対する機械学習モデルの構築を進めてきた。その過程で、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つの重点課題を選定し、予後予測モデルの性能を検証した。しかし、個別の予後予測においては、AUC、感度、特異度、陽性的中率等の観点から、臨床現場での意思決定支援や高リスク者抽出に直ちに活用できる水準の精度を得ることは困難であることが明らかとなった。

2025年度は、これらの結果を踏まえ、政策的・実務的価値が高く、かつ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に焦点を当てた。具体的には、大腿骨近位部骨折後の要介護度悪化または死亡の予測、および高齢者の高額医療・介護費予測の2課題を重点課題として選定した。これらの課題に対して、臨床専門家の知見を取り入れ、併存疾患、薬剤、術式、診療プロセス、要介護認定情報等の予測変数を精緻化し、対象集団およびアウトカム定義の再検討を行った。

本年度の目的は、第一に、NDB・介護DB連結データを用いた機械学習モデルにおいて、臨床的に重要な変数の追加や対象集団の再定義により予測精度が改善するかを検証することである。第二に、大腿骨近位部骨折後の要介護度悪化または死

亡、および高額医療・介護費発生という政策的に重要な転帰について、個別予測AIとしての実装可能性を評価することである。第三に、現行の公的データベースを用いたAI分析の課題を整理し、今後の医療DX、公的データベース連結、生活機能情報を含むデータ基盤整備に向けた示唆を得ることである。

B. 研究方法

本研究では、NDBおよび介護DBを連結した公的データを用いて、医療と介護の特性を総合的に考慮した機械学習モデルを構築し、その予測性能を評価した。研究デザインは、既存の匿名化された公的データベースを用いた後ろ向き観察研究である。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、介護DBからは要介護認定情報および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。

前年度までに、本研究班では医療・介護現場における実務上の課題を踏まえ、20個の臨床研究疑問を立案した。その後、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、7つの重点研究疑問に整理した。2025年度は、これら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課題として、(1)大腿骨近位部骨折後の要介護度悪化または死亡の予測、(2)高齢者の高額医療・介護費予測、の2課題を重点課題として選定した。

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を行い、臨床現場で重要と考えられる予測変数を精緻化した。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を細分化し、内服薬の種類も詳細化した。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に関する情報を予測変数として追加した。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、対象集団は、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とした。

高齢者の高額医療・介護費予測では、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、医療情報および介護情報の精緻化を行った。具体的には、傷病名、診療行為、薬剤情報等の医療情報を詳細化するとともに、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。

各重点課題では、基準日前の傷病名、薬剤、診療行為、要介護認定情報、介護サービス利用情報等から説明変数を作成した。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設定し、機械学習モデルにより予測を行った。予測モデルには、Random Forest、eXtreme Gradient Boosting (XGBoost) 等の複数の手法を用いた。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。特に、ハイリスク者抽出やスクリーニングへの活用可能性を検討するため、AUCや正解率のみではなく、感度および陽性的中率を含めて総合的に評価した。

解析にあたっては、請求情報由来の傷病名の解釈、死亡を含めた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、イベント不均衡下での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定されるアウトカムでは、正解率が高くても感度が低い場合があるため、実用化可能性の判断では、目的に応じた評価指標の妥当性を重視した。

倫理面への配慮として、本研究は、匿名化されたNDBおよび介護DB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さらに低減するため、詳細な地域情報を使用しないこととした。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

本年度は、前年度までに実施したNDB・介護DB連結データを用いた機械学習モデルの構築と検証結果を踏まえ、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。具体的には、(1)大腿骨近位部骨折後の要介護度悪化または死亡の予測、(2)高齢者の高額医療・介護費予測、の2課題である。これらの課題について、臨床専門家の知見を反映した予測変数の追加、対象集団の再定義、アウトカム設定の見直しを行い、機械学習モデルの予測性能が改善するかを検証した。

1) 大腿骨近位部骨折後の要介護度悪化または死亡の予測

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、解析対象者217,482名を対象とした。整形外科専門医へのヒアリングを踏まえ、臨床現場で予後に関連すると考えられる情報を追加した。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を細分化し、内服薬の種類を詳細化した。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に関する情報を予測変数として追加した。加えて、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として再設定した。

術後365日時点の状態をみると、術前要支援1では維持が21.8%、要介護度悪化が68.1%、死亡が10.1%であった。術前要支援2では、改善が2.2%、維持が27.4%、要介護度悪化が59.7%、死亡が10.7%であった。術前要介護1では、改善が1.6%、維持が29.3%、要介護度悪化が53.9%、死亡が15.2%であった。術前要介護2では、改善が3.6%、維持が32.6%、要介護度悪化が45.6%、死亡が18.1%であった(表1)。この結果から、術前の要介護度にかかわらず、術後365日時点で要介護度悪化

または死亡が一定割合で発生していることが確認された。

これらの精緻化を行ったうえで、Random Forest、eXtreme Gradient Boosting(XGBoost)等の機械学習モデルを用いて予測性能を検証した。しかし、詳細な医療情報や診療プロセス情報を追加しても、モデル性能の明確な改善は認められなかった。Generalized Linear ModelのAUCは0.690、Random Forestは0.677、Neural Networkは0.683、XGBoostは0.687、Stochastic Gradient Boostingは0.683であり、いずれのモデルもAUCは0.67～0.69程度にとどまった(表2)。また、感度は71.5～78.6%であった一方、陽性的中率は63.1～65.1%であり、陽性と判定された対象者の一定割合には実際のアウトカムが認められなかった。したがって、大腿骨近位部骨折後の要介護度悪化または死亡については、現行のNDB・介護DB連結データに詳細な医療情報を追加しても、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。

2) 高齢者の高額医療・介護費予測

高齢者の高額医療・介護費予測では、解析対象者1,243,877名を対象とした。介護予防や早期支援の観点から、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、予測モデルの精緻化を行った。具体的には、傷病名、診療行為、薬剤情報等の医療情報を詳細化するとともに、要介護認定情報を予測変数として追加した。解析対象は、介護予防や早期介入の対象として想定される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。

年間医療・介護費の分布をみると、累積患者割合50%時点の年間医療・介護費は874,170円であり、累積年間医療・介護費割合は14.6%であった。累積患者割合75%時点では、年間医療・介護費は1,931,250円、累積年間医療・介護費割合は34.6%であった。累積患者割合90%時点では、年間医療・介護費は4,035,142円、累積年間医療・介護費割合は60.5%であった。さらに、累積患者割合95%時点では、年間医療・介護費は5,632,444円、累積年間医療・介護費割合は75.3%であった(表3)。この結果から、基準月の翌月から12か月間の医療・介護

費は一部の対象者に集中していることが確認された。特に、年間医療・介護費が上位5%の対象者が、年間医療・介護費総額の約24.7%を占めていた。

モデル精緻化後の予測性能を検証した結果、高額医療・介護費発生者の抽出性能は十分ではなかった。Generalized Linear ModelのAUCは0.766、Neural Networkは0.738、XGBoostは0.765、Stochastic Gradient Boostingは0.762であり、AUCは一定程度の値を示した。一方で、感度はGeneralized Linear Modelで11.1%、XGBoostで11.8%、Stochastic Gradient Boostingで11.3%にとどまった(表4)。Neural Networkでは感度が0.0%であり、陽性的中率は算出不能であった。特異度や正解率が高い場合であっても、感度が低い場合には、支援を必要とする高額費用発生者を十分に抽出できない。そのため、本モデルは、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。

3) 2025年度の結果のまとめ

本年度の解析では、2つの重点課題のいずれにおいても、臨床専門家の知見を取り入れた予測変数の精緻化、対象集団の限定、アウトカム定義の再設定を行った。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、解析対象者217,482名を対象に、術後365日時点の要介護度推移および死亡を把握した。その結果、術前要支援1から要介護2までのいずれの群においても、要介護度悪化または死亡が一定割合で発生していた(表1)。しかし、機械学習モデルのAUCは0.67～0.69程度にとどまり、精緻化前と比較して改善は認められなかった(表2)。

高齢者の高額医療・介護費予測では、解析対象者1,243,877名を対象に、基準月の翌月から12か月間の医療・介護費分布を確認した。その結果、医療・介護費は一部の対象者に集中する傾向がみられ、累積患者割合95%時点で累積年間医療・介護費割合は75.3%であった。すなわち、年間医療・介護費が上位5%の対象者が、年間医療・介護費総額の約24.7%を占めていた(表3)。一方で、予測モデルではAUCが約0.76であったものの、感度は11.1～11.8%にとどまり、高額医療・介護費発生者を

抽出するスクリーニングモデルとしては十分ではなかった(表4)。

以上より、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高精度に予測することには課題が残ると考えられた。一方で、特定の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。これらの結果は、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、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基礎資料として活用可能である。

D. 考察

本年度は、前年度までのNDB・介護DB連結データを用いた機械学習モデルの構築と検証結果を踏まえ、政策的・実務的価値が高い2課題を重点課題として選定し、予測モデルの精緻化を行った。具体的には、大腿骨近位部骨折後の要介護度悪化または死亡の予測と、高齢者の高額医療・介護費予測を対象とし、臨床専門家の知見を取り入れて、併存疾患、薬剤、術式、診療プロセス、要介護認定情報等の予測変数を詳細化した。また、対象集団を再定義し、アウトカム設定も見直した。しかし、いずれの課題においても、実用化が推奨される水準の予測精度には至らないと判断した。この結果は、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高精度に予測することには課題が残ることを示している。

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を踏まえ、臨床的に重要と考えられる変数を追加した。併存疾患や薬剤情報を細分化し、術式、入院日数、外来受診回数等の診療プロセスに関する情報も取り入れた。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。それにもかかわらず、AUCは精緻化前より改善しなかった。このこ

とは、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。術後の生活機能は、受傷前のフレイル、歩行能力、栄養状態、認知機能、リハビリテーションの実施状況、住環境、家族介護力など、多面的な要因に左右される可能性が高い。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、要介護認定情報を詳細化し、解析対象を要支援認定者に限定した。しかし、感度は11.1～11.8%にとどまり、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十分ではなかった。高額費用の発生は、基礎疾患や既存の医療利用だけでなく、急性増悪、骨折、感染症、入院、施設入所、介護サービス利用の変化など、複数の要因が時間的に連鎖して生じる。そのため、基準日前の請求情報を中心とした説明変数だけでは、1年後の高額費用発生を十分に識別することが難しかったと考えられる。また、イベント割合が低いアウトカムでは、正解率や特異度が高く見えても、感度が低ければ支援対象者の抽出にはつながりにくい。一方で、陽性的中率は、陽性と判定された対象者のうち実際に高額医療・介護費が発生した者の割合を示すため、介入対象者を選定する際の実用性を判断するうえで重要である。したがって、AIモデルの実用性を評価する際には、AUCや正解率だけでなく、利用目的に応じて感度、陽性的中率、陰性的中率を含めて判断する必要がある。

本研究で予測精度が十分に改善しなかった背景には、公的データベースの構造的な制約も関係している。NDBおよび介護DBは、全国規模で医療・介護サービスの利用実態を把握できる点で大きな強みを有する。一方で、これらのデータは主として請求情報および要介護認定情報に基づくため、臨床検査値、身体機能、認知機能、栄養状態、生活環境、家族介護力など、予後やサービス利用に影響する重要な情報を十分に含んでいない。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変化や状態悪

化の時点を細かく捉えることには制約がある。これらの制約は、長期予後や高額費用発生の個別予測において、モデル性能の上限を規定する要因になり得る。

以上を踏まえると、現時点では、NDB・介護DB連結データのみを用いた個別予測AIを、臨床現場の意思決定支援ツールとして直ちに実装するには課題が残る。特に、手術適応、退院支援、ACP、ケアプランの見直しなど、個別患者の意思決定に関わる場面では、予測モデルの感度や陽性的中率が十分でない場合、支援対象者の抽出に課題が残る。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。特定の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示すことは、個別予測AIの直接実装とは異なる形で、支援を要する患者群の可視化、地域や施設における支援体制の検討、政策介入対象の優先順位付けに活用できる。

今後の公的データ活用においては、個別予測AIの開発だけを目的とするのではなく、データ特性に適した活用目的を設定することが重要である。NDB・介護DBは、患者群の層別化、地域差・施設差の可視化、医療・介護サービス利用パターンの抽出、高額費用群や急性増悪群の候補把握、政策介入対象の検討において有用性を発揮する可能性がある。これらの活用では、個人単位で高精度な予測を行うことよりも、集団レベルでリスクやサービス利用の分布を把握し、介入の必要性が高い領域を明らかにすることが重視される。本研究の結果は、公的データベースを用いたAI・データサイエンスの社会実装において、個別予測と集団レベルの政策支援を区別して設計する必要性を示している。

また、医療DXおよび全国医療情報プラットフォームの推進に向けては、NDB・介護DBに加え、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。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可能性がある。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、

褥瘡の発生・治癒、排泄状態の変化など、よりケアの質に近いアウトカムを設定できる可能性がある。ただし、データを追加すれば直ちに予測精度が向上するとは限らないため、時点情報の標準化、アウトカム定義の精緻化、データ入力の質の確保を並行して進める必要がある。

さらに、AI開発に資するデータ提供体制の整備も重要である。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、モデル構築後の変数再定義、再抽出、再解析といった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いて変数を見直し、対象集団やアウトカムを再定義し、再度モデルを評価する過程が不可欠である。そのため、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や、追加変数・再抽出の審査を効率的に進める仕組みが求められる。これは、施策に直結するエビデンスをタイムリーに創出するためにも重要である。

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NDB・介護DBに含まれる請求情報および要介護認定情報に限られており、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、1年程度先の要介護度悪化、死亡、高額医療・介護費発生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実用性を判断できなかった。これらの限界を踏まえ、今後は、より短期の転帰、生活機能に近いアウトカム、DPC・LIFE・健診データ等を含む多層的なデータ連結を用いた検証が必要である。

以上より、本年度の研究は、NDB・介護DB連結データを用いた個別予測AIの課題を明確化するとともに、公的データの現実的な活用方向を示した点に意義がある。個別患者の長期予後を高精度に予測するには、現行データのみでは十分ではない可能性が高い。一方で、全国規模の公的データを用いて、支援を要する患者群の特徴や医療・介護費の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の推進、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および介護DBを連結した公的データを用いて、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。2025年度は、大腿骨近位部骨折後の要介護度悪化または死亡の予測、および高齢者の高額医療・介護費予測を重点課題として選定し、臨床専門家の知見を取り入れて、予測変数、対象集団、アウトカム定義を精緻化した。その結果、いずれの課題においても、実用化が推奨される水準の予測精度には至らないと判断した。

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、詳細な医療情報や診療プロセス情報を追加してもAUCの改善は認められなかった。また、高齢者の高額医療・介護費予測では、感度の改善は限定的であり、高額費用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十分ではなかった。

以上より、現行のNDB・介護DB連結データのみを用いて、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高精度に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、特定の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、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重要な成果である。

今後は、個別予測AIの直接的な社会実装について慎重に検討しつつ、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、データ項目の追加のみで予測精度が大幅に改善するとは限らないため、時点情報の標準化、データ品質、アウトカム定義の精緻化を並行して進め

る必要がある。さらに、生活機能やケアの質に近いアウトカムを設定し、時点情報の標準化とデータ提供体制の整備を進めることで、医療DXおよび公的データ利活用基盤の充実に資する研究を継続する必要がある。

本研究の成果は、医療介護分野における機械学習活用の可能性と課題を示すものであり、今後の公的データ利活用基盤の整備に向けた基礎資料となると考えられる。

F. 健康危険情報

該当なし

G. 研究発表

- 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. *Geriatr Nurs*. 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 Ageing*. 2025; 54(2): afaf015.

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